The Journal of Immunology

The Evolution and Regulation of the Mucosal Immune Complexity in the Basal Chordate Amphioxus

Shengfeng Huang, Xin Wang, Qingyu Yan, Lei Guo, Shaochun Yuan, Guangrui Huang, Huiqing Huang, Jun Li, Meiling Dong, Shangwu Chen and Anlong Xu

J Immunol 2011; 186:2042-2055; Prepublished online 19 January 2011;
doi: 10.4049/jimmunol.1001824
http://www.jimmunol.org/content/186/4/2042

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

The Evolution and Regulation of the Mucosal Immune Complexity in the Basal Chordate Amphioxus

Shengfeng Huang,[1] Xin Wang,[1] Qingyu Yan, Lei Guo, Shaochun Yuan, Guangrui Huang, Huiqing Huang, Jun Li, Meiling Dong, Shangwu Chen, and Anlong Xu

Both amphioxus and the sea urchin encode a complex innate immune gene repertoire in their genomes, but the composition and mechanisms of their innate immune systems, as well as the fundamental differences between two systems, remain largely unexplored. In this study, we dissect the mucosal immune complexity of amphioxus into different evolutionary-functional modes and regulatory patterns by integrating information from phylogenetic inferences, genome-wide digital expression profiles, time course expression dynamics, and functional analyses. With these rich data, we reconstruct several major immune subsystems in amphioxus and analyze their regulation during mucosal infection. These include the TNF/IL-1R network, TLR and NLR networks, complement system, apoptosis network, oxidative pathways, and other effector genes (e.g., peptidoglycan recognition proteins, Gram-negative binding proteins, and chitin-binding proteins). We show that beneath the superficial similarity to that of the sea urchin, the amphioxus innate system, despite preserving critical invertebrate components, is more similar to that of the vertebrates in terms of composition, expression regulation, and functional strategies. For example, major effectors in amphioxus gut mucous tissue are the well-developed complement and oxidative-burst systems, and the signaling network in amphioxus seems to emphasize signal transduction/modulation more than initiation. In conclusion, we suggest that the innate immune systems of amphioxus and the sea urchin are strategically different, possibly representing two successful cases among many expanded immune systems that arose at the age of the Cambrian explosion. We further suggest that the vertebrate innate immune system should be derived from one of these expanded systems, most likely from the same one that was shared by amphioxus. ***The Journal of Immunology*****, 2011, 186: 2042–2055.**

Metazoans require an effective immune system to either suppress hostile microbes or maintain a beneficial microbial flora, but launching an immune response comes at a cost, either consuming extra resources/energy or causing damage to tissues. Thus, between killing and maintaining, and between efficiency and cost, metazoans are driven to evolve immune mechanisms and strategies to achieve a dynamic balance. Toward this goal, vertebrates developed both innate and adaptive immunity (1, 2). Innate immunity provides recognition of relatively invariable molecular features present in microbes of certain classes through a limited number of germ line-encoded pathogen-associated molecular pattern recognition receptors (PRRs; e.g., TLR and NACHT-leucine-rich repeat receptor [NLR]) (3). Adaptive immunity offers nearly unlimited recognition capacity by generating somatically diversified Ag receptors (BCR, Ig, TCR, and variable lymphocyte receptor [VLR]) for specific molecular features in microbes, which, together with the mechanism of selective clonal expression and expansion, forms the basis of immune memory (4). Innate and adaptive immunity interweave on different levels (5, 6). On the signal initiation level, T/B cell activation and proliferation often require multiple signal input from both innate and adaptive receptors (7). On the effector level, the complement system, as a major innate effector, incorporates Abs (Igs) as its primary sensors and elicitors through the C1 complex (8). Insects such as *Drosophila* have only innate immunity but exploit it in a manner different from vertebrates (9). A hallmark of insect immunity is the systemic immune response, which relies on peptidoglycan recognition proteins (PGRPs) and Gram-negative binding proteins (GNBPs) for recognition, uses Toll, Imd, and JAK/STAT pathways for signaling, and activates the secretion of large amounts of antimicrobial peptides into the hemolymph for microbial clearance (9). Another hallmark of insect immunity is prophenoloxidase-mediated melanization (9). Genome sequences of transitional species between protostomes and vertebrates, such as amphioxus and the sea urchin, reveal that they possess neither the vertebrate-type adaptive immunity nor the insect-style innate

State Key Laboratory of Biocontrol, Guangdong Key Laboratory of Pharmaceutical Functional Genes, College of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China

[1]S.H and X.W. contributed equally to this work.

Received for publication June 17, 2010. Accepted for publication December 1, 2010.

This work was supported by Projects 2007CB815800 (973), 2008AA092601 (863), and 2007DFA30840 (International S&T Cooperation Program) from the Ministry of Science and Technology of China; a Key Project (0107) from the Ministry of Education; Project 30901103 from the National Natural Science Foundation of China; and projects from the Commission of Science and Technology of Guangdong Province and Guangzhou City and the Sun Yet-sen Univeristy Science Foundation. A.X. is a recipient of an Outstanding Young Scientist award from the National Nature Science Foundation of China.

Address correspondence and reprint requests to Dr. Anlong Xu, Department of Molecular Biology and Immunology, State Key Laboratory of Biocontrol, College of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China. E-mail address: lssxal@mail.sysu.edu.cn

The online version of this article contains supplemental material.

Abbreviations used in this article: BC, bacterially challenged; BF, *Branchiostoma floridae* embryonic; bjPGRP1, peptidoglycan recognition protein 1 from *Branchiostoma japonicum*; CBP, chitin-binding protein; CCP, complement control protein; CLR, C-type lectin receptor; C_T, cycle threshold; DAMP, damage-associated molecular pattern; DAP, diaminopimelic acid; DFD, death-fold domain; EST, expressed sequence tag; GNBP, Gram-negative bacteria-binding protein; IRF, IFN regulatory factor; LRR, leucine-rich repeat; LRRIG, leucine-rich repeat-Ig protein; LTA, lipoteichoic acid; Lys, L-lysine; MACPF, membrane attack complex/perforin; MASP, mannose-binding lectin-associated serine protease; ME, minimum evolution; MIX, bacterial mixture; NACHT, NTPase domain named after NAIP, CIITA, HET-E, and TP1; NCF, neutrophil cytosol factor; NLR, NACHT-leucine-rich repeat receptor; PGN, peptidoglycan; PGRP, peptidoglycan recognition protein; PRR, pattern recognition receptor; qPCR, quantitative PCR; RLR, retinoic acid-inducible gene I-like helicase receptor; ROS, reactive oxygen species; SRCR, scavenger receptor with cysteine-rich repeat; TIR, Toll/IL-1R resistance; TPO, thyroid peroxidase; TRAF, TNFR-associated factor; UC, unchallenged; VCBP, V region containing chitin-binding protein; VLR, variable lymphocyte receptor; VTLR, vertebrate-type TLR.

immunity; instead, they develop greatly expanded innate gene repertoires (10–12). For example, the family sizes of their TLRs, NLRs, scavenger receptors with cysteine-rich repeats (SRCRs), and C-type lectin receptors (CLRs) are 5–20 times those of vertebrates and insects. Hence, it has been proposed that the deuterostome immune system was primitively elaborate (13). However, it remains elusive about the reasons why ancestral deuterostomes produced such great innate complexity, why the modern vertebrate immunity abandoned such innate complexity, and how the evolutionary transition between two types of immune systems was realized.

Despite the superficial similarities, we note that the innate repertoires of amphioxus and the sea urchin are different in several aspects: 1) the sea urchin exhibits great expansion of TLRs, NLRs, and SRCRs, whereas amphioxus shows its greatest increase in leucine-rich repeat-Ig proteins (LRRIGs), SRCRs, CLRs, and fibrinogen genes but only slight to medium expansion of TLRs and NLRs; 2) amphioxus shows a greater expansion of complement control proteins (CCPs), caspases, Toll/IL-1R resistance (TIR) adaptors, and death-fold domain (DFD) genes than does the sea urchin; and 3) of other important gene families in vertebrate immunity, such as TNFs/TNFRs, TNFR-associated factors (TRAFs), IFN regulatory factors (IRFs), and mannose-binding lectin-associated serine proteases (MASPs), amphioxus has an equal number or more as compared with vertebrates, whereas the sea urchin has much fewer. So far, the mechanisms and evolutionary differences underlying these differences are unexplored, and since amphioxus represents basal chordates, it is relevant to ask whether these differences represent amphioxus specialties or primitive chordate features. On the other hand, many important facets of amphioxus innate immunity are not understood on a genome-wide scale, including PGRPs, GNBPs, transcription factors, oxidative pathways, and other effectors. Moreover, the innate immune complexity of amphioxus, which was inferred from a draft genome sequence, has not been confirmed at the transcriptome level. Even if such complexity can be demonstrated, there is little available information about its organization and regulation during immune responses.

Addressing these questions will further our understanding of the ancient deuterostome immune complexity. The genome sequence of the amphioxus *Branchiostoma floridae* has been proved to be a good starting point (10), but it cannot alone offer further insight. In this study, we attempt to draw information from dynamic transcriptomes. We used massively parallel mRNA sequencing to monitor the genome-wide expression profile of the amphioxus *Branchiostoma belcheri* intestine in response to bacterial challenge. Based on the transcriptomic analysis, we selected >200 candidate genes for real-time quantitative RT-PCR to monitor their time course modulation in response to three types of bacterial challenges (LPS only, lipoteichoic acid [LTA] only, and a mixture of LPS, LTA, and Gram-positive and Gram-negative bacteria). We show that these expression data, together with those from genome sequences and functional data, when integrated into an evolutionary framework, provide novel insights into the inner working of the amphioxus immune system and the origin of modern vertebrate immunity.

Materials and Methods

Gene annotation and phylogenetic analyses

The haploid draft genome of *B. floridae* and the corresponding predicted transcripts were downloaded from the Joint Genome Institute Web site (http://genome.jgi-psf.org/Brafl1/Brafl1.home.html). Gene annotation, protein domain search, and protein-based phylogenetic reconstruction were performed as described in our previous genomic survey (10). Three molecular tree building methods were used. Minimum-evolution (ME) trees were built by using Mega v4.1, with 1000 bootstrap tests and handling gaps by pairwise deletion. The maximum parsimony method was conducted by using Phylip v3.65, with 100 bootstrap tests and 10 times of jumble. Bayesian trees were carried out using MrBayes v3.1, with a mixed fixed-rate amino acid substitution model. These methods provide similar results, so in this study only bootstrapped consensus ME trees are presented.

Bacterial infection, cDNA preparation, and pyrosequencing

The procedure of bacterial challenge has been described previously (14, 15). In this study, we made two changes: 1) instead of a single stimulant, a mixture of stimulants was injected into the intestine (15 μl/animal), which included live *Vibrio vulnificus* (Gram-negative, 5×10^7 cells/ml), live *Staphylococcus aureus* (Gram-positive, 5×10^7 cells/ml), bacterial cell wall component LPS (from *Escherichia coli*; Sigma-Aldrich, 1 mg/ml) and LTA (from *S. aureus*; Sigma-Aldrich, 1 mg/ml); 2) naive (unchallenged) animals were used to represent the inactivated state, instead of animals injected with PBS. To avoid errors introduced by injection procedure, we conducted the injection under the microscope and compared the procedures that led to severe injury (defined by faster injection speed and visible wound) or milder injury (slow injection and no visible wound). We found that the procedure causing severe injury generally caused dramatic but inconsistent expression changes. Accordingly, we chose the milder way of injection. Six hours later, intestines from 30 bacterial-challenged animals or 30 naive animals were harvested and processed to bacterially challenged (BC) and unchallenged (UC) cDNA libraries. Approximately 5 μg cDNA of BC and UC samples was used for nitrogen nebulization and subsequent massive parallel pyrosequencing (16). Obtained expressed sequence tags (ESTs) were deposited in National Center for Biotechnology Information (NCBI) Sequence Read Archive database (http://www.ncbi.nlm.nih.gov/) under accession number SRA001002.

Mapping of EST to the genome of B. floridae

All EST sequences were mapped to predicted gene models of *B. floridae* by using NCBI BLASTN with filtering and mismatch penalty set to −1. For BC and UC libraries, EST sequences that hit to predicted transcripts with at least 50 bp alignment length and an E value of $<1 \times 10^{-5}$ were considered as successful mappings, which provided an average of 89% nucleotide identity and at least 80% translated protein sequence identity. EST sequences were also mapped to the genome sequence by using NCBI BLASTN with filtering and mismatch penalty set to −1. Additionally, ~269,739 (after quality control) *B. floridae* EST sequences derived from different stages of embryos were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/) and used for comparison.

Digital expression profiling

Because the sequencing depth in this experiment was in the range of 160,000–270,000 transcripts per cDNA library, the digitalized expression level was normalized to 100,000 transcripts per library, which is easy to compare with the widely used digital expression unit of transcripts per million. The statistical test of the significance for expression level change was performed using the software IDEG6, including the Audic and Claverie method (17). The Gene Ontology analysis was performed using in-house script.

Time course expression analysis

Immune challenges and tissue harvest were performed as described above, but the injection was increased to 25 μl per animal and three different treatments were attempted: LPS only (4 mg/ml, from *E. coli* B4; Sigma-Aldrich), LTA only (4 mg/ml, from *S. aureus*; Sigma-Aldrich), and a mixture of live *E. coli* (5×10^7 cells/ml), live *S. aureus* (5×10^7 cells/ml), LPS (4 mg/ml), and LTA (4 mg/ml). Noted that in this study we replaced *V. vulnificus* with *E. coli*. The naive state and seven postinfection time points had been sampled (1st, 2nd, 4th, 8th, 12th, and 24th h; however, the 1st h samples of mixture treatment were not available). We collected six adult animals for each sample.

Total RNA was purified from the intestine using a Qiagen RNeasy Plus Mini kit and then treated with Promega DNaseI. Double-stranded cDNA was synthesized from total RNA by using the SYBR Perfect real-time series kits (Takara Bio). The real-time quantitative RT-PCR was performed on Roche's LightCycler 480 real-time PCR system (using the 384-well module). Quantitative PCR (qPCR) testing for each sample was performed in two replicates using a 10-μl reaction system containing 50 ng initial total RNA (internal controls/reference genes were performed in double replicates). Both annealing and extending temperature were set to 60°C. Forty PCR cycles were run and the melting curve was recorded. All other parameters for the reaction system and the PCR program were set according to the manufacturer's protocol of the SYBR PrimeScript RT-PCR kit (Takara Bio). All results were confirmed by repeating the assays by one or two more times.

According to the digital expression analysis, we selected 259 genes for further time course analyses. A total of 361 primer pairs had been designed using the software Beacon Designer 7. The primer searching parameters were set as followed: melting temperature, 57 ± 2°C; primer length, 19–23 bp; amplicon length, 70–250 bp; and 3′ end bias. All other parameters were set as default. Standard curve analysis, melting curve analysis, and electrophoresis analysis were performed to check the quality of the primer pairs, which eliminated the primer pairs with low amplification efficiency, incorrect amplicon length, and nonspecific products. Finally, only 214 primer pairs for 214 individual genes were considered qualified from further analysis. All these primer pairs are present in Supplemental Table IV.

The 18S rRNA is the best internal control for quantitative RT-PCR if applicable, but its abundance often leads to low cycle threshold (C_T) value (<5), especially when we used high starting template concentration (50 ng mRNA per 10 μl reaction), which made the direct use of 18S rRNA infeasible. In this study, we used three genes as internal controls (two primer pairs for 18S rRNA, two primer pairs for β-actin, and one primer pair for GAPDH). We compared the expression level of these genes and found that GAPDH showed the least variation relative to 18S rRNA (Supplemental Fig. 2). Thus, GAPDH was used to normalize the expression levels. The expression level was calculated using the comparative C_T method.

We considered a target gene to have truly undergone significant expression upregulation (downregulation) at a certain time point when the *t* test *p* value was <0.01 and the change was >1.5-fold (>2-fold for downregulation). The rationale to choose 1.5-fold as the minimal requirement for significant upregulation was as follows: More than 99% of the observed C_T value differences between two duplicates were <0.22 cycle, which caused a maximum error of expression level of ~[(2 × 0.22)(2 × 0.22)], or 1.36-fold (variations on both target genes and reference genes were considered). We also observed that the difference of amplification efficiency between target genes and reference genes was <0.1 (to achieve an upper bound threshold, in this study we used 2 and 1.85 as the efficiency for the reference gene and the target gene, respectively). We also observed that the change of template concentration in different samples varied <1.5-fold (to achieve an upper-bound threshold, we used 2-fold). Using these values we calculated that the maximum false upregulation was 1.47-fold.

Therefore, in this study we chose 1.5-fold as the minimal requirement for a significant upregulation change and 2-fold as the more reliable change. As for the downregulation, because the expression of many genes may appear to be "downregulated" due to the upregulation of other genes, even though they were not actually downregulated, we consider 2-fold downregulation as the minimum change and 4-fold as the reliable change.

To compare qPCR expression levels of different genes and the digital expression levels, we arbitrarily set the basal expression level of GAPDH to 100 per 100,000 transcripts in this study and then normalized other qPCR expression levels to this value. The basal expression number for GAPDH is consistent with the estimation from the UC library (91 per 100,000 ESTs), and it is roughly in agreement with reports from other species (800–2000 per million transcripts). Therefore, in this study, the digital expression levels can be roughly compared with the qPCR expression levels. We further estimated that the largest error of these comparisons should be a lot <1 order of magnitude, calculated as follows: In this study the differences between duplicates of the same primer pairs should be <1.36-fold (as calculated above), the observed differences caused by different primer pairs of the same genes were <2.5-fold, and the transcript (and the PCR product) length difference between a target gene and GAPDH was <2-fold, so the product was <1.36 × 2.5 × 2, or 5.8-fold.

However, note that too low or too high C_T values led to higher deviation because those values had fallen out of the linear increasing range.

Finally, the unsupervised expression profile clustering analysis was performed using the dChip software.

Peptidoglycan binding assay and chloroform-induced bacteriolysis assay

The fusion protein His-tagged TRX-bjPGRP1 was incubated with the insoluble peptidoglycan (PGN) from two bacterial species. After incubation and subsequent centrifugation, the unbound protein in supernatant and the bound protein in deposition were collected separately. Both were used for Western blotting. The TRX protein was used as control.

As for chloroform-induced bacteriolysis assay, a construct pGEX-bjPGRP1 was first transformed into the *E. coli* strain BL21(DE3). The expression of fusion protein GST-bjPGRP1 in the cultured *E. coli* was induced by using isopropyl β-D-thiogalactoside. The fresh culture was collected and washed using PBS. Then, 1% chloroform (of total volume) was added to the bacterial suspension. After 10 min, the cell culture was diluted 10-fold and its OD_{600} value was determined by spectrophotometry. The GST protein was used as control. The difference between the experimental and control groups could be easily detected visually. Vector pGEX-4T contains GST as chaperone (purchased from Amersham Biosciences).

Results

Gene annotation and general considerations

We previously identified and curated 5000 candidate immune genes from the diploid genome assembly of *B. floridae* (10). This list contains redundant items because it includes both gene alleles. Because the haploid assembly has been published (18), we updated the list by removing alleles and including 12 new families and >200 new genes. This new list includes >3300 genes (Supplemental Table I) in >50 gene families (Fig. 1*C*). We tentatively classified these genes into four molecular immune functions: "pattern recognition," "effectors," "cytokines and signal transducers," and "kinases and transcription factors." This classification is inevitably subjective and overlapping, as some genes can be assigned to different categories, and the functions of many genes are simply inferred from sequence similarity or other nonexperimental clues. For example, the expanded families LRRIG and

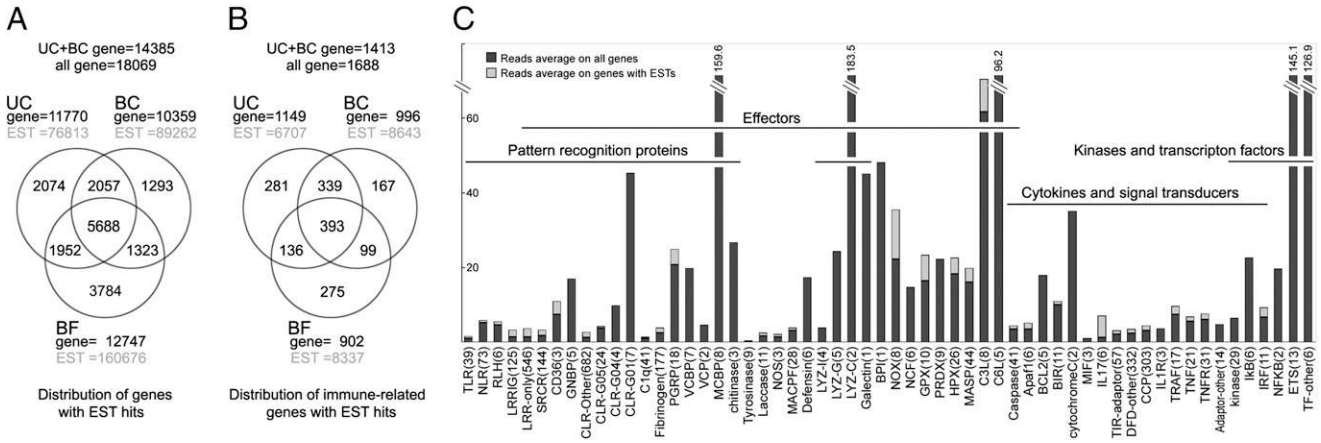

FIGURE 1. Distribution of all genes (*A*) and candidate immune genes (*B*) with EST hits in the UC library of *B. belcheri*, the BC library of *B. belcheri*, and the BF library. *C*, Average EST (UC plus BC) abundance for different immune gene families. Refer to Supplemental Fig. 1 to see the total EST abundance for different gene families. Numbers in parentheses indicate the family size. "Adaptor-other" includes Bam32, Diablo, ECSIT, TAB1/2, Tollip, FADD-like, PEA15, and PIDD-like. "TF-other" includes CEBP, STAT, Jun, Fos, and NFAT. To guarantee the representativeness of the abundance estimation, in the family with >60 members, the average abundance did not take into account the members that contribute >15% of total ESTs.

Table I. Alteration of transcriptome at the sixth hour after bacterial infection

Function Category	Relative Expression in Naive Condition (%)	Alteration after Bacterial Challenge (%)
Translation[a]	17.83	18.16
Generation of precursor metabolites and energy	1.77	10.15
Catabolic process	4.15	17.51
Biosynthetic process (translation excluded)	5.03	−4.19
Secondary metabolic process	0.75	−14.01
Developmental process	25.26	−3.67
Growth	8.26	−4.68
Reproduction	9.13	−4.80
Aging	1.24	−9.30
Digestion	0.96	−34.03
Circulatory system process	1.81	−47.33
Neurologic system process	3.60	−24.86
Cell adhesion	4.47	−22.89
Endocytosis	3.95	−14.75
Cell motility	4.80	−26.03
Cellular homeostasis	1.84	−27.14
Cell division	3.15	−14.34
Cell growth	0.94	−13.56
Cell differentiation	13.72	−5.27
Cell proliferation	6.61	−5.87
Apoptosis	5.51	−1.52
Anti-apoptosis	1.70	15.16
Induction of apoptosis	1.96	−20.09
Inflammatory response	0.64	38.51
Candidate immune genes (not a Gene Ontology term)	8.73	11.5

[a]RNA processing and transcription are not changed

LRR-only have a classical PRR motif (LRRs), but their function has not been explored. Moreover, many proteins containing DFD, CCP, and C-type lectin-like domains may have broad implications either inside or outside the immune system. Notably, this study is designed to compare the naive gut (mucous tissue) transcriptome with that stimulated by bacterial substances. Visible injury was not allowed in our injection of stimulants into the amphioxus intestine, but microinjury was unavoidable (see *Materials and Methods*). Thus, our result is in fact a combination of responses to the immune challenge plus the injury.

mRNA sequencing and cross-species EST mapping

Transcriptomic analyses of human and *Drosophila* antibacterial responses have shown that most genes attain a substantial expression change at the sixth hour (19–21). To monitor the transcriptome change of infected amphioxus, we sequenced the naive, UC transcriptome, and the BC transcriptome of the intestine of amphioxus *B. belcheri* at 6 h postinfection. For the infection, a mixture of LPS, LTA, *V. vulnificus* (Gram-negative), and *S. aureus* (Gram-positive) was used to produce a maximum gene spectrum of expression change. In total, 215,571 and 165,433 high-quality EST reads were obtained from UC and BC cDNA libraries, respectively. Approximately 52% of UC plus BC ESTs were mapped to the genome of *B. floridae* by using NCBI BLASTN (E value of $<1 \times 10^{-5}$ and alignment length of $\geq$50 bp, equivalent to an average 89% nucleotide identity and >80% amino acid identity). Because this is a cross-species EST mapping, we needed to assess whether the mapping was sufficiently accurate to allow digital expression analysis. In this study, we define an ambiguous mapping event as an EST that has identical BLASTN scores for its first-best and second-best hits to the genome, and the sequences (2.5 kb) flanking the genomic loci of the first-best and second-best hits are not duplicates. Analysis of the ESTs with 90–110 bp alignment length showed that only 6% of ESTs exhibited ambiguous mapping, among which <42% were located in genic region (exons and introns). Hence, the level of mapping accuracy permits accurate expression profiling, which is further supported by the consistency between digital and the qPCR analyses (discussed later).

Genome-wide expression activation against infection

We compared the EST mapping results of UC, BC, and *B. floridae* embryonic (BF) libraries (Fig. 1*A*, 1*B*). Approximately 8.7% of UC ESTs and 9.7% of BC ESTs were mapped to immune genes, whereas this was the case with only 5% of BF ESTs, suggesting that the adult gut has more abundant immune transcripts than does the embryo. In both total gene and immune gene comparison, the BC library covered fewer genes with more ESTs when compared with the UC library, indicating that postinfection both global and immune expression profiles were skewed to gene induction rather than gene suppression. The normalized expression levels for all genes and immune genes are presented in Supplemental Fig. 2 and Supplemental Table I. Based on the normalized expression profiles, we used Gene Ontology to evaluate the transcriptome change from the perspective of biological processes (Table I). Following infection, the irrelevant processes and regular systems were suppressed through either active means (e.g., the apoptosis network was suppressed via the increased expression of anti-apoptotic genes) or passive means (e.g., the apoptosis network was suppressed via the decreased expression of proapoptotic genes), whereas the relevant processes, including the immune system, energy, catabolism, and protein synthesis, were upregulated to intensify immune defense.

Transcriptional evidence for innate immune complexity

With >3300 annotated candidate immune genes (Supplemental Table I), amphioxus encodes the most complex innate immune system described thus far, but inference from the draft genome does not necessarily correspond to functional complexity, since assembly errors and the rapid gene turnover rate may produce nonfunctional genes. For instance, only 400 of 1000 human olfactory receptors are functional (22), and 30% of 222 sea urchin TLR genes were classified as pseudogenes (13). In this study, we demonstrated

Table II. EST evidence for expanded gene families of potential immune functions

Gene Family[b]	Gene Number[a] Hs	Sp	Bf	Hit by BF ESTs[c]	Hit by UC or BC ESTs	Hit by Any ESTs
TLR	10	222	39	6	17	22
NLR	≈25	203	73	28	60	63
SRCR	16	218	144	34	68	76
LRRIG	≈30	≈22	125	19	40	48
LRR only	<200	<297	≈546	113	163	195
RLR	3	12	6	1	5	5
PGRP	4	5	18	8	14	15
GNBP	0	3	5	5	5	5
CLR	81	104	717	172	216	295
C1q-like	29	4	41	14	15	27
Fibrinogen	26	<90	177	45	97	113
CCP	53	<247	303	135	175	212
MACPF	12	22	28	14	18	22
MASP-like	7	2	44	26	30	37
Other TIR	5	26	57	15	36	40
TNF	20	4	21	2	17	17
TNFR	26	7	31	13	23	25
TRAF	6	4	17	4	12	13
Caspase	10	31	41	11	29	32
Other DFDs	60	116	332	89	198	218
IRF	9	2	11	3	7	7

[a]ESTs of *B. floridae*, available on GenBank.
[b]Gene numbers in Hs, Sp, and Bf.
[c]LRR only, genes containing only LRR modules; Fibrinogen, fibrinogen domain is the building block of ficolin; Other TIR, TIR genes other than TIR receptors (e.g., TLR and IL-1R), many of which are predicted to be adaptors; Other DFDs, proteins containing DFD other than NLRs, RLHs, Apaf-like, caspases, and DRs.
Bf, amphioxus *B. floridae*; Hs, human; Sp, sea urchin.

that a substantial portion of the annotated genes, especially those expanded, can be processed into mature mRNA (Table II). Intestinal libraries (UC and BC) provided expression evidence for more immune genes than did BF libraries, suggesting the adult gut as the region of intense immune activity. However, the number of expressible genes of potential immune functions shown in Table II is an underestimate due to the limited sample size and the divergence between two amphioxus species. Nevertheless, the current EST evidence has sufficiently demonstrated that the innate immune system of amphioxus is among the most complex known.

Differential regulatory patterns between expanded gene families

Although a large portion of the expanded innate repertoire was shown to be expressible, the regulatory pattern varied greatly among families (Fig. 1*C*). Compared with those nonexpanded PRR families (PGRP, GNBP, retinoic acid-inducible gene I-like helicase receptor [RLR], CD36, galectin, V region containing chitin-binding protein [VCBP], chitinase, other chitin-binding proteins, and lysozymes), the expanded PRR families (TLR, LRRIG, LRR-only, SRCR, CLR, C1q, and fibrinogen) were expressed at a low level, that is, <1.7 EST reads per gene (fibrinogen genes at 2.5 reads per gene). Even when considering only the genes with ESTs, the expression was fewer than 3.4 reads per gene (3.9 for fibrinogen). TLR was an extreme example, which was represented by 0.9 reads per gene (1.6 in expressed genes). These expression patterns are presented in Fig. 1*C* and Supplemental Fig. 1. The transcript length may affect EST abundance in a shotgun sequencing strategy, but similar results were obtained when restricting the calculation to the 3′-most 1000 bp of transcripts. Interestingly, NLR genes were an exception among the expanded PRR families, which were expressed in a higher proportion (63 of 73 genes) and at a higher level (>5 reads per gene). This is probably related to the function of NLRs. First, NLRs are intracellular receptors typically and specifically expressed in the gut; second, many NLRs in vertebrates perform regulatory functions or damage-associated molecular pattern (DAMP) sensing rather than microbial sensing (23, 24). In this analysis, using different BLASTN cutoff *E* values (1×10^{-7}, 1×10^{-6}, 1×10^{-4}, and 1×10^{-3}) did not affect the conclusions, suggesting that our conclusions were robust to the cross-species sequence mapping and the complicated evolutionary modes of gene families. As for other expanded families, heme peroxidases and MASP-like serine proteases were expressed at high levels (>16 reads per gene), and TNFs, TNFRs, and TRAFs were expressed at medium levels (5.4–7.3 reads per gene); however, caspases, TIR, and DFD genes were underrepresented (2–3 reads per gene). Heme peroxidases and MASP genes play catalytic/regulatory roles in the oxidative system and in the complement system, respectively. TNFs/TNFRs and TRAFs represent important cytokines/receptors and conserved adaptors, respectively. Most caspases, TIR, and DFD genes are presumed to be cytosolic proteins involved in the signal transduction for receptors such as TLRs, NLRs, RLRs, and TNFRs. Taken together, except NLRs, TNFs/TNFRs, and TRAFs, the expanded PRR and signal transducer families appear to be controlled at very low expression levels, which may reflect a requirement for controlling the self-reactive activity, the cost of mRNA synthesis, and the cost of signal transduction for great innate complexity.

Time course expression dynamics of 214 immune genes

On the basis of global digital analysis, we selected 214 immune genes for expression time course (0–24 h) analyses using real-time qRT-PCR (Fig. 2*A*, 2*B*; graphic representation with full details is shown in Supplemental Figs. 3 and 4; the raw data are presented in Supplemental Table III). The qPCR and digital analyses produced consistent results, with a Pearson correlation coefficient of 0.83 on the mixed bacterial treatment ($p < 0.01$). Specifically, at the

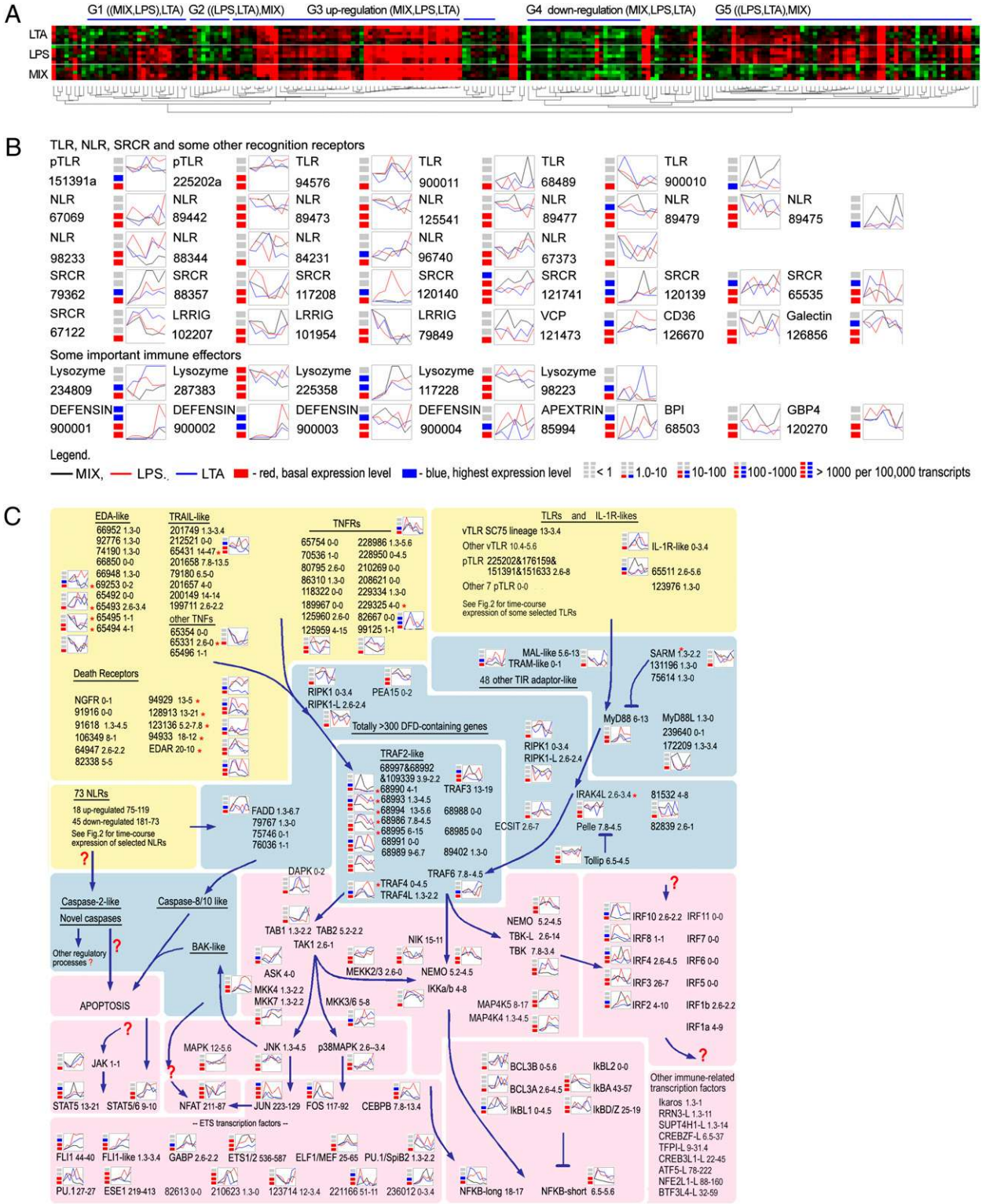

FIGURE 2. *A*, Unsupervised hierarchical clustering analysis of the expression time course of 214 candidate immune genes based on a Pearson correlation coefficient. Five groups of distinct expression patterns are marked. Refer to Supplemental Fig. 3 for a detailed version of this analysis. *B*, Graphic representation of the time course expression dynamics of selected immune receptors and effectors using quantitative RT-PCR. The expression levels of each gene have been normalized by the highest expression level of the gene. *C*, The putative immunity-related signaling network in amphioxus. The digits after gene symbols are the normalized UC and BC expression levels. The time course expression dynamics are also provided when available. Question marks indicate the pathways lacking sufficient information or robust functional evidence. A yellow background marks the layer of cytokines and receptors, a blue background marks the layer of adaptors and intermediate signal transducers, and a pink background marks the terminal signaling network.

significance level of 0.01 (*t* test) for digital analysis, 85% of the genes analyzed by both methods showed a similar trend of expression change. However, this dropped to 77% when the significance level was 0.05, possibly due to the relatively small sampling number of EST reads, which provided a poor basis for estimation for low-expressed genes.

Three types of treatment were conducted in our time-course analyses: bacterial mixture (MIX), LPS, and LTA, which mimick a complex infection, a Gram-negative bacterial infection, and a Gram-positive bacterial infection, respectively. To compare expression levels of different genes, we normalized them to the basal expression level of GAPDH in noninfected animals. The basal expression level of GAPDH was arbitrarily set to 100 per 100,000 transcripts. This basal number for GAPDH is consistent with the estimate from the UC library (91 per 100,000 ESTs) and roughly in agreement with reports of other species (800–2000 per million transcripts). Therefore, in this study, the digital expression level can be roughly compared with the qPCR expression level (see *Materials and Methods*). Given that the action of a gene may have been significantly altered or reached peak expression before its transcript concentration peaked, we also recorded the time points of the first significant change of expression in addition to the peak expression.

In agreement with reports of other species (20, 21, 25), the expression profiles of the three infection models largely overlapped; however, five groups of distinct regulatory patterns could be identified through unsupervised hierarchical clustering analysis (Fig. 2*A*, Supplemental Fig. 3). In group 1, genes had similar expression patterns in MIX and LPS treatments, but they had a different one in LTA treatment. Genes in group 2 had lower expression levels in LPS and LTA treatments than in MIX treatment; conversely, genes in group 5 tended to be upregulated or have higher expression levels in LPS and LTA treatments than in MIX treatment. In groups 3 and 4, genes exhibited similar expression patterns in all three treatments, except that group 3 tended to be upregulated whereas group 4 tended to be downregulated. Remarkably, group 3 was abundant in effector genes ($p < 0.015$, Fisher exact test), whereas group 5 was enriched with transcription factors, kinases, and adaptors ($p < 0.0015$). Despite some genes having lower expression in MIX treatment, the overall upregulation of the MIX treatment was much higher than that of the LPS treatment (15-fold versus 7-fold, $p < 0.0025$), and LPS was also higher than LTA (7-fold versus 5-fold, $p < 0.0006$). These observations may reflect the true effect of the treatments, but other factors cannot be ruled out, such as batch-to-batch variation, the effective dosage of stimulants, and the persistence time of stimulants in the intestine. However, the most significant upregulation associated with LPS occurred earlier than that of LTA (100 genes at 1 h versus 66 genes at 1 h, $p < 0.0002$), which cannot be caused by stimulant dosage effect or batch-to-batch variation and therefore suggests that the amphioxus immune system reacts to LPS more rapidly than to LTA. Furthermore, the significant expression alterations (mostly upregulations) of transcription factors also occurred earlier than that of other genes ($p < 0.035$), and, on average, kinases and adaptors were expressed at the lowest basal level and exhibited the lowest upregulation among all genes and gene classes (Supplemental Fig. 4, Supplemental Table III).

The evolution and regulation of the terminal signaling network

We analyzed >13 gene families of kinases and transcription factors that are relevant to the immune system (Fig. 2*C*, pink background). Except for the IRF family, most studied genes underwent neither whole-genome duplications nor species-specific expansion, with the ETS family being a typical example (Fig. 3*A*). Other analyzed families included TBKs/IKKs, MAP3Ks, MKKs, MAPKs, TAKs, JUNs/FOSs, STATs, NFAT, Ikaros, NF-κBs, and IκBs (Supplemental Figs. 5–15). This result suggests that kinases and transcription factors comprise the most primitive and concise part of the amphioxus immune system. As an exception, amphioxus IRF family has undergone species-specific duplications

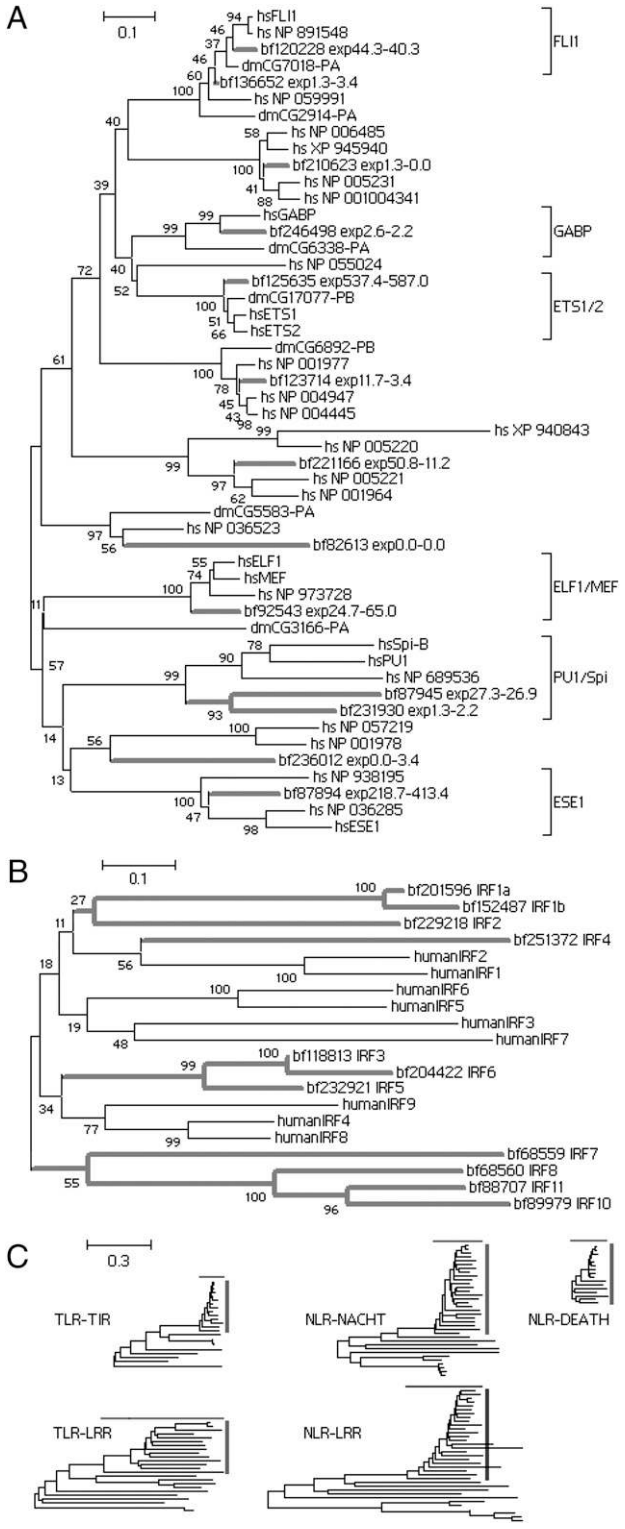

FIGURE 3. *A*, ME protein tree of ETS transcription factors from amphioxus, humans, and *Drosophila*. *B*, ME protein tree of IRF transcription factors from amphioxus and humans. *C*, Comparison of the divergence between domains of amphioxus TLRs and NLRs (the ME protein trees). Note that all trees share the same scale bar so that the branch length of trees can be cross-compared. Potential diversified lineages are marked by vertical bars.

and shares no reliable orthology with any human IRFs (Fig. 3*B*), but both amphioxus IRFs (11 genes) and human IRFs (9 genes) have undergone a medium expansion when compared with the sea

urchin (2 genes) and *Drosophila* (0 gene). We think that the IRF expansion in amphioxus is correlated with the expansion of its innate immune system.

As mentioned earlier, on average kinases were expressed at the lowest basal level and showed the lowest level of upregulation (Fig. 2*C*, Supplemental Fig. 4). A presumable reason for this is that kinases may not require high expression levels, as they are catalytic and reusable. We also mentioned above that if the expression of the studied transcription factors was significantly altered (mostly upregulated), they tended to be altered earlier than other genes (Fig. 2*C*, Supplemental Fig. 4), which should be a reflection of their leading role in the subsequent regulation of other genes. Time course expression dynamics and hierarchical analysis further revealed that a third of the studied transcription factors assumed similar expression patterns in different treatments (LPS, LTA, and MIX), and half of the studied transcription factors responded to LPS/LTA treatments and MIX treatment in different expression patterns (mostly being lower expression level associated with MIX treatment) (Fig. 2, Supplemental Figs. 3, 4). The former case implicated IRF2, STAT5, STAT5/6, GABP, ETS1/2, ELF1/MEF, and four other ETSs (models 210623, 123714, 221166 and 236012), which we suggested should play a pivotal and consistent role in different immune responses. However, the latter case that implicated IRF8, IRF10, NFAT, JUN, FOS, CEBPB, two NF-κBs, three IκBs, and four ETSs (FLI1, FLI1-like, SpiB1, SpiB2) was quite unexpected. We inferred that multiple signal inputs offered by MIX treatment should account for the MIX-specific expression patterns of these transcription factors, as the only difference between LPS/LTA and MIX treatments is that MIX treatment produced many more types of stimulants. Remarkably, these MIX-specific patterns seemed to be associated with positive consequence, because, overall, we observed much greater upregulation in MIX treatment than in LPS or LTA treatments (see above).

Initial signal receptors are not as diversified as their expansion suggests

TLR and NLR are major innate receptor families capable of signaling initiation. In amphioxus these two families are expanded and hence become potential "innate diversified receptors" (1, 10). We define "diversified" as "a state of being rapidly duplicated and rapidly diversified that is driven by pathogens" for those germ line-encoded families. In this study, we reevaluated amphioxus TLRs and NLRs under this definition. Amphioxus has 11 protostome-type TLRs, at the same level with the sea urchin (8 genes) and *Drosophila* (9 genes). Protostome-type TLRs are lost in bony vertebrates, and their exact function in amphioxus and the sea urchin is unknown, but in protostomes they are cytokine receptors (9). Amphioxus has 28 vertebrate-type TLRs (VTLRs), which is a lot less than the sea urchin (211 genes) and not much more than vertebrates (10–18 genes). Unlike the situation in the sea urchin (11, 26), in amphioxus only a third of VTLRs (10 genes) meet our definition of diversified. These VTLRs belong to the so-called SC75 lineage, encoded in single exons and subjected to rapid diversifying selection (10). There is no information about the pathogen-associated molecular pattern ligands of amphioxus VTLRs (nor on the sea urchin VTLRs). Despite some being upregulated by immune challenge, the expression levels of amphioxus TLRs were low in the gut (Fig. 2*B*, 2*C*; also discussed above), suggesting that either their role in mucosal immunity is not essential, their function (signaling initiation) does not require high expression, or their expression is restricted to particular immunocytes rather than any epithelial cells.

Amphioxus had a medium expansion on NLRs (73 genes, compared with 25–40 in vertebrates and 203 in the sea urchin). The sea urchin NLRs tend to be under rapid diversification (26), whereas the case for amphioxus NLRs is not so simple. Amphioxus NLRs can be separated into two categories. One category is similar to the vertebrate NLR family, which contains diverse domain architectures (10). Apparently, diversifying selection pressures from pathogens are difficult as an explanation for the origin of diverse architectures. Another category is similar to the sea urchin NLR family, which includes 41 genes adopting a unified NACHT-LRR structure (with the typical DEATH-NACHT-LRR as the predominant structure). Thirty NLRs of this category are clearly derived from one lineage by rapid duplications, but further study indicates that this lineage may not have undergone rapid diversification, because the LRR divergence between lineage members is much lower than that of the SC75 lineage of TLRs, and actually it is even not much greater than that of their NACHT or DEATH domains (Fig. 3*C*). We speculated that these LRR regions may be dedicated to microbe recognition, but the diversifying pressure from microbes is not high enough to significantly drive up the sequence divergence. Alternatively, these LRR regions may be dedicated to recognize partner proteins or endogenous DAMP molecules that generally impose less diversifying selection. Whatever the case, both explanations do not meet our definition of diversified. Amphioxus NLRs had high expression levels as well as high upregulation of folds in the gut (Fig. 2*B*; also discussed above), suggesting that they are essential innate receptors in the gut as is the case with their vertebrate counterparts. In fact, many vertebrate NLRs perform DAMP sensing and regulatory functions in bowel inflammation rather than microbe detection (23, 24).

Differential expression patterns observed within TNF and IL-1R systems

Unlike the undeveloped state in other invertebrates, amphioxus has as large a TNF/TNFR system as do vertebrates despite having quite different phylogenetic patterns between them (10). As a typical cytokine-receptor system, the TNF/TNFR network does not directly sense pathogens, but it regulates immune response and inflammation by facilitating intercellular communication. On average, TNF/TNFR genes were expressed at medium levels in the gut (see above), but expression levels and changing patterns varied considerably in different genes and treatments (LPS, LTA, or MIX) (Fig. 2*C*). We therefore concluded that the amphioxus TNF/TNFR system has active and important, but differential, roles in regulating the gut mucosal immune responses. Additionally, two of three amphioxus IL-1R–like cytokine receptors were substantially upregulated during infection, suggesting their active role in the regulation of gut mucosal immunity (Fig. 2*C*).

Differential roles between conserved and nonconserved intermediate signal transducers

An early study has shown that in amphioxus, downstream of TLRs, NLRs, TNFRs, and other receptors (e.g., IL-1Rs and RLRs) lies a huge intracellular intermediate signal-transducing network not previously seen in other genomes (10). This network may include 57 TIR genes, 332 DFD genes, 17 TRAFs, 36 initial caspases, and some other genes (Fig. 2*C*). As discussed earlier, on average the expression of expanded adaptor families (except TRAFs) tended to be underrepresented. In this case, we wanted to know the specific situations on individual genes. We assume that since adaptor proteins are less reusable and lack catalytic power, their essentiality should be correlated with their expression levels or upregulation folds. Following this assumption, we found that the highest expression levels of those conserved adaptors (adaptors with unambiguous orthologs in vertebrates, protostomes, or other chordates)

were 5–10 times higher than those of the nonconserved adaptors. These conserved adaptors include TAB1, TAB2, MyD88, TRAF2s, TRAF3, TRAF4s, TRAF6, RIPK1s, IRAK4, IRAK/pelle, FADD, caspase-8/10, Tollip, CRADD, SARM, ECSIT, Bam32, and MALT1. Furthermore, through in vitro functional analyses the presence of two conserved pathways has been confirmed, the VTLR-MyD88-IRAK-TRAF6-NF-κB pathway (27, 28) and the TNFRs/death receptors-FADD-(caspases) pathway (29). The first pathway is also shown to be negatively regulated by SARM, which suppresses MyD88 and TRAF6 by physically interacting with them (30). Therefore, given high expression levels, sequence conservation, and protein-interaction conservation, we suggested that conserved intermediate signal transducers have a primary role in signaling of the gut mucosal immunity. As for the nonconserved adaptors, we suggested that they might either play complementary or subsidiary roles for the conserved adaptors, or they might play important roles in specific situations or fashions.

Apoptosis network is regulated to favor cell survival and activation

The amphioxus genome encodes an expanded caspase-dependent apoptosis network (Fig. 4*A*). Both genome-wide expression profiles and time course dynamics suggested that apoptosis in the gut mucosal tissue was suppressed after bacterial challenges (Fig. 4*A*, Table I, Supplemental Fig. 4). In particular, the pathway from caspase-8/10 to caspase-3/6/7 was significantly suppressed. This implied that caspase-8–dependent TNF-induced apoptosis was blocked, and the TNF/TNFR signals might be all routed to NF-κB, IRFs, and other transcription factors that favor cell survival and activation. In vertebrates, some caspases may also interact with NLRs and lead to inflammation. We speculated that there could be similar interaction in amphioxus (Fig. 4*A*). Consistent with this speculation, some caspases (e.g., 105741, 105738, and 81881) may have a high possibility of interacting with some NLRs (e.g.,

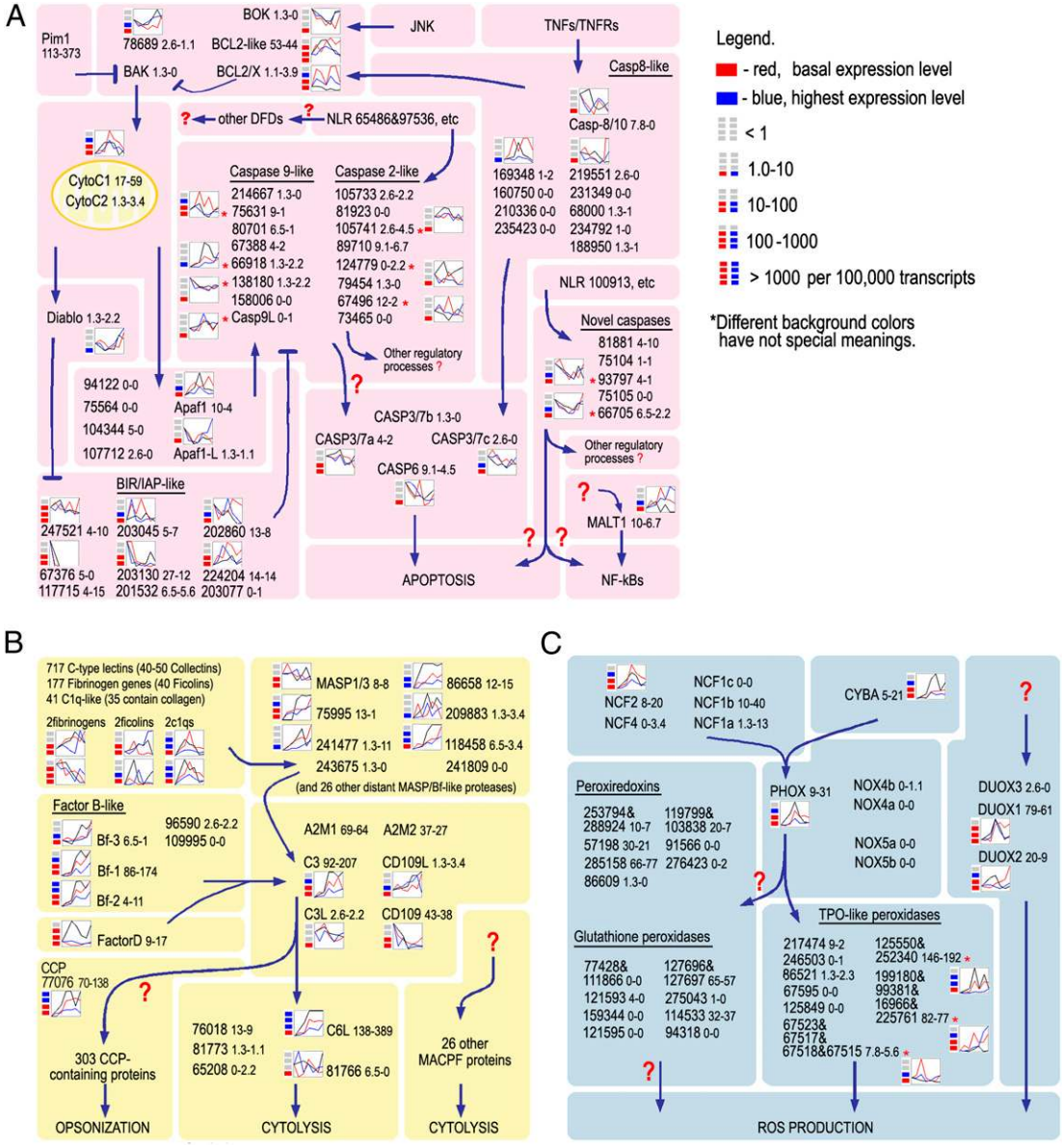

FIGURE 4. *A*, The putative apoptotic network. The phylogenetic analysis of BCL2 family is provided in Supplemental Fig. 16. *B*, The putative complement pathways. The phylogenetic analysis of C3-like, C6-like, and MACPF family is provided in Supplemental Figs. 17 and 18. *C*, The putative oxidative pathways. The phylogenetic analysis of peroxiredoxins, glutathione peroxidases, NCFs, NOXs, and TPO-like families is provided in Supplemental Figs. 19–23. The digits after gene symbols are the normalized UC and BC expression levels. The time course expression dynamics are provided when available. Question marks indicate the pathways lacking sufficient information or robust functional evidence.

65486, 97356 and 100913) due to their high amino acid identity (55–65%) in the DFD domains. Finally, we also observed the induction of MALT1, a unique caspase containing two Ig domains, which in vertebrates is able to promote NF-κB and lymphocyte activation.

Diversified PRRs are expressed at low levels in the gut

There are eight expanded PRR-like families in amphioxus, including NLR, TLR, SRCR, CLR, C1q, fibrinogen, LRRIG and LRR-only (Table II). LRR-only genes are not further discussed here because they are of diverse origins and are difficult to be modeled. The other seven families except NLR were expressed at very low levels in the gut (Fig. 1*C*; also see above). Amphioxus NLRs actually do not behave like typical pathogen-driven diversified PRRs, although we do not know exact reasons (discussed previously). However, a real diversified lineage of 10 genes is found in the TLR family (see above), and more potential diversified lineages can be found in six other families (data not shown). LRRIGs appear to be highly diversified, but it is unknown whether their evolution is driven by pathogens. SRCRs are known to primarily function in the immune system (31, 32) and be substantially expressed in the gut (Fig. 2*B*), whereas CLRs, fibrinogen genes, and C1q genes may have broad implications outside the immune system. The good news is that each of these three families has one gene shown to have immune functions (15, 33, 34), which represent a 54-membered CLR subfamily, a 28-membered fibrinogen subfamily, and probably the entire C1q family, respectively. In any case, if these families (except NLRs) really serve as diversified PRRs, then their low expression levels are justifiable (their expression is supposed to be induced when needed), because recognition diversity not only poses the same magnitude of self-reactive potential, but also costs much more resources with slightly higher expression.

A full-fledged complement system is the major immune effector in the gut

Amphioxus has the most developed complement system known in invertebrates (Fig. 4*B*) (10). As a comparison, the sea urchin complement system lacks ficolins, MASPs, and C6 proteins (11). Compared with amphioxus, the complement system of vertebrates is more developed in some aspects. First, its activation mechanism splits into three pathways (classical, alternative, and lectin pathways), with the classical pathway engaging with adaptive immune receptors; second, its cytolytic machinery is more elaborated because C6 proteins have been duplicated into four functionally related paralogs. However, compared with vertebrates, the amphioxus complement system also has special advanced features, which include specific expansions on sensor/elicitor PRRs, Bf/C2-like proteases, MASP-like serine proteases, and CCP genes (Fig. 4*B*) (10). Because PRRs lacking collagen repeats may activate complement as well (35), those diversified PRR families in amphioxus such CLR, fibrinogen, and C1q may be potential sensor/elicitor PRRs for complement activation. We particularly mention a set of ∼116 collagen-containing PRRs (Fig. 4*B*), because all mammalian complement PRRs use collagen for MASP binding. Furthermore, we have observed that recombinant C1q-like and ficolin-like proteins are capable of binding with both carbohydrates and MASP1/3 in vitro (Ref. 34 and unpublished data), and therefore we thought that amphioxus has greatly diversified lectin pathways for complement activation. Expression dynamics analysis suggested that both lectin pathways and Bf-mediated alternative pathways were activated in the gut mucosal immune responses (Fig. 4*B*, Supplemental Fig. 4). Additionally, the putative C6-mediated cytolytic pathway also appeared to be in action (Fig. 4*B*). All genes involved in these active pathways contributed >1% of the transcriptome under infection (excluding membrane attack complex/perforin [MACPF] genes, CCP genes, and non–collagen-containing PRRs). Because we found no other immune mechanisms that get closer to this expression level, we suggested that the complement system is the major effector in the amphioxus gut immunity.

A well-developed oxidative burstlike system plays an important role in the gut mucosal immunity

Oxidative burst occurs in the activated mammalian phagocytes, which rapidly release a large amount of reactive oxygen species (ROS) into the phagosome to kill ingested bacteria (36). In oxidative burst, NADPH-oxidase PHOX, together with CYBA and neutrophil cytosol factors 1 and 2 (NCF1 and NCF2), produces H_2O_2, and then myeloperoxidase converts H_2O_2 into ROS. Furthermore, NADPH-oxidase NOX1, NOX4, DUOX1 and DUOX2, lactoperoxidase, eosinophil peroxides, and thyroid peroxidase (TPO) also produce ROS and play a role in innate immunity (37–39). NOX1 and NOX4 are both LPS inducible; DUOX2 and TPO are responsible for thyroid hormone synthesis; and DUOX1 and DUOX2 are expressed in mucous epithelia and provide H_2O_2 to lactoperoxidase to produce extracellular ROS. As for insects, it is known that *Drosophila* lacks all components of oxidative burst, but *Drosophila* nevertheless produces ROS during encapsulation, melanization, and mucosal immune response (9). There are two NOX genes (NOX and DUOX) in *Drosophila*. Downregulation of *Dropsophila* DUOX reduces ROS production in the gut epithelia and leads to rapid lethal gut infection (40). As for the sea urchin, we found that it lacks CYBA, myeloperoxidase, and NCF1 and has a questionable NCF2 homolog. In contrast with other investigated invertebrates, amphioxus has all key components of oxidative burst (Fig. 4*C*). Amphioxus lacks myeloperoxidase, but it has an expanded family of TPO-like peroxidases instead (Fig. 4*C*, Supplemental Fig. 23). Expression analysis indicated that some of these key components were simultaneously upregulated and expressed at high levels (much higher than that of GAPDH) (Fig. 4*C*, Supplemental Fig. 4). This suggests that a functional oxidative burstlike pathway exists in amphioxus and has an important role in gut immunity (Fig. 4*C*). Besides, amphioxus has three DUOX genes, two of which were shown to be upregulated and expressed at high levels (much higher than those of GAPDH) (Fig. 4*C*, Supplemental Fig. 4), suggesting their roles in gut immune responses.

Many amphioxus PGRPs are among the important effectors in the gut

PGRPs are important for antimicrobial defense (41, 42). *Drosophila* has 13 PGRPs that function as either sensors or effectors. Sensor PGRPs recognize pathogens and activate innate signaling pathways such as Toll, Imd, and prophenoloxidase, whereas effector PGRPs have either direct bactericidal or amidase activities. PGRP amidases can hydrolyze peptidoglycan to reduce its immunostimulatory activity. Mammals possess four PGRPs, all of which serve as effectors (41, 42). We found that amphioxus has 17–18 PGRP genes, none of which reliably clusters with insect or mammalian PGRPs (Fig. 5*A*). Therefore, phylogentic analysis provided no clues for the function of amphioxus PGRPs. In contrast, sequence analysis indicated that all amphioxus PGRPs have Zn^+ binding and amidase active sites, suggesting their potential amidase activity (Supplemental Fig. 25). There are two types of peptidoglycan, the diaminopimelic acid (DAP) type and the L-lysine (Lys) type. The DAP type is found in all Gram-negative bacteria and Gram-positive *Bacillus*, whereas the Lys type is found

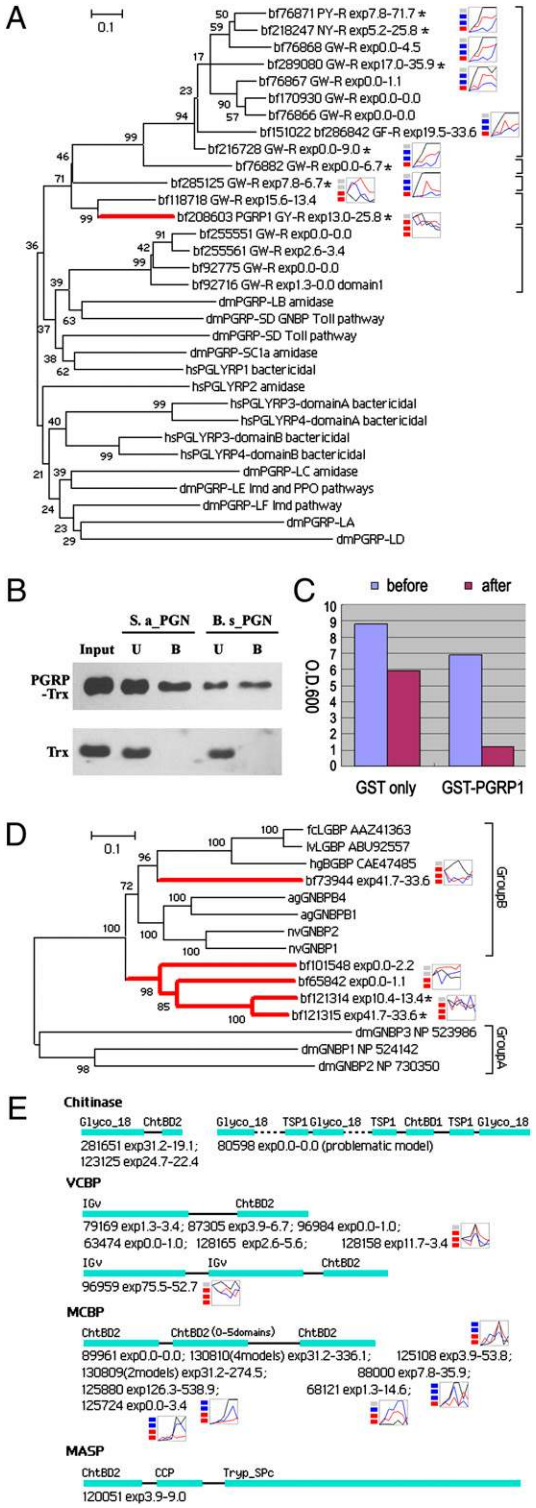

FIGURE 5. *A*, The ME protein tree of PGRPs of amphioxus, humans, and *Drosophila*. The motif for peptidoglycan-binding specificity is provided for amphioxus PGRPs. The function of humans and *Drosophila* PGRPs is also provided. *B*, The PGN binding assays of fusion protein His-tagged TRX-bjPGRP1 showed the PGN-binding activity of the recombinant bjPGRP1. The TRX protein was used as control. B, binding protein; B.s_PGN, PGN from *B. subtilis* (DAP type); S.a_PGN, PGN from *S. aureus* (Lys type); U, unbinding protein. *C*, The chloroform-induced bacteriolysis assay showed the bacteriolytic activity of the recombinant bjPGRP1. *D*, The ME protein tree of GNBPs of amphioxus and arthropods. *E*, Protein architectures of those amphioxus CBPs. Numbers with an "exp" prefix are the normalized UC and BC expression levels. The time course expression dynamics are also provided for amphioxus PGRPs, GNBPs and CBPs when available. ag, *Anopheles gambiae*; bf, *B. floridae*; dm, *Drosophila melanogaster*; fc, *Fenneropenaeus chinensi*; hg, *Homarus gammarus*; hs, *Homo sapiens*; lv, *Litopenaeus vannamei*; nv, *Nasutitermes comatus*.

in most Gram-positive bacteria. It has been proposed that PGRPs that prefer binding to the DAP type possess a GW-R motif, whereas PGRPs that prefer the Lys type have an NF-V motif (42). In amphioxus, most PGRPs bear the GW-R motif and none bears the NF-V motif, but other variants such as GY/F-R, NY/W-R, and PY-R exist, suggesting a certain degree of recognition diversity for amphioxus PGRPs (Fig. 5*A*, Supplemental Fig. 25). Many amphioxus PGRPs were greatly upregulated during the gut immune responses, with the peak expression level several times higher than that of GAPDH (Fig. 5*A*, Supplemental Fig. 4). The top 10 highly expressed PGRPs contributed >0.3% of all transcripts in the BC cDNA library. If we assume that sensor PRRs are generally expressed at lower levels than effector PRRs, we would infer that those PGRPs of extremely high expression are likely effectors. Guided by this prediction, we chose PGRP1 (i.d. 208603) for functional analysis. PGRP1 is among the highly expressed PGRPs and bears a nonconserved binding motif. Functional analysis with recombinant PGRP1 proteins showed that PGRP1 could bind both DAP and Lys types of peptidoglycan, with higher affinity toward the Lys type (Fig. 5*B*). Remarkably, PGRP1 was able to lyse the cell wall of *E. coli* (Fig. 5*C*). Taken together, we suggested that many amphioxus PGRPs should function as effectors, and they should be one of the major effectors in the gut mucosal immunity considering their high expression levels.

GNBPs

GNBPs, also known as LPS- and β-1,3-glucan recognition proteins, represent another major PRR family in protostomes (43). GNBPs can be divided into two groups. Group A is restricted to *Drosophila* and has lost the key residues for glucanase activity (44), whereas group B is present in various invertebrates and has predicted glucanase activity. Notably, *Drosophila* GNBP1 and GNBP3 act as sensor PRRs and work with PGRPs in the Toll pathway (9). GNBPs have been lost in jawed vertebrates, but five GNBPs are found in amphioxus, suggesting the presence of GNBPs in the chordate ancestor. One amphioxus GNBP is clustered with the decapod lineage in group B, while the remaining four form an independent lineage closely related to group B (Fig. 5*D*). Amphioxus GNBPs showed little change in expression, although three of them maintained high expression levels (higher than that of GAPDH) during the gut infection (Fig. 5*D*, Supplemental Fig. 4). Similar expression patterns were also observed in protostomes, so we think that amphioxus GNBPs have some kind of immune functions.

Chitin-binding proteins

Chitin is the second most abundant biopolymer in nature and can be found in fungi, algae, and protostomes. In mammals, some TLRs (e.g., TLR2) and CLRs (e.g., macrophage mannose receptors) can sense chitin and produce an immune response (45). Additionally, mammals have a set of dedicated CBP for digestion and immune regulation (45). CBPs are known to be more abundant and have greater function in arthropods, with roles in digestion, development, structural formation, and host defense (46). There are reports suggesting that arthropod CBPs are not only capable of binding chitin, but also binding and inhibiting bacteria (47, 48). Three distinct functional domains can be found in CBPs: the chitinase catalytic domain, capable of chitin hydrolysis, and the

chtbd1 and chitbd2 (peritrophin-A) domains, capable only of chitin binding. A set of CBPs of varying architectures has been identified in the amphioxus genome, including 3 chitinases, 7 VCBPs, and 12 multiple chitin-binding domain-containing proteins (Fig. 5*E*). VCBP genes encode a peritrophin domain and one or two IgV-like domains. The IgV domains of amphioxus VCBPs exhibit high sequence diversification within individual animals, reminiscent of the somatically diversified BCR/TCR in jawed vertebrates (49). In this study, multiple chitin-binding domain-containing proteins represented one of the highest expressed families in the gut immune response (contributing 1.25% of total transcripts in the BC library). Moreover, chitin-binding domains are also found in other candidate immune genes such as MACPF proteins, MASP-like proteases, and CLRs. The presence of the chitin-binding domain-containing MASP gene (having EST evidence for its architecture) may represent a shortcut activation pathway to the complement cascade against the chitin-containing microbes (Fig. 5*E*). Many amphioxus CBPs showed high expression levels post-infection, suggesting that they may have an important role in gut immunity. Notably, mammals have no chitin synthases, and therefore we can be certain that the role of mammalian CBPs is dedicated to digestion and immunity. On the contrary, in amphioxus we have identified several chitin synthases (data not shown), suggesting that amphioxus may use chitin as a structural component. Therefore, further experiments are needed to clarify amphioxus-specific functions of CBPs.

Other effector genes

Several other effector PRRs play important roles in host defense, including lysozymes, CD36, galectins, and apextrins (Fig. 2*B*, Supplemental Table I). There are three subfamilies of lysozymes: C-type, I-type, and G-type. Mammals and insects have a small expansion of C-type lysozymes, whereas amphioxus preserves only two C-type lysozymes and encodes multiple G-type and I-type lysozymes (Supplemental Fig. 24). Amphioxus C-type and G-type lysozymes were highly expressed during the gut mucosal immune response (Fig. 2*B*, Supplemental Fig. 4). Additionally, the non-PRR effectors bactericidal/permeability-increasing protein, defensins, guanylate-binding proteins, and matrix metalloproteinases were also found to have high expression levels following bacterial infection (Fig. 2*B*, Supplemental Table I). Finally, it is well known that prophenoloxidase-mediated melanization is one of the major effector systems in protostomes, but no prophenoloxidases have been identified in amphioxus. Tyrosinases and laccases were found in amphioxus, but were weakly expressed and were not shown to be induced by bacterial infection in this study.

Discussion

The amphioxus effector system is an important step toward the modern vertebrate innate effector system

If we use expression level as the standard to rate the importance of immune effectors, we find that the complement system, the oxidative system, PGRPs, GNBPs, CBPs, lysozymes, and defensins are major effectors in the gut mucosal immunity in amphioxus. PGRPs, GNBPs, and CBPs are major immune effectors in invertebrates, but they have been significantly downplayed in vertebrates. The origin of complement can be traced back to cnidaria (50), but in this study we show that the prototype of the modern vertebrate complement system first emerged in amphioxus. Unlike the vertebrate system, the amphioxus complement system has no somatically diversified Ag receptors (Igs) as sensors and elicitors, but instead it develops a huge repertoire of germ line-encoded sensors and elicitors (e.g., collectins, ficolins, C1qs, and other noncollagen-containing proteins) as well as elicitor-associated serine proteases. In a certain sense, the amphioxus complement system represents a milestone in the transition from the antimicrobial peptide-based systemic immunity in arthropods to the complement-based humoral immunity in mammals. Production of ROS is also an ancient defense mechanism, but we find that amphioxus may be the first invertebrate having a prototypic oxidative burst system, which marks another major advance in chordate immunity. Taken together, we suggest that from the evolutionary view, the amphioxus effector system is an important step toward the formation of the modern vertebrate innate effector system.

Amphioxus focus on immune signal modulation rather than just on signaling initiation

TLRs and NLRs are two major types of innate receptors capable of initiating immune signaling, but in amphioxus these receptors are not as diversified as their expansions suggest. Thus, despite no consistent phylogenetic patterns between TLRs from amphioxus and other subphyla, most amphioxus TLRs and NLRs are as likely to function as normal innate receptors as do their vertebrate/invertebrate counterparts. As a comparison, the sea urchin contains 222 TLRs and 203 NLRs (11). Following signal initiation, there is signal transduction and modulation, which can be separated into intercellular communication and intracellular modulation. The TNF/TNFR system and the IL-1/IL-1R system are dedicated to intercellular signal communication. The TNF/TNFR system of amphioxus is as large as the system of vertebrates, despite almost no consistent phylogenetic patterns between them. As a comparison, the TNF/TNFR systems from insects and the sea urchin are undeveloped (10). As for intracellular signal transduction/modulation, amphioxus has a huge cytosolic adaptor (and signal transducer) repertoire thus far not seen in other genomes. Expression analyses suggest that those conserved adaptors and signal transducers should occupy critical nodes in the signaling network and form the major pathways. For nonconserved adaptors, we speculate that there exist at least three functional mechanisms: 1) fine tuning the major pathways positively or negatively; 2) mediating subsidiary pathways for the primary pathways or mediating independent and novel pathways, which are likely roles for nonconserved adaptors expressed at a high level; and 3) bridging different pathways or building shortcut pathways, a likely role for genes with novel domain combinations (as elaborated in Ref. 10). As a comparison, despite a greater number of TLRs and NLRs, the sea urchin has fewer adaptors than does amphioxus (10). Taken together, we propose that while the sea urchin immune system may focus more on signal initiation, the amphioxus immune system heavily reinforces its signal transduction and modulation. Remarkably, if we contrast those highly diversified PRRs (SRCRs, CLRs, C1q, fibrinogen genes) and the reinforced signal regulatory system in amphioxus, it reminds us of the vertebrate immune system, where somatically diversified receptors are superimposed on a robust innate regulatory system.

Possible modes for terminal transcriptional regulations

In amphioxus, the terminal signaling network composed of kinases and transcriptional factors remains in a primitive simple form. One may therefore have two questions: Why did this network not expand as did the other parts of the immune system or as the vertebrate counterparts? How does such a primitive simple network efficiently control the expanded immune complexity? The first question may be explained by the gene dosage balance hypothesis. This hypothesis assumes that most conserved kinases and transcription factors also have important roles in biological processes other than immunity and are so-called duplication resistant genes

according to the dosage balance hypothesis (51, 52). This means that even if these genes have the potential to be adapted for new functions after duplication, and the system does have a need for that (e.g., the modulation of a greatly expanded gene repertoire), the initial duplication will have an immediate detrimental effect (e.g., increased dosage that interferes with the original highly constrained cellular system) that prevents the fixation of the duplicates. However, this obstacle may be bypassed through whole genome duplications, since dosage balance is not initially altered by such duplications (52). This explains why 2- to 4-fold increases in kinases and transcription factors can be seen in vertebrates after two rounds of whole genome duplications (18).

As for the second question, an intuitive explanation is that most of the observed immune complexity is pseudogenic, but this idea has been largely disproved in this study. An alternative possibility is the use of selective expression as an effective regulatory strategy. As a paradigm, in the vertebrate adaptive system, the expression or signaling of those diversified BCRs/TCRs is controlled by cell type-specific, cell clone-specific, and even cell phase-specific selective expression. This mode breaks down the overall complexity into small, simple, and easy-to-coordinate parts and hence effectively coordinates the expression of somatically diversified BCRs/TCRs. We expected that amphioxus may use a similar strategy. A third possibility is the development of a new transcription factor family to meet the increased demand for regulation. Rapid duplication and diversifying selection that happened on a new family would result in less deleterious interference than on those conserved and heavily occupied families. The IRF family could be a typical example of this strategy. Basal deuterostomes have only two (in the sea urchin) or three (in hemichordates) IRFs (53), but amphioxus and vertebrates expand this family to more than nine members through distinct evolutionary paths.

Speculations on the evolution of immunological strategies used by basal deuterostomes

Because both the sea urchin and amphioxus have developed a greatly expanded innate immune repertoire, we think that these immunological "big bangs" were common scenes in ancient deuterostomes 500–600 million years ago, probably coupled with the Cambrian big bang of species. Because of the apparent differences between two systems, we speculate that there were actually various kinds of expanded immune systems 500–600 million years ago and that the amphioxus and the sea urchin systems appear to be descendants of two strategically distinct systems. As for the relationship between the innate immune systems of amphioxus and vertebrates, there could be two possibilities. The first possibility is that the big bang of the immune system only happened to the lineage leading to amphioxus (i.e., the cephalochordate lineage). This possibility suggests fundamental differences between amphioxus and vertebrates in terms of immunological strategies. However, this scenario is not consistent with the findings in this study. For example, the prototypes of the vertebrate complement system and oxidative burst system have been formed in amphioxus and serve as major effectors; both amphioxus and vertebrates expanded their IRF families and TNF/TNFR systems; both amphioxus and vertebrate NLR families underwent intense domain reshuffling; and the function of many amphioxus VTLRs and NLRs is close to those of nondiversified innate receptors in vertebrates. Therefore, we favor an alternative possibility; that is, that the big bang occurred in the common ancestor of amphioxus and vertebrates (i.e., in the chordate ancestor). One may ask that if this was true, then why are gene expansions in vertebrates not as significant as in amphioxus, and why are phylogenetic patterns of those expanded families not consistent between vertebrates and amphioxus? For the first question, we think that the rise of adaptive immune system might cause the erosion of those expanded innate immune gene families. For the second question, we argue that it was a natural outcome for those families with high turnover rates.

In conclusion, this study confirms the amphioxus innate immune complexity on the transcriptome level and describes its composition, evolutionary-functional modes, and expression regulatory patterns on a genome-wide scale in the context of gut mucosal antibacterial responses. These results provide many insights into the innerworking of the amphioxus immune system and the immunological differences or similarities between major model deuterostome species. On this basis, we conclude that amphioxus is more similar to vertebrates but fundamentally different from the sea urchin in terms of immunological mechanisms and strategies, and it can serve as a unique model for understanding the evolution of the immunity of vertebrates.

Disclosures

The authors have no financial conflicts of interest.

References

1. Pancer, Z., and M. D. Cooper. 2006. The evolution of adaptive immunity. *Annu. Rev. Immunol.* 24: 497–518.
2. Flajnik, M. F., and M. Kasahara. 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* 11: 47–59.
3. Akira, S., S. Uematsu, and O. Takeuchi. 2006. Pathogen recognition and innate immunity. *Cell* 124: 783–801.
4. Cooper, M. D., and M. N. Alder. 2006. The evolution of adaptive immune systems. *Cell* 124: 815–822.
5. Slack, E., S. Hapfelmeier, B. Stecher, Y. Velykoredko, M. Stoel, M. A. Lawson, M. B. Geuking, B. Beutler, T. F. Tedder, W. D. Hardt, et al. 2009. Innate and adaptive immunity cooperate flexibly to maintain host-microbiota mutualism. *Science* 325: 617–620.
6. Iwasaki, A., and R. Medzhitov. 2010. Regulation of adaptive immunity by the innate immune system. *Science* 327: 291–295.
7. Hoebe, K., E. Janssen, and B. Beutler. 2004. The interface between innate and adaptive immunity. *Nat. Immunol.* 5: 971–974.
8. Carroll, M. C. 2004. The complement system in regulation of adaptive immunity. *Nat. Immunol.* 5: 981–986.
9. Lemaitre, B., and J. Hoffmann. 2007. The host defense of *Drosophila melanogaster. Annu. Rev. Immunol.* 25: 697–743.
10. Huang, S., S. Yuan, L. Guo, Y. Yu, J. Li, T. Wu, T. Liu, M. Yang, K. Wu, H. Liu, et al. 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res.* 18: 1112–1126.
11. Hibino, T., M. Loza-Coll, C. Messier, A. J. Majeske, A. H. Cohen, D. P. Terwilliger, K. M. Buckley, V. Brockton, S. V. Nair, K. Berney, et al. 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev. Biol.* 300: 349–365.
12. Rast, J. P., L. C. Smith, M. Loza-Coll, T. Hibino, and G. W. Litman. 2006. Genomic insights into the immune system of the sea urchin. *Science* 314: 952–956.
13. Rast, J. P., and C. Messier-Solek. 2008. Marine invertebrate genome sequences and our evolving understanding of animal immunity. *Biol. Bull.* 214: 274–283.
14. Yuan, S., Y. Yu, S. Huang, T. Liu, T. Wu, M. Dong, S. Chen, Y. Yu, and A. Xu. 2007. Bbt-TNFR1 and Bbt-TNFR2, two tumor necrosis factor receptors from Chinese amphioxus involve in host defense. *Mol. Immunol.* 44: 756–762.
15. Yu, Y., Y. Yu, H. Huang, K. Feng, M. Pan, S. Yuan, S. Huang, T. Wu, L. Guo, M. Dong, et al. 2007. A short-form C-type lectin from amphioxus acts as a direct microbial killing protein via interaction with peptidoglycan and glucan. *J. Immunol.* 179: 8425–8434.
16. Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
17. Romualdi, C., S. Bortoluzzi, F. D'Alessi, and G. A. Danieli. 2003. IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genomics* 12: 159–162.
18. Putnam, N. H., T. Butts, D. E. Ferrier, R. F. Furlong, U. Hellsten, T. Kawashima, M. Robinson-Rechavi, E. Shoguchi, A. Terry, J. K. Yu, et al. 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453: 1064–1071.
19. Roach, J. C., K. D. Smith, K. L. Strobe, S. M. Nissen, C. D. Haudenschild, D. Zhou, T. J. Vasicek, G. A. Held, G. A. Stolovitzky, L. E. Hood, and A. Aderem. 2007. Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells. *Proc. Natl. Acad. Sci. USA* 104: 16245–16250.

20. De Gregorio, E., P. T. Spellman, G. M. Rubin, and B. Lemaitre. 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc. Natl. Acad. Sci. USA* 98: 12590–12595.
21. Irving, P., L. Troxler, T. S. Heuer, M. Belvin, C. Kopczynski, J. M. Reichhart, J. A. Hoffmann, and C. Hetru. 2001. A genome-wide analysis of immune responses in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 98: 15119–15124.
22. Gilad, Y., and D. Lancet. 2003. Population differences in the human functional olfactory repertoire. *Mol. Biol. Evol.* 20: 307–314.
23. Inohara, N., and G. Nuñez. 2003. NODs: intracellular proteins involved in inflammation and apoptosis. *Nat. Rev. Immunol.* 3: 371–382.
24. Ting, J. P., J. A. Duncan, and Y. Lei. 2010. How the noninflammasome NLRs function in the innate immune system. *Science* 327: 286–290.
25. Boutros, M., H. Agaisse, and N. Perrimon. 2002. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev. Cell* 3: 711–722.
26. Messier-Solek, C., K. M. Buckley, and J. P. Rast. 2010. Highly diversified innate receptor systems and new forms of animal immunity. *Semin. Immunol.* 22: 39–47.
27. Yuan, S., T. Liu, S. Huang, T. Wu, L. Huang, H. Liu, X. Tao, M. Yang, K. Wu, Y. Yu, et al. 2009. Genomic and functional uniqueness of the TNF receptor-associated factor gene family in amphioxus, the basal chordate. *J. Immunol.* 183: 4560–4568.
28. Yuan, S., S. Huang, W. Zhang, T. Wu, M. Dong, Y. Yu, T. Liu, K. Wu, H. Liu, M. Yang, et al. 2009. An amphioxus TLR with dynamic embryonic expression pattern responses to pathogens and activates NF-κB pathway via MyD88. *Mol. Immunol.* 46: 2348–2356.
29. Yuan, S., H. Liu, M. Gu, L. Xu, S. Huang, Z. Ren, and A. Xu. 2010. Characterization of the extrinsic apoptotic pathway in the basal chordate amphioxus. *Sci. Signal.* 3: ra66.
30. Yuan, S., K. Wu, M. Yang, L. Xu, L. Huang, H. Liu, X. Tao, S. Huang, and A. Xu. 2010. Amphioxus SARM involved in neural development may function as a suppressor of TLR signaling. *J. Immunol.* 184: 6874–6881.
31. Sarrias, M. R., J. Grønlund, O. Padilla, J. Madsen, U. Holmskov, and F. Lozano. 2004. The scavenger receptor cysteine-rich (SRCR) domain: an ancient and highly conserved protein module of the innate immune system. *Crit. Rev. Immunol.* 24: 1–37.
32. Bowdish, D. M., and S. Gordon. 2009. Conserved domains of the class A scavenger receptors: evolution and function. *Immunol. Rev.* 227: 19–31.
33. Fan, C., S. Zhang, L. Li, and Y. Chao. 2008. Fibrinogen-related protein from amphioxus *Branchiostoma belcheri* is a multivalent pattern recognition receptor with a bacteriolytic activity. *Mol. Immunol.* 45: 3338–3346.
34. Yu, Y., H. Huang, Y. Wang, Y. Yu, S. Yuan, S. Huang, M. Pan, K. Feng, and A. Xu. 2008. A novel C1q family member of amphioxus was revealed to have a partial function of vertebrate C1q molecule. *J. Immunol.* 181: 7024–7032.
35. Sekine, H., A. Kenjo, K. Azumi, G. Ohi, M. Takahashi, R. Kasukawa, N. Ichikawa, M. Nakata, T. Mizuochi, M. Matsushita, et al. 2001. An ancient lectin-dependent complement system in an ascidian: novel lectin isolated from the plasma of the solitary ascidian, *Halocynthia roretzi*. *J. Immunol.* 167: 4504–4510.
36. Forman, H. J., and M. Torres. 2002. Reactive oxygen species and cell signaling: respiratory burst in macrophage signaling. *Am. J. Respir. Crit. Care Med.* 166: S4–S8.
37. Rada, B., and T. L. Leto. 2008. Oxidative innate immune defenses by Nox/Duox family NADPH oxidases. *Contrib. Microbiol.* 15: 164–187.
38. Leto, T. L., and M. Geiszt. 2006. Role of Nox family NADPH oxidases in host defense. *Antioxid. Redox Signal.* 8: 1549–1561.
39. Klebanoff, S. J. 2005. Myeloperoxidase: friend and foe. *J. Leukoc. Biol.* 77: 598–625.
40. Ha, E. M., C. T. Oh, Y. S. Bae, and W. J. Lee. 2005. A direct role for dual oxidase in *Drosophila* gut immunity. *Science* 310: 847–850.
41. Royet, J., and R. Dziarski. 2007. Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. *Nat. Rev. Microbiol.* 5: 264–277.
42. Dziarski, R., and D. Gupta. 2006. The peptidoglycan recognition proteins (PGRPs). *Genome Biol.* 7: 232.
43. Iwanaga, S., and B. L. Lee. 2005. Recent advances in the innate immunity of invertebrate animals. *J. Biochem. Mol. Biol.* 38: 128–150.
44. Zhang, S. M., Y. Zeng, and E. S. Loker. 2007. Characterization of immune genes from the schistosome host snail *Biomphalaria glabrata* that encode peptidoglycan recognition proteins and Gram-negative bacteria binding protein. *Immunogenetics* 59: 883–898.
45. Lee, C. G. 2009. Chitin, chitinases and chitinase-like proteins in allergic inflammation and tissue remodeling. *Yonsei Med. J.* 50: 22–30.
46. Arakane, Y., and S. Muthukrishnan. 2010. Insect chitinase and chitinase-like proteins. *Cell. Mol. Life Sci.* 67: 201–216.
47. Du, X. J., J. X. Wang, N. Liu, X. F. Zhao, F. H. Li, and J. H. Xiang. 2006. Identification and molecular characterization of a peritrophin-like protein from fleshy prawn (*Fenneropenaeus chinensis*). *Mol. Immunol.* 43: 1633–1644.
48. Loongyai, W., J. C. Avarre, M. Cerutti, E. Lubzens, and W. Chotigeat. 2007. Isolation and functional characterization of a new shrimp ovarian peritrophin with antimicrobial activity from *Fenneropenaeus merguiensis*. *Mar. Biotechnol.* 9: 624–637.
49. Cannon, J. P., R. N. Haire, and G. W. Litman. 2002. Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. *Nat. Immunol.* 3: 1200–1207.
50. Miller, D. J., G. Hemmrich, E. E. Ball, D. C. Hayward, K. Khalturin, N. Funayama, K. Agata, and T. C. Bosch. 2007. The innate immune repertoire in cnidaria: ancestral complexity and stochastic gene loss. *Genome Biol.* 8: R59.
51. Papp, B., C. Pál, and L. D. Hurst. 2003. Dosage sensitivity and the evolution of gene families in yeast. *Nature* 424: 194–197.
52. Conant, G. C., and K. H. Wolfe. 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* 9: 938–950.
53. Nehyba, J., R. Hrdlicková, and H. R. Bose. 2009. Dynamic evolution of immune system regulators: the history of the interferon regulatory factor family. *Mol. Biol. Evol.* 26: 2539–2550.