ISJ 2: 105-113, 2005 ISSN 1824-307X

Minireview

Toll-like receptors in invertebrate innate immunity

L Zheng[1], L Zhang[1, 2], H Lin[1, 3], MT McIntosh[1], AR Malacrida[4]

[1] *Yale University School of Medicine, Epidemiology and Public Health, 60 College Street, New Haven, CT 06520, USA*
[2] *Department of Parasitology, Medical College, Jinan University, Shipai, Guangzhou 510632, Guangdong, P. R. China*
[3] *Fujian Provincial Centers for Disease Control and Prevention, 76 Jintai Avenue, Fuzhou, Fujian, 350001, P.R.China*
[4] *Dipartimento di Biologia Animale, Università di Pavia, Piazza Botta 9, 27100 Pavia, Italy*

Accepted August 4, 2005

Abstract

Among invertebrates, innate immunity is the only defense mechanism against harmful non-self agents. In response to recognition of microbial pattern molecules, *Drosophila melanogaster* activates either the Toll or Imd pathway, leading to the translocation of NF-kB (or Rel) transcription factors from the cytoplasm to the nucleus and the subsequent production of antimicrobial peptides, which provide systemic innate immunity. Toll-like receptors (TLRs) are characterized by an extracellular leucine rich repeat (LRR) domain and an intracellular Toll/interleukin-1 receptor (TIR) domain. TLRs are found from cnidarians to mammals. Here we argue that TLR mediated innate immunity developed during an early stage of evolution when organisms acquired a body cavity. This is supported by the distributions of TLR and Rel genes in the animal kingdom. Further, TLR mediated immunity appears to have developed independently in invertebrates and vertebrates. Recent studies have shown that microbial molecules, with the potential to signal through TLR, can be beneficial to host survival. Studies on this signaling pathway could open doors to a better understanding of the origins of innate immunity in invertebrates and potential transmission blocking strategies aimed at ameliorating vector-borne diseases.

Key words: Toll; innate immunity; antimicrobial peptides; invertebrates; coelom

Introduction

In invertebrates, innate immunity is the sole defense mechanism against infectious non-self agents. There are three manifestations of innate immunity: phagocytosis, activation of humoral responses leading to coagulation, melanization or opsonization, and the systemic production of antimicrobial peptides, or AMPs (Hoffmann *et al.*, 1999). Production of AMPs may represent a more recent evolutionary adaptation of innate immunity, allowing for an effective and efficient means of protection against systemic infection.

Corresponding Author:
Liangbiao Zheng
Yale University School of Medicine, Epidemiology and Public Health, 60 College Street, New Haven, CT 06520, USA
E-mail: liangbiao.zheng@yale.edu

Phagocytosis, being the most primitive feature, developed probably as a means of acquiring nutrients and was probably co-opted later as a defense mechanism. While origins of the other two forms of innate responses are not clear, it is likely they too developed quite early when life forms co-existed in primordial oceans and had to contend with a milieu of diverse microbial agents (friends and foes). That said, for primitive aquatic organisms lacking a real body cavity it would not have been efficient to produce and secrete AMPs onto the surfaces as these would be subjected to dilution by the environment and thus never reach the critical concentration required for effective clearance of microbes. It is therefore likely that systemic immunity in the form of production of AMPs mediated by Toll-like receptors (TLRs) first developed in organisms with a body cavity.

Systemic immunity in the model system *Drosophila melanogaster*

Systemic production of AMPs is best characterized in the fruit fly, *Drosophila melanogaster* (reviewed in Hoffmann, 2003). Generally speaking, immunity against fungi and Gram(+) bacteria is mediated by the Toll-Dif/Dorsal pathway while anti Gram(-) bacterial responses are mediated by the Imd-Relish pathway, leading to the activation and expression of a different set of antimicrobial peptide genes.

Imd was first identified genetically as a key factor for survival against bacterial challenge (Lemaitre *et al.*, 1995). How Gram(-) bacteria are recognized remains to be determined; though it is known that the recognition eventually leads to the activation of a member of the peptidoglycan recognition protein (PGRP) family, PGRP-LC. PGRP-LC interacts directly with Imd (Choe *et al.*, 2005), which then sends the signal through two separate channels. One leads to the activation of a signalsome consisting of IKKβ and IKKγ via Tak1. The other channel leads to the activation of Dredd, a member of the caspase family. Both channels converge on the NF-kB factor, Relish. IKKβ and IKKγ complex phosphorylate the I-kB domain of Relish, while Dredd presumably cleaves the I-kB domain (Stoven *et al.*, 2000, 2003). This allows the amino terminal portion of Relish, carrying both DNA binding and transcriptional activation domains, to enter into the nucleus and activate expression of antimicrobial peptide genes such as diptericin and cecropin.

In *D. melanogaster*, fungal pathogens lead to an activation of an extracellular protease, *Persephone*, while anti-Gram(+) bacterial responses require the products of *PGRP-SA* and *GNBP1*, members PGRP and Gram negative binding protein (GNBP) families, respectively. Both fungal and Gram(+) pathogens eventually converged on the cytokine-like molecule Spaetzle, a protein with cysteine knots. Proteolytic cleavage of pre-protein Spaetzle produces the ligand for Toll. Toll then sends the signal through a multimeric protein complex, consisting of MyD88, Tube and Pelle, leading to the phosphorylation of Cactus, an I-kB like molecule, and subsequent activation and nuclear translocation of Dif or Dorsal, both belonging to the NF-kB family of transcription factors. In adult flies, Dif activates transcription of antimicrobial peptide genes such as drosomycin and metchnikowin (reviewed in Hoffmann, 2003).

Toll was originally identified as a gene required for dorsal ventral patterning (Nusslein-Volhard *et al.*, 1980) but was then found to be required for antifungal immunity in adults (Lemaitre *et al.*, 1996). *Toll* is one of nine TLR genes in *D. melanogaster* that encode proteins with extracellular leucine rich repeat (LRR) arrays and an intracellular Toll-interleukin-1 receptor (TIR) domain (Tauszig *et al.*, 2000). Thus far, genetic studies have shown that Toll (or Toll1) is the primary receptor in innate immunity. However, experiments in cell lines also suggest that Toll5 (Luo *et al.*, 2001) and Toll9 (Ooi *et al.*, 2002) can activate expression of antimicrobial peptide genes. Toll1 and Toll5 (also known as Tehao) are similar in the TIR domain, though Toll1 carries a carboxyl terminal extension which Toll5 lacks (Luo and Zheng, 2000). To date, screening by whole genome RNA interference has not elucidated the functions of the other eight TLR genes in *D. melanogaster*.

Toll and Imd signaling pathways in insects

The origin of the innate immune systems will be better understood as more and more genomes of invertebrates are determined. The completion of the genome sequencing of *Anopheles gambiae,* the main vector for human malaria (Holt *et al.*, 2002), in addition to the resolved genome of *D. melanogaster*, allows for comparative immunology to be examined between these two dipterans. Although most of the Toll and Imd signaling components are conserved in *A. gambiae*, significant differences have been found (Christophides *et al.*, 2002). Similar conclusions can be drawn from the ongoing genome sequencing of another major vector of human diseases, *Aedes aegypti*, which transmits dengue and yellow fever viruses (unpublished observations).

For the Imd pathway, it was found that *Rel2,* in both *A. gambiae* and *A. aegypti,* is both structurally and functionally different from its orthologue, Relish, in *D. melanogaster*. Rel2 in mosquitoes contains an additional death domain which is absent in the fly orthologue (Shin *et al.*, 2002). Furthermore, the *Rel2* gene in mosquitoes produces different spliced variants. In *A. aegypti*, *Rel2* produces three spliced forms, yielding a full length protein, with both rel-homology domain (RHD) and I-kB ankyrin repeats, and partial length polypeptides, containing either the I-kB domain or the RHD domain, respectively (Shin *et al.*, 2002). In *A. gambiae*, two spliced variants have been found thus far, corresponding to the full length and RHD domain containing polypeptides from *A. aegypti*, respectively. More importantly, it has been shown in *A. gambiae* that the Imd pathway functions in immune responses against both Gram(+) and Gram (-) bacteria (Meister *et al.*, 2005).

Components of the Toll pathway are also found in *A. gambiae*, though once again significant differences have been found between it and *D. melanogaster*. The most conspicuous being the absence of a *Dif* orthologue in both *A. gambiae* and *A. aegypti*. It was found recently that the Dorsal homologue, *Rel1* from *A. aegypti* also produces spliced variants, Rel1A and Rel1B, differing at the carboxyl terminal domains (Shin *et al.*, 2005). Similar spliced variants have been found for Drosophila Dorsal. Concomitant over-expression of the two variants *Rel1A* and *Rel1B* from *A. aegypti* leads to the up-regulation of *DefA* and *CecA* in a cell line (Shin *et al.*, 2005). Alternative splicing also appears to occur in the *A. gambiae Dorsal* homologue known as Gambif (Barillas-Mury *et al.*, 1996) or *Rel1*. One expressed sequence tag (ENSANGESTT00000367694) of *Rel1* from *A. gambiae* represents a partial cDNA clone which uses different splice sites than the full-length transcript *Rel1* transcript characterized previously (Barillas-Mury *et al.*, 1996). At present, whether or not spliced forms similar to the

Aedes Rel1A and Rel1B exist in *A. gambiae* remains to be determined.

Another major difference is the presence of four genes (*Toll1A*, *1B*, *5A* and *5B*) that share similarities in the TIR domain with *D. melanogaster Toll1* and *Toll5*. *Toll1A* and *5A* are arranged in tandem on the X chromosome, while Toll1B and 5B are situated in tandem on the third chromosome (Christophides *et al.*, 2002). While it has been shown that over-expression of Toll related genes from mosquitoes in *D. melanogaster* cell lines leads to elevated expression from the promoter of drosomycin (Luna *et al.*, 2002, 2003), there remains no additional functional evidence for the involvement of mosquito Tolls in innate immunity.

Structure and functions of Toll-like receptors (TLRs)

As indicated above, Toll receptors are transmembrane proteins with an extracellular LRR domain interspersed with cysteine knots. The intracellular domain shares sequence similarities with the Interleukin-1 receptors from mammals, and is called the Toll/Interleukin-1 receptor (TIR) domain (reviewed in Imler and Zheng, 2004).

The TIR domain is composed of approximately 150 amino acid residues. It is present not only in TLRs, but also in other intracellular signaling components, such as MyD88. These intracellular TIR-containing adaptor molecules, however, do not contain a leucine rich repeat. In addition, TIR domains and leucine rich repeats are also present in some plant resistance (*R*) genes, typified by *RPP5*, which also has a nucleotide binding domain (NBD) (Parker *et al.*, 1997). In contrast to TLRs, the plant TIR-NBD-LRR resistant proteins are intracellular. Furthermore, their TIR domains are phylogenetically distinct from TLRs. Thus, it seems likely that the TIR domain-containing adaptor and plant TIR-NBD-LRR genes evolved independently from TLRs.

The molecular structures of the TIR domains of human TLR1 and TLR2 have been determined and show a five-stranded parallel β-sheet surrounded by five helices on both sides (Xu *et al.*, 2000). There is a conserved surface which primarily consists of a loop, known as the BB loop, which protrudes from the rest of the structure. Mutations in the BB loop result in a loss or significant reduction in activity of TLRs from both man and *Drosophila*. Within this region are a few highly conserved charged amino acid residues, including an arginine at loop position 3 and an aspartic acid at loop position 4. A proline residue at loop position 7 is also known to be essential for LPS signaling by human TLR4. As eluded to previously, this proline residue is not conserved in all TLRs (Luna *et al.*, 2003) and functional studies suggest that it is not a key structural determinant (Xu *et al.*, 2000). Instead, substitution of this residue may disrupt interactions with other signaling components needed for activation of the pathway. Mutation at a histidine residue in the BB loop position 1 in *D. melanogaster*, represented by allele Tl^{rB1} or Tl^{r6} in the flybase database (www.flybase.org), also results in a loss of function of Toll in dorsal ventral formation in early embryogenesis (Schneider *et al.*, 1991). This histidine residue is highly conserved in insect TLRs, but not in vertebrate TLRs. Loss of function mutations outside the BB loop yet still within the TIR domain have also been documented in *D. melanogaster*. How these residues fully contribute to the function of Toll signaling has yet to be determined. As TIR domains of MyD88 genes from different insects vary widely in sequence, variation on this interacting surface may contribute to species specific signaling in insects (unpublished observations).

Distribution of TLRs in the animal kingdom

Thus far, on the tree of life, TLRs have been found to be present in animals ranging from cnidarians to mammals, though they appear to be absent in platyhelminths (Table 1). TLR genes have been found in all vertebrates where they also play an important role in bridging innate and adaptive immunity, reviewed in (Pasare and Medzhitov, 2004). For example, human and mouse genomes contain 10 and 12 TLR genes, respectively. These TLRs are activated, individually or in combinations, by different bacterial pattern molecules, such as lipopolysaccharide, peptidoglycan, or naked DNA (for a recent review see Kaisho and Akira, 2004).

TLRs are also found in invertebrates that have body cavities (Table 2). *D. melanogaster* possesses 9 TLRs while *A. gambiae* encodes 10. Thus far, at least one TLR gene has been found in all insect genomes. Within the phylum Arthropoda, one TLR gene has been found in the horseshoe crab *Tachypleus* (Inamori *et al.*, 2000, 2004) and the lobster *Homarus americanus* (accession number: CN852754, unpublished observation). Limited genomic information has shown that at least one TLR gene is present in *Euprymna*, the Hawaiian squid which belongs to the phylum Mollusca (Hoa NT *et al.*, unpublished observations). Likewise, Nematoda genome sequencing has revealed that *Caenorhbiditis elegans* and *Caenorhbiditis briggsea* each encode one TLR. Interestingly, functional studies showed that this unique TLR gene (*CeTol-1*) appears to have no function in innate immunity in *C. elegans* (Pujol *et al.*, 2001).

A TLR gene has yet to be found in the phylum Platyhelminth, though significant genomic information is available for *Schistosoma japonicum*, *Schistosoma mansoni* and *Schmidtea mediterranea*. Surprisingly though, on-going genome sequencing of a cnidarian recently showed that *Hydra magnipapillata* has at least 3 TLR genes. These expressed sequence tags (accession number CB889349, CN630662 and CN630303) encode proteins with a TIR domain and a transmembrane segment, thus establishing them as likely TLR genes. Phylogenetic analysis provides further evidence that these TIR domains cluster with animal TLRs, rather than with the intracellular TIR domain containing genes. While it is conceivable that TLRs were lost during the evolutionary period leading to platyhelminths, as apparent in *Schistosoma* and *Schmidtea,* further genome

Table 1. Distribution of Toll related receptors (TLR) and NF-kB genes in the animal kingdom

Phyla	Representative	Toll	Rel
Cnidaria	*Hydra*	3	-?
Platyhelminth	*Schistosoma*	-?	-?
Nematoda	*Caenorhbiditis*	1	-
Annelida	*Tubifex*	?	1
Mollusca	*Euprymna/Crassostrea*	1?	(+)
Arthropoda	*Drosophila,*	9	3
	Anopheles	10	2
	Tachypleus	1	?
	Homarus	1	?
Echinodermata	*Strongylocentrotus*	2	1
Chordata	*Mus*	12	5
	Homo	10	5

Table 2. Toll like receptor genes in arthropods[†]

Subclass	Infraclass	Order	Suborder	# of genes	Representative
Neoptera	Orthopteroida	Orthoptera	Caelifera	2*	*Schistocerca*
			Caelifera	>1*	*Locusta*
	Paraneoptera	Hemiptera		1*	*Homalodisca*
	Endopterygota	Hymenoptera	Aculeata	5*	*Apis*
		Coleoptera	Polyphaga	1*	*Tribolium*
		Lepidoptera	Glossata	7*	*Bombyx*
			Glossata	1*	*Manduca*
		Diptera	Brachycera	9	*Drosophila*
			Brachycera	1*	*Ceratitis*
			Brachycera	1*	*Glossina*
			Nematocera	4*	*Aedes*
			Nematocera	1*	*Culex*
			Nematocera	3*	*Clogmia*
			Nematocera	10	*Anopheles*

* indicates genome has not been fully sequenced or annotated.
[†]Not included here are one TLR gene each found in a lobster and a horseshoe crab, respectively.

sequencing may yet reveal TLRs within the phylum Platyhelminth.

Phylogeny of TLR genes

The numbers of extracellular LRR arrays in different TLRs are different, making sequence alignment in this region highly subjective. Therefore, we based the phylogenetic relationships among TLRs on the intracellular TIR domain. The TIR domain was identified using the SMART algorithm (Schultz *et al.*, 2000; Letunic *et al.*, 2004) and aligned by Clustal X (Thompson *et al.*, 1997). Neighbor-joining showed that all TLRs from Mollusca and Arthropoda form a clade, except Toll9s from *Drosophila* and *Anopheles*, which cluster with mammalian TLRs. TLRs from Nematoda form a separate clade (Fig. 1). This is consistent with the proposal that innate immunity developed independently in invertebrates and vertebrates (Hughes, 1998; Friedman and Hughes, 2002). Interestingly, the TLRs from Cnidarian cluster with mammalian TLRs (Fig. 1).

Distribution of Rel transcription factors in the animal kingdom

From studies in *Drosophila* and mammals, Toll signaling pathways in innate immunity require key transcriptional factors belonging to the NF-kB (or Rel) family. These proteins are characterized by a RHD domain. Thus far, NF-kB factors have been found in an

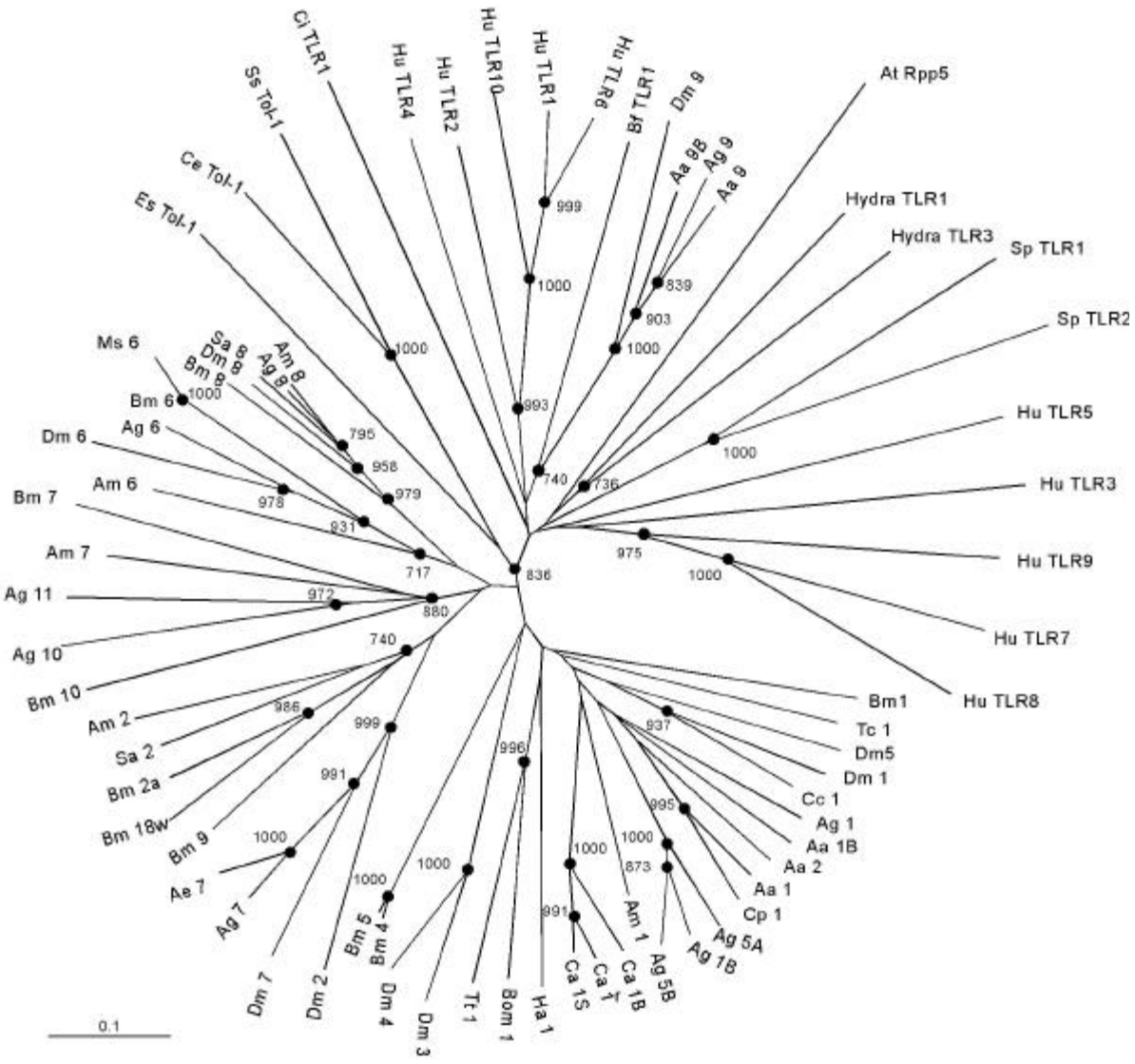

Fig. 1 Phylogenetic relationships among the TLRs. TIR domains were identified using the SMART algorithm and sequences before the conserved motif (F/Y)DA and after the motif FW(E/D) were removed. Alignment was performed by Clustal X, followed by visual inspection. Edited alignment was used for tree generation using 1000 bootstrap iterations. Nodes with bootstrap values >700/1000 are highlighted with filled circles. Scale represents 0.1 amino acid change/residue. Abbreviations are as follows: Hu (human), Dm (*Drosophila*), Aa (*Aedes*), Ag (*Anopheles*), Am (*Apis*), At (*Arabidopsis*), Bf (Branchiostoma), Bm (*Bombyx*), Bom (*Boophilus*), Ca (*Clogmia*), Cc (*Ceratitis*), Ce (*Caenorhbiditis*), Ci (*Ciona*), Cp (*Culex*), Es (*Euprymna*), Ha (*Homarus*), Mx (*Manduca*), Sa (*Schistocerca*), Ss (*Strongyloides*), Tc (*Tribolium*), Tt (*Techypleus*). Due to the high throughput nature of the sequences analyzed, some insect TLR genes may be allelic. For simplicity, arthropod TLRs are indicated by a number after the species abbreviation, while the single TLR genes from *Caenorhbiditis, Strongyloides* and *Euprymna* each are designated as Tol-1. The alignment file is available upon request.

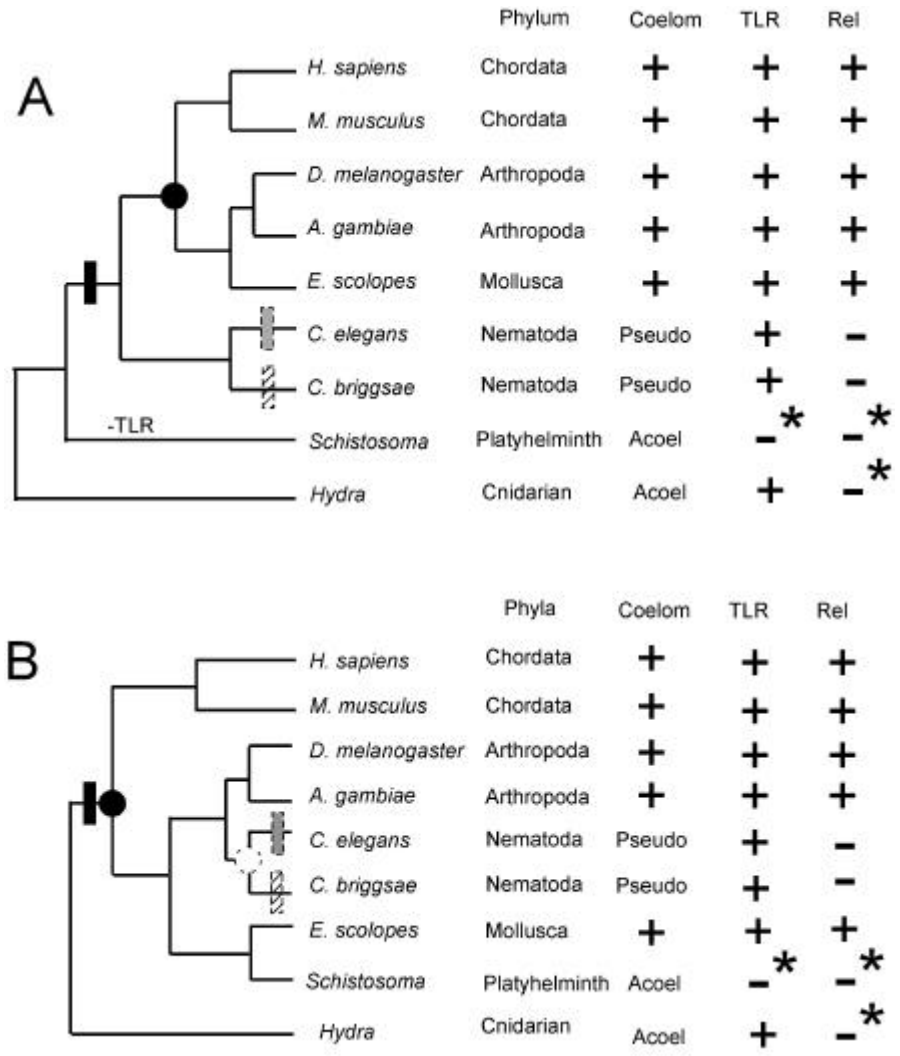

Fig. 2 Coelomata (panel A) versus Ecdysozoa (panel B). If genes shared among organisms with a body cavity are absent in nematodes (black filled circle), they can be easily explained by the Coelomata hypothesis (panel A). The Ecdysozoa hypothesis would need to invoke a loss of genes in the nematodes (dashed open circle). Genes shared among coelomates but absent in one of the two nematode species (filled rectangle) can be explained equally well by either the Coelomata or Ecdysozoa hypotheses (where gene losses are represented by dashed rectangles with different fill patterns for the two nematode species). The distribution of TLR and Rel genes (last two columns in panel A) is consistent with the Coelomata hypothesis. An asterisk indicates the absence of a corresponding gene, though it is conceivable that the gene could be found when the complete genome is sequenced. For simplicity, the *Rel* gene from another coelomate annelid, *Tubifex* is not shown here.

annelid, mollusks, arthropods and vertebrates, all of which possess real body cavities. None have been found in nematodes, platyhelminths or cnidarian (the former being pseudocoelomates while the later two acoelomates; see Table 1 and Fig. 2). This was based on a tBLASTn analysis (Altschul *et al.*, 1997) performed on April 7, 2005, using *A. gambiae* Rel1 and *D. melanogaster* Relish as queries against the updated non-redundant GenBank database.

TLR mediated systemic immunity developed in coelomates: a hypothesis

Among invertebrates, antimicrobial peptide production is the main form of systemic immunity. We argue that such immunity developed in coelomates (Kanzok *et al.*, 2004). Before animals became terrestrial, it would not have been efficient to produce antimicrobial peptide on the surface, where the activity would have been diluted by sea water. When a coelom was developed in an animal, it provided a nutrient-rich and sterile environment that would have been envied by microbes. To protect against internal infection, the first coelomate patched together a signaling pathway that includes the existing TLR genes and may have acquired a novel transcription factor of the NF-kB family. This is consistent with the presence of both TLR and NF-kB genes in the phyla where sufficient genomic sequences are available: Echinodermata, Chordata, Arthropoda and Mollusca. Further, antimicrobial peptides such as defensin have also been found in mammals, arthropods and mollusks.

Although one TLR gene each has been found in *C. elegans* and *C. briggsae*, no NF-kB-like transcription factor has been found in nematodes. Further, genetic evidence with mutations in the TLR gene of *C. elegans* suggests that it is essential for development and is not involved in innate immunity (Pujol *et al.*, 2001).

Coelomata and Ecdysozoa

Traditional classification of metazoans (or Coelomata hypothesis) is based on the presence of a coelom and divides bilaterians into acoels, pseudocoels and coels (Hyman, 1951). Our hypothesis implies that classical division of the animal kingdom based on the presence or absence of a coelom (body cavity) is correct and promoted us to examine the relationships among taxa at the deep root of the evolutionary tree. Both Mollusca and Arthropoda are Protostomia coels and are considered more closely related to each other than they are to a pseudocoel phylum, Nematoda. Likewise, newer morphology-based classifications recognize Protostomia and Deuterostomia, with Protostomia divided into Spiralia and Cycloneuralia, and these place Mollusca and Arthropoda as more closely related to each other than to Nematoda (Nielsen, 2001). A radical view proposed recently separates Mollusca from Arthropoda and groups Arthropoda with Nematoda and others, forming a clade named Ecdysozoa. A separate clade, Lophotrochozoa includes organisms such as annelids, mollusks and platyhelminths. This later view is primarily based on phylogenetic analysis of nuclear 18S ribosomal RNA sequences (Aguinaldo *et al.*, 1997).

Evidence since has been accumulated in favor of (e.g. Dopazo and Dopazo, 2005; Roy and Gilbert, 2005) and against (e.g. Hughes and Friedman, 2004; Steinauer *et al.*, 2005) the Ecdysozoa hypothesis. Instead of examining sequence variations over time, which are subjected to varying evolutionary rates for different genes and homoplasy, recent studies show that large numbers of genes are shared between mollusks and arthropods yet are absent in the nematodes *C. elegans* and *C. briggsae* (Hoa NT et al, unpublished observations). Even taking into consideration gene loss that is known to occur in *C. elegans*, Coelomata theory provides the most parsimonious explanation for the sharing of genes between mollusks and arthropods.

If the Coelomata theory is indeed correct, it would explain with the distributions of TLR and NF-kB genes in the animal kingdom. TLR genes first appeared in cnidarian and were subsequently lost in platyhelminths. The NF-kB transcription factor evolved once in coelomates and was retained in organisms such as annelids, mollusks, arthropods and vertebrates. The combination of TLR and NF-kB transcription factor genes provided the first coelomate a systemic immune response to internal infection by harmful bacteria.

Future perspectives

We argue here that TLR mediated innate immunity developed after the appearance of coelomates. Whether the Imd pathway evolved in the same way remains to be examined. As more and more genome sequences become available, comparative immunology will shed further light on the origin of the systemic innate immune responses.

Interactions between microbes and their animal hosts range from pathogenic, to mutualistic, to symbiotic. Thus far, we have examined host immune responses to microbes as non-self. In invertebrates, symbiotic microbes are essential to the survival and success of their hosts (Rio *et al.*, 2004). In a mutualistic interaction, the mollusk, *Euprymna scolopes* or Hawaiian squid, depends on its microbial partner, *Vibrio fisheri*, to build a light-emitting organ. Recent studies suggested that microbial pattern molecules are essential to the generation of this organ (Koropatnick *et al.*, 2004), though how these pattern molecules trigger organ development remains a mystery. Whether Toll or Imd, both or additional pathways are involved in the formation of this unique organ or even other bacterial associated structures in invertebrates are questions which also remain unsolved. It is worth noting however that recent studies in mammals have shown that commensal microbes can activate host TLR receptors which in turn are required to maintain intestinal homeostasis (Rakoff-Nahoum *et al.*, 2004). In summary, understanding how microbial organisms interact with their hosts will provide

insights into the innate immunity of invertebrates and could lead to practical applications such as blocking the transmission of vector-borne pathogens.

Acknowledgements

Research in L Zheng's laboratory is supported by an NIH grant (AI43053). L Zhang and H Lin received generous support from the Chinese government.

References

Aguinaldo AM, *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387: 489-493, 1997.

Altschul SF, *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402, 1997.

Barillas-Mury C, Charlesworth A, Gross I, Richman A, Hoffmann JA, Kafatos FC. Immune factor Gambif1, a new rel family member from the human malaria vector, *Anopheles gambiae*. EMBO J. 15: 4961-4701, 1996.

Choe KM, Lee H, Anderson KV. *Drosophila* peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. Proc. Natl. Acad. Sci. USA 102: 1122-1126, 2005.

Christophides GK, *et al*. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298: 159-165, 2002.

Dopazo H, Dopazo J. Genome-scale evidence of the nematode-arthropod clade. Genome Biol. 6: R41, 2005.

Friedman R, Hughes AL. Molecular evolution of the NF-kappaB signaling system. Immunogenetics 53: 964-974, 2002.

Hoffmann JA. The immune response of *Drosophila*. Nature: 426, 33-38, 2003.

Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA. Phylogenetic perspectives in innate immunity. Science 284: 1313-1318, 1999.

Holt RA, *et al*. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298: 129-149, 2002.

Hughes AL. Protein phylogenies provide evidence of a radical discontinuity between arthropod and vertebrate immune systems. Immunogenetics 47: 283-296, 1998.

Hughes AL, Friedman R. Differential loss of ancestral gene families as a source of genomic divergence in animals. Proc. Biol. Sci. 271 (Suppl. 3): S107-109, 2004.

Hyman LH. The Invertebrates, Vol 1 (New York, McGraw-Hill) 1951.

Imler JL, Zheng L. Biology of Toll receptors: lessons from insects and mammals. J. Leukoc. Biol: 75, 18-26, 2004.

Inamori K, Ariki S, Kawabata S. A Toll-like receptor in horseshoe crabs. Immunol. Rev. 198: 106-115, 2004.

Inamori K, Koori K, Mishima C, Muta T, Kawabata S. A horseshoe crab receptor structurally related to *Drosophila* Toll. J. Endotoxin Res. 6: 397-399, 2000.

Kaisho T, Akira S. Pleiotropic function of Toll-like receptors. Microbes Infect. 6: 1388-1394, 2004.

Kanzok SM, Hoa NT, Bonizzoni M, Luna C, Huang Y, Malacrida AR, Zheng L. Origin of Toll like receptor mediated innate immunity. J. Mol. Evol. 58: 442-448, 2004.

Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ. Microbial factor-mediated development in a host-bacterial mutualism. Science 306: 1186-1188, 2004.

Lemaitre B, *et al.* A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the *Drosophila* host defense. Proc. Natl. Acad. Sci. USA 92: 9465-9469, 1995.

Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86: 973-983, 1996.

Letunic I, *et al.* SMART 4.0: towards genomic data integration. Nucleic Acids Res. 32: D142-144, 2004.

Luna C, *et al.* Characterization of three Toll-like genes from mosquito *Aedes aegypti*. Insect Mol. Biol. 12: 67-74, 2003.

Luna C, Wang X, Huang Y, Zhang J, Zheng L. Characterization of four Toll related genes during development and immune responses in *Anopheles gambiae*. Insect Biochem. Mol. Biol. 32: 1171-1179, 2002.

Luo C, Shen B, Manley JL, Zheng L. Tehao functions in the Toll pathway in *Drosophila melanogaster*: possible roles in development and innate immunity. Insect Mol. Biol. 10: 457-461, 2001.

Luo C, Zheng L. Independent evolution of Toll and related genes in insects and mammals. Immunogenetics 51: 92-98, 2000.

Meister S, *et al.* Functional characterization of Rel2 from *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA 2005 (in press).

Nielsen C. Animal evolution: interrelationships of the living phyla, 2nd, Oxford University Press, Oxford, 2001.

Nusslein-Volhard C, Lohs-Schardin M, Sander K, Cremer C. A dorso-ventral shift of embryonic primordia in a new maternal-effect mutant of *Drosophila*. Nature 283: 474-476, 1980.

Ooi JY, Yagi Y, Hu X, Ip YT. The *Drosophila* Toll-9 activates a constitutive antimicrobial defense. EMBO Rep. 3: 82-87, 2002.

Parker JE, *et al.* The *Arabidopsis* downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. Plant Cell 9: 879-894, 1997.

Pasare C, Medzhitov R. Toll-like receptors: linking innate and adaptive immunity. Microbes Infect. 6: 1382-1387, 2004.

Pujol N, Link EM, Liu LX, Kurz CL, Alloing G, Tan MW, Ray KP, Solari R, Johnson CD, Ewbank JJ. A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. Curr. Biol. 11: 809-821, 2001.

Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell 118: 229-241, 2004.

Rio RV, Hu Y, Aksoy S. Strategies of the home-team: symbioses exploited for vector-borne disease control. Trends Microbiol. 12: 325-336, 2004.

Roy SW, Gilbert W. Resolution of a deep animal divergence by the pattern of intron conservation. Proc. Natl. Acad. Sci. USA 102: 4403-4408, 2005.

Schneider DS, Hudson KL, Lin TY, Anderson KV. Dominant and recessive mutations define functional domains of Toll, a transmembrane protein required for dorsal-ventral polarity in the *Drosophila* embryo. Genes Dev. 5: 797-807, 1991.

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 28: 231-234, 2000.

Shin SW, Kokoza V, Ahmed A, Raikhel AS. Characterization of three alternatively spliced isoforms of the Rel/NF-kappa B transcription factor Relish from the mosquito *Aedes aegypti*. Proc. Natl. Acad. Sci. USA 99: 9978-9983, 2002.

Shin SW, Kokoza V, Bian G, Cheon H-M, Kim YJ, Raikhel AS. Rel1, a homologue of *Drosophila Dorsal*, regulates Toll antifungal immune pathway in the female mosquito *Aedes aegypti*. J. Biol. Chem. 280: 16499-16507, 2005.

Steinauer ML, Nickol BB, Broughton R, Orti G. First sequenced mitochondrial genome from the phylum Acanthocephala (*Leptorhynchoides thecatus*) and its phylogenetic position within Metazoa. J. Mol. Evol. 60: 706-715, 2005.

Stoven S, Ando I, Kadalayil L, Engstrom Y, Hultmark D. Activation of the *Drosophila* NF-kappaB factor Relish by

rapid endoproteolytic cleavage. EMBO Rep. 1: 347-352, 2000.

Stoven S, *et al.* Caspase-mediated processing of the *Drosophila* NF-kappa B factor Relish. Proc. Natl. Acad. Sci. USA 100: 5991-5996, 2003.

Tauszig S, Jouanguy E, Hoffmann JA, Imler J-L. Toll related receptors and the control of antimicrobial peptide expression in *Drosophila*. Proc. Natl. Acad. Sci. USA 97: 10520-10525, 2000.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882, 1997.

Xu Y, *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. Nature 408: 111-115, 2000.