TIR-domain-containing protein repertoire of nine anthozoan species reveals coral–specific expansions and uncharacterized proteins

Angela Z. Poole *, Virginia M. Weis

Oregon State University, Department of Integrative Biology, 3029 Cordley Hall, Corvallis, OR 97331, United States

A R T I C L E I N F O

Article history:
Received 23 March 2014
Revised 4 June 2014
Accepted 5 June 2014
Available online 13 June 2014

Keywords:
- TIR domain
- TLR
- IL-1R
- Anthozoa

A B S T R A C T

The intracellular toll/interleukin–1 receptor (TIR) domain plays an important role in vertebrate immunity, but the evolution and function of invertebrate TIR-domain-containing proteins is not fully understood. This study characterized and compared the TIR-domain-containing protein repertoire of nine cnidarians in class Anthozoa. A diverse set of proteins, including MyD88 (myeloid differentiation primary response protein 88), toll-like receptor (TLR)-like, interleukin–1 receptor (IL-1R)-like, and TIR-only proteins are present in the species surveyed. Increased numbers of TIR-only proteins were observed in corals compared to anemones, especially in the Acroporid and Pocilloporid coral families. This expansion could be linked to diversity of the microbial community or in hosts and managing both positive and negative associations. Phylogenetic analysis indicates there are two groups of proteins with IL-1R-like domain architecture in anthozoans that potentially evolved independently of the vertebrate family. Bacterial-like TIR_2 domain proteins are also present, including one sequence with novel domain architecture. Overall this work promotes a better understanding of the anthozoan immune repertoire, which is important in the context learning about ancestral immune pathways and host–microbe interactions.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The intracellular TIR domain is found in many proteins that participate in innate immune pathways, including TLRs, the IL–1R family, and the TIR adaptor proteins (Bowie and O’Neill, 2000; Leulier and Lemaître, 2008). These immune signaling pathways are generally initiated by cytokines or microbial surface molecules binding to the extracellular regions of either TLRs or IL–1R family members. Ligand binding results in conformational changes that allow the TIR domain of the receptor to interact with the TIR domain of adaptor proteins through heterotypic protein–protein interactions (Dunne et al., 2003; O’Neill and Bowie, 2007). The TIR adaptor proteins then recruit downstream signaling molecules to ultimately activate transcription factors (O’Neill and Bowie, 2007; Sims and Smith, 2010).

In vertebrates, there are two major families of TIR-domain-containing receptors, the TLRs and IL–1R (Fig. 1). TLRs are pattern recognition receptors that detect a variety of microbial and viral signatures, such as lipopolysaccharide, dsRNA, flagellin, and non-CpG DNA (O’Neill and Bowie, 2007; Uematsu and Akira, 2008). Most TLRs localize to the plasma membrane, while others are present in the membranes of endosomes (Trinchieri and Sher, 2007). They are characterized by extracellular LRRs, which dictate ligand specificity, while the intracellular TIR domain initiates downstream signaling events through interactions with TIR adaptor proteins (Fig. 1) (Coscia et al., 2011; Takeda et al., 2003). Invertebrates, signaling downstream of ligand binding and TLR activation operates through two distinct mechanisms: MyD88-dependent and MyD88-independent pathways (Kawai et al., 2001; O’Neill, 2003). MyD88 is a TIR adaptor protein and contains a death domain that recruits downstream signaling molecules that lead to activation of NF-κB and AP-1, resulting in expression of pro-inflammatory genes (Akira, 2006; Kenny and O’Neill, 2008; Trinchieri and Sher, 2007). In contrast, MyD88-independent
signaling is unique to TLR3 and TLR4 and utilizes another TIR adaptor protein, TRIF, to activate downstream interferon production (Oshiumi et al., 2003a; Yamamoto et al., 2003). Although they use different adaptor proteins, the end result from either TLR signaling mechanism is the expression of pro-inflammatory cytokines that result in production of pro-inflammatory cytokines that influence T-cell development and functioning (Akira, 2000; Sims and Smith, 2010).

Although TLR and IL-1R signaling is best understood in vertebrate systems, these pathways are also of interest in invertebrates. TLR-like sequences are present in most invertebrate groups, and functional studies have revealed roles for these proteins in both immunity and development (Coscia et al., 2011; Halfon et al., 1995; Rosetto et al., 1995; Tenor and Aballay, 2007; Wang et al., 2011; Yuan et al., 2009). Unlike TLRs, there is little evidence for the existence of true homologs to IL-1R family members in invertebrates. There have been descriptions of IL-1R-like molecules in several phyla, but the evidence is either indirect, or based only on protein domain structure (Beck et al., 2000; Huang et al., 2008; Shinzato et al., 2011). Therefore, the evolutionary history of the Ig-TIR domain combination has not been fully resolved and requires further investigation.

In addition to the TIR-domain-containing receptors, there are several intracellular TIR adaptor proteins that act downstream of TLR or IL-1R family activation. In mammals, there are five TIR-domain adaptor proteins: MyD88 and TRIF (both described above), TIRAP, TRAM, and SARM (O'Neill, 2003). Each of these adaptor proteins plays a specific role in intracellular signaling cascades. MyD88, as previously discussed, is recruited to activated TLRs, IL-1R, and IL-18R at the plasma membrane to propagate downstream signaling cascades (Adachi et al., 1998; Medzhitov et al., 1998; Muzio et al., 1997). In some cases, a second adaptor protein, TIRAP, acts as a bridge between TLRs and MyD88 (Horng et al., 2001; Sheedy and O'Neill, 2007; Yamamoto et al., 2002). In a similar manner to TIRAP, TRAM acts as a bridging protein for TRIF in the TLR MyD88-independent pathway, (Oshiumi et al., 2003b) and MyD88 in IL-18R signaling (Ohnishi et al., 2012). Finally, the only adaptor that negatively regulates TLRs is SARM, which specifically blocks TRIF-dependent signaling (Carty et al., 2006).

Several of the TIR adaptor proteins are present in invertebrates, while others likely evolved after or at the emergence of chordates. MyD88 homologs have been identified in basal metazoans such as sponges and cnidarians and are also present in all invertebrate groups investigated, with the exception of Caenorhabditis elegans (reviewed in Coscia et al., 2011). SARM has homologs in many invertebrate groups, but the function is not entirely conserved. For example, in cephalochordates and horseshoe crab, SARM homologs negatively regulate TLR signaling, while in C. elegans they promote antimicrobial peptide production, and therefore have the opposite function of their vertebrate counterpart (Belinda et al., 2008; Couillault et al., 2004; Yuan et al., 2010). At present, the role of the ancestral SARM protein remains unclear, and it may have been co-opted for a variety of functions throughout evolution. The remaining TIR adaptors, TRIF, TRAM, and TIRAP have been found only in chordates (Wu et al., 2011b).

Research on TIR-domain-containing proteins in cnidarians is of interest in the context of learning about the function of ancestral immune pathways and the role of immunity in cnidarian-dinoflagellate symbiosis. Cnidarians serve as hosts to a diversity of microbes including symbiotic dinoflagellates, bacteria, viruses, and apicomplexans (Kirk et al., 2013; Reshef et al., 2006; Thurrell and Correa, 2011). It is unknown how cnidarians manage this balance of positive and negative associations, and characterization of the cnidarian immune repertoire will allow for the generation of...
testable hypotheses to address this question. Through bioinformatics analysis, TIR-domain-containing proteins have been identified in several cnidarian species. The hydrozoan *Hydra magnipapillata* contains four TIR-domain-containing proteins, two TIR-only proteins (Hm-TR1 and 2) and two that are similar to MyD88 (Miller et al., 2007). Hm-TR1 was found to interact with an LRR-containing protein to promote the production of antimicrobial peptides, which is reminiscent of TLR signaling (Bosch et al., 2009). In addition, *Hydra* MyD88 plays an important role in defense against pathogens, and in the establishment of a commensal bacterial community (Franzenburg et al., 2012). From this research, it is clear TLR signaling has a conserved role in immunity in cnidarians.

The repertoire of TIR-domain-containing proteins has also been characterized in the anemone *Nematostella vectensis* and the coral *Acropora digitifera* (Miller et al., 2007; Shinzato et al., 2011). In comparing the number of TIR-domain-containing proteins among the three cnidian genomes available, *A. digitifera* has a clear expansion compared to *N. vectensis* and *H. magnipapillata* (Shinzato et al., 2011). This trend was also observed for other domains commonly found in immune proteins including the NACH/NB-ARC domain (found in NLRs), TNF domain, and the death domain (found in MyD88 and apoptosis related proteins) (Hamada et al., 2013; Shinzato et al., 2011). Two hypotheses that the authors propose to explain these observations is that immune repertoire complexity is related to either symbiotic capabilities or the colonial lifestyle. Like most corals, *A. digitifera* is colonial and harbors intracellular dinoflagellate symbionts. In contrast, *N. vectensis* and *H. magnipapillata* are solitary and non-symbiotic. Therefore, *A. digitifera* may need a more complex immune repertoire to manage self/non-self recognition and the higher diversity of microbes within the holobiont. All three of these organisms belong to different taxonomic groups within Phylum Cnidaria, and therefore a third hypothesis that emerges from these data is that immune repertoire complexity is attributable to class or order distinctions within the phylum. Since this comparison used only three species, the conclusions that can be drawn from these data are limited. Therefore, the goal of the research presented here is to compare the TIR-domain-containing protein repertoire of nine anthozoan species to determine whether symbiotic state, colonial lifestyle, or order-level distinctions can explain the patterns previously observed.

### 2. Materials and methods

#### 2.1. Cnidarian genomic and transcriptomic resources

To investigate the hypotheses proposed above, nine anthozoan species with publicly available genomic or transcriptomic resources were searched for TIR-domain-containing proteins. These included three anemone species: *Anthopleura elegantissima* (Kitchen and co-workers, in preparation: http://people.oregonstate.edu/~meyere/data.html), *Aiptasia pallida* (Lehnert et al., 2012), and *N. vectensis* (Putnam et al., 2007), and six coral species; *A. digitifera* (Shinzato et al., 2011), *Acropora millepora* (Moya et al., 2012), *Fungia scutaria* (Kitchen and co-workers, in preparation), *Montastrea cavernosa* (Kitchen and co-workers, in preparation), *Pocillopora damicornis* (Taylor-Knowles et al., 2011), and *Seriatopora hystrix* (Kitchen and co-workers, in preparation). These resources represent various developmental stages, symbiotic states, and lifestyles (Table 1). All resources were used without manipulation, with the exception of *A. pallida* for which raw Illumina sequence reads for accession SRR056721 were downloaded from the sequence read archive (SRA) entry for the aposymbiotic C7 transcriptome (http://www.ncbi.nlm.nih.gov/sra/SRX231866) and reassembled using Trinity (Grabherr et al., 2011).

#### 2.2. Sequence searching and verification

To search for TIR-domain-containing proteins, the consensus sequence for Pfam01582 (TIR domain) from the conserved domain database (CDD) (http://www.ncbi.nlm.nih.gov/cdd) (Marchler-Bauer et al., 2013) was used as a query in tBLASTn searches of each anthozoan resource. All BLAST searches were performed using Geneious pro version 5.4.3 (Drummond et al., 2011) with the exception of *A. digitifera* and *N. vectensis*, which were performed through the OIST Marine Genomics Genome Browser (http://marinogenomes.oist.jp/genomes/ncbiblast/search), the Comparative genomics platform for early branching Metazoa (http://compagen.zoologie.uni-kiel.de) (Hemmrich and Bosch, 2008) and the Joint Genome Institute Genome portal (http://genome.jgi-psf.org/pages/blast.jsf?db=Nemv1), respectively. A high E-value cutoff (1 × 10⁻¹) was used in the tBLASTn searches to recover more divergent sequences. Since *A. digitifera* TIR-domain-containing proteins have previously been annotated (Shinzato et al., 2011), the Pfam search feature of the genome browser was also used to retrieve protein sequences for this species. This additional search was used to check the effectiveness of the search method, given that the TIR consensus sequence was generated from data from organisms that are distantly related to cnidarians. To ensure that as many TIR-domain-containing sequences as possible were recovered, representative *A. digitifera* sequences of each protein type (IL-1R-like, TLR-like, TIR-only, and MyD88) were also used as queries for tBLASTn searches of the other eight cnidian resources. To confirm that the sequences obtained contained the TIR domain, nucleotide sequences were translated using the ExPASy translate tool (http://web.expasy.org/translate/) and then used as query sequences for protein domain databases including Pfam (http://pfam.sanger.ac.uk) (Punta et al., 2012), SMART (http://smart.embl-heidelberg.de) (Letunic et al., 2012) and CD-search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer and Bryant, 2004). Only sequences that had hits to the TIR domain with an E-value of less than 1 × 10⁻² in two out of the three databases were considered for further analyses. Sequences were also used as queries for BLASTp searches of the non-redundant protein database and the top cnidian, non-cnidarian, and human hits were recorded. Sequences for each species were aligned and those that were identical or almost identical (less than 5aa different in the TIR domain) were thrown out of the analysis due to the fact that they likely represented artifacts of assembly issues or different isoforms of the same protein. To rule out the fact that for symbiotic organisms some of these sequences could be due to dinoflagellate contamination, the *Symbiodinium minutum* genome (Shoguchi et al., 2013) was screened for TIR-domain-containing proteins and none were found.

#### 2.3. Sequence alignment and phylogenetic analysis

A multiple sequence alignment was performed with the MAFFT v7.017 plug-in through Geneious using the default settings (Drummond et al., 2011; Katoh et al., 2002). The alignment was trimmed to contain only the region of the TIR domain based on SMART annotations. Sequences that did not span the majority (missing more than 20 aa) of the TIR domain, or that did not visually align well were removed and regions with gaps were trimmed manually. ProtTest v 2.4 (http://darwin.uvigo.es/software/prottest_server.html) (Abascal et al., 2005) was run to determine the best model of protein evolution for the alignment, using an overall comparison of statistical frameworks that considered all possible substitution matrices and improvements. All frameworks suggested that LG + G (Le and Gascuel, 2008) was the best model to fit the data. Maximum likelihood phylogenetic analysis was performed using the PhyML 3.0 web server (http://www.atgc-montpellier.fr/phyml)
In the initial searches for TIR-domain-containing proteins, it was discovered that some of the sequences showed greater similarity to the TIR_2 domain (pfam13676) than the TIR domain (pfam01582). Although described as a bacterial domain by Pfam, it is also present in a variety of metazoan sequences (Huang et al., 2008; Zhang et al., 2011). Bacterial TIR domain proteins have received a great deal of attention in the last decade for their role in negative regulation of TLR signaling and promoting bacterial evasion of the immune system (Radhakrishnan et al., 2009; Rana et al., 2011; Yadav et al., 2010). Anthozoan genomic and transcriptomic resources were searched for TIR-domain-containing proteins with the CDD consensus sequence for pfam13676 using tBLASTn searches as previously described above (http://www.ncbi.nlm.nih.gov/cdd) (Marchler-Bauer et al., 2013). Confirmation of TIR_2-domain-containing proteins was difficult because only Pfam, and not SMART, recognizes TIR_2 as a distinct domain. A protein sequence was classified as a TIR_2-domain-containing protein if TIR_2 was significant (E-value less than 1 x 10^{-5}) when searched in pfam (http://pfam.sanger.ac.uk) (Punta et al., 2012) and the hit to pfam01582 (TIR) was less significant or not significant according to CDD. As the classification of some sequences was ambiguous, another criterion was added: whether the sequence had BLASTp hits to known proteins (TIR) or was uncharacterized (TIR_2). The TIR_2-domain-containing proteins were not included in the phylogenetic analysis. A separate analysis was conducted on the TIR-2-domain-containing proteins from a variety of invertebrate and bacterial species, but due to extremely low bootstrap values, this proved to be largely uninformative (data not shown). Confirmation that these sequences were not bacterial contamination was obtained through BLASTp searches of the non-redundant protein database in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) by verifying that the top ten hits were all to eukaryotes. The sequences from genomic resources were also checked for introns to provide further evidence that they were eukaryotic and not bacterial contamination. The top cnidarian, non-cnidarian, and human hits were recorded for each sequence. To rule out the fact that for symbiotic organisms some of these sequences could be due to the dinoflagellate contamination, the S. minutum genome (Shoguchi et al., 2013) was screened for TIR_2-domain-containing proteins and none were found.

### 3. Results

#### 3.1. TIR-domain-containing proteins form clusters on phylogenetic tree according to domain structure

The search for TIR-domain-containing proteins in cnidarians revealed that many of the sequences were incomplete and only contained the C-terminal TIR domain. As this limited the number of sequences that could be assigned a particular protein type, maximum likelihood phylogenetic analysis was conducted in order to make a prediction about the identity of the incomplete sequences. The resulting tree shows that the major groups of TIR-domain-containing proteins form distinct clusters on the tree, most of which have moderate to high bootstrap support (Fig. 2). All MyD88 and TLR-like sequences form well-supported groups, but the IL-1R-like and TIR-only (contain only the TIR-domain) sequences are more ambiguous. Sequences with domain structures similar to IL-1R family members separate into two general groups on the tree. One set forms a well-supported group with the TIR-only proteins, while the other group is more closely related to the MyD88 sequences. For the TIR-only proteins, there are smaller groups with high support, but how these are interrelated is unclear. Overall however, proteins with similar domain structure formed distinct groups on the tree, which allowed the sequences missing the N-terminal portion of the protein to be assigned a predicted protein type. The resulting diversity of domain structures for full-length proteins and the number of TIR-domain-containing proteins in each species based on the phylogenetic tree are shown in Fig. 3A and B.

#### 3.2. Distribution of TIR-domain-containing proteins in cnidarians

Among the nine species surveyed in this study, there is variation in the number of TIR-domain-containing proteins present. The two Acropora species have the highest number of TIR-domain-containing proteins, followed closely by P. damicornis.
and S. hystrix (Fig. 3B). The remaining corals, M. cavernosa and F. scutaria, and all the anemone species had fewer (8–9), but similar numbers of TIR-domain-containing proteins (Fig. 3B). All species searched had MyD88 and IL-1R-like sequences, but TLRs and TIR-only proteins were missing from some of the transcriptomes. Full-length TLRs had one or three LRR domains, and full-length IL-1R-like sequences contained one, two, three, or four extracellular Ig domains (Fig. 3A).

Regardless of the total number of TIR-domain-containing proteins present in each coral species, it is clear there is an expansion of the TIR-only proteins in corals compared to anemones. There is one pair of anemone sequences (Ae_TIR and Ap_TIR) that grouped weakly with the TIR-only coral sequences, but because these are not both complete sequences, it is possible that they are IL-1R-like proteins (Fig. 2). In contrast, corals have multiple TIR-only proteins and this group is particularly abundant in the Acroporids and Pocilloporids (Fig. 3B).

3.3. Anthozoan TIR_2-containing proteins

The search for TIR_2-domain-containing proteins in anthozoans revealed several different domain combinations and variable numbers across species (Fig. 4A and B). For example some proteins contain only the TIR_2 domain, while others contain two TIR_2 domains, SAM domains, ARM domains, death domains, or protein tyrosine kinase domains (Fig. 4A).

4. Discussion

4.1. Number of TIR-domain-containing proteins can be explained in a phylogenetic context, but not by symbiotic capabilities or a colonial lifestyle

Three hypotheses that could explain the differences in the number of TIR-domain-containing proteins in A. digitifera, N. vectensis, and H. magnipapillata are that TIR-domain-containing protein repertoire diversity is related to symbiotic capabilities, a colonial lifestyle (Shinzato et al., 2011), or class/order distinctions. The data suggest that none of these completely explain the patterns of TIR-domain-containing proteins in the anthozoan species described in this study. The similarity in the total number of TIR-domain proteins between F. scutaria, M. cavernosa, and the anemones suggests that symbiotic state or colonial lifestyle alone does not explain the expansion of TIR-domain-containing proteins observed in the A. digitifera genome. It also suggests that the expansion cannot be explained by a simple difference between corals and anemones (order-level distinction) since not all coral species surveyed have
an expanded repertoire of TIR-domain-containing proteins compared to anemones.

The patterns observed may instead be explained by placing them in the context of current anthozoan phylogenies (Fukami et al., 2008; Kitahara et al., 2010). Sea anemones, which are members of Order Actinaria, are basal to Order Scleractinia, which contains the stony corals. The scleractinians are divided into the basal Scleractinia, the complex clades that include the acroporids, and the more derived robust clade, which includes *P. damicornis*, *S. hystrix*, *F. scutaria*, and *M. cavernosa* (Fig. 5) (Kitahara et al., 2010). Within the robust clade, *P. damicornis* and *S. hystrix* both belong to the family Pocilloporidae, a relatively basal group of the robust clade, while *F. scutaria* and *M. cavernosa* are more derived species (Fig. 5) (Fukami et al., 2008; Kitahara et al., 2010). Therefore it appears that for the species in this study, either expansion of TIR-domain-containing proteins occurred in certain coral groups, or alternatively gene loss in *F. scutaria* and *M. cavernosa.* From the limited species sampled in this study it is difficult to definitively say which of these occurred.

### 4.2. Expansion of coral TIR-domain-containing proteins occurred specifically in TIR-only proteins

The results indicate that corals, particularly the Acroporids and the Pocilloporids have increased complexity in their repertoire of TIR-only proteins. Therefore it appears as though this expansion has occurred in both complex corals and in a basal group of the robust clade corals. Many of these sequences cluster by family on the tree and are nearly identical between and within species, suggesting that these may be the result of recent lineage-specific expansions, rather than gene loss. One hypothesis that could explain this complexity is that a greater number of TIR-only proteins gives corals the ability to create a more diverse immune response to manage their microbiome. The coral microbial community has previously been shown to vary across site, environment, and species, and these differences may shape the immune repertoire of a given species (Kooperman et al., 2007; Kvennefors et al., 2010; Lee et al., 2012). Diversity in the innate immune repertoire may allow cnidarians to create a specific immune response in the absence of acquired immunity in order to differentiate between beneficial and pathogenic microbes.

The TIR-only protein expansion in corals is particularly interesting given the domain structure similarity of these proteins to the *Hydra* TIR-only proteins. According to SMART annotations, these proteins have a transmembrane domain, which positions them in the correct location to interact with LRR-only proteins to mimic TLR signaling events. TIR-only proteins have also been found in the sponge *Subrites domuncula* and deuterostome species such as *Branchiostoma floridae*, *Saccoglossus kowalevksi*, and *Strongylocentrotus purpuratus* (Wiens et al., 2005; Wu et al., 2011a). It has previously been suggested that domain shuffling occurred between ancestral eumetazoan LRR- and TIR-domain-containing proteins to create the classical TLRs, which could also be an explanation for the evolutionary fate of the invertebrate TIR-only proteins (Coscia et al., 2011). This is further supported by phylogenetic analysis that demonstrated TIR-only proteins from a variety of invertebrates, including cnidarians, grouped within the human TLRs instead of falling outside of them (Wu et al., 2011a). In addition BLASTp searches with many of the anthozoan TIR-only proteins in this study showed the top non-cnidarian hits as TLRs from a variety of metazoans (Supplemental Table 1).

### 4.3. IL-1R-like molecules

Two general groups of IL-1R-like molecules formed on the phylogenetic tree, but are not well supported by bootstrap values (Fig. 2). Within these groups, there are full-length sequences that contain one, two, three, or four Ig domains (Fig. 3A). Therefore,
cnidarians appear to have a more diverse set of extracellular domains than those found in mammalian IL-1R family members. The presence of two independent groups of IL-1R-like molecules is consistent with previous work suggesting that the Ig-TIR domain combination may have evolved multiple times over the course of evolution (Zhang et al., 2010). Specifically, in phylogenetic analysis comparing TIR-domain-containing proteins across metazoans, Ig-TIR-domain molecules from *B. floridae* and *N. vectensis* did not group with the vertebrate IL-1R family sequences, but instead were scattered across the TIR adaptor proteins and TLRs, suggesting that they evolved independently from the IL-1R family (Zhang et al., 2010). Further evidence that invertebrate and vertebrate Ig-TIR proteins have evolved separately is that no invertebrate sequences used in the study grouped with the vertebrate IL-1R family (Zhang et al., 2010). Therefore, it is possible that some, if not all, cnidarian Ig-TIR-domain proteins are not IL-1R homologs. To further support this, most of the top non-cnidarian and human BLASTp hits for these proteins from both clusters are TLRs from a variety of metazoans, not IL-1R family members (Supplemental Table 1).

4.4. TIR_2-domain-containing proteins in cnidarians

TIR_2-domain-containing proteins with a variety of domain architectures were found in anthozoans. The number in each species was variable and unlike the TIR-domain-containing proteins, there was no obvious pattern to their distribution. Many of the same domain combinations found in this study are also present in the genomes of *B. floridae* (Putnam et al., 2008) and the choanoflagellate *Salpingoeca rosetta* (Fairclough et al., 2013), but the anthozoan domain architectures are not always identical. For example, the SAM-TIR_2-TIR_2 domain combination is very similar to that of the TIR adaptor SARM, but the anthozoan sequences have an extra TIR_2 domain and are missing the second SAM domain. In addition their top BLASTp hits are not to SARMs. Therefore it seems unlikely that these are SARM homologs, despite the similarity in domain structure. Another examples is the PTK-DD-TIR_2 domain architecture, that is thus far unique to cnidarians, and interestingly was only found in anemone species in this study. The identity of these proteins remains unclear as BLASTp searches for most sequences revealed top hits to hypothetically uncharacterized proteins, and did not have significant hits to human proteins.

There is much about these sequences that remains unknown. The function of the TIR_2-domain-containing proteins have not been determined, but the lack of transmembrane domains suggests that these are intracellular proteins that may interact with other TIR-domain-containing proteins to modulate immune signaling cascades. The evolutionary history of these sequences is also unclear. It has been proposed that some metazoan TIR adaptor sequences are the result of lateral gene transfer from bacterial species as they phylogenetically group with bacterial sequences (Zhang et al., 2011). Therefore, future research on these sequences could reveal interesting insight into their function in immunity and evolutionary history.

![Fig. 4.](image1.png)

**Fig. 4.** Distribution and diversity of TIR_2-domain-containing proteins in cnidarians. (A) Diversity of domain structure of full-length TIR_2-domain-containing proteins. ARM, armadillo/beta-catenin-like repeat; SAM, sterile alpha motif domain; Pk_Tyr, protein tyrosine kinase domain (B) number of TIR_2-domain-containing proteins in each cnidarian species.

![Fig. 5.](image2.png)

**Fig. 5.** Basic anthozoan phylogeny indicating the relationship of the species used in the TIR-domain-containing protein repertoire comparison (redrawn from Kitahara et al., 2010).
5. Conclusions

Overall, the results of this study demonstrate the diversity of TIR and TIR-2-domain-containing proteins in anthozoans. One limitation to this study is that the majority of the anthozoan species had only transcriptomes available for analysis. Although the trends observed for TIR-domain-containing protein repertoire seem robust to symbiotic state, developmental stage, and coloniality, further analysis at the genomic level will allow confirmation of the results given that not all TIR-domain-containing proteins are likely to be expressed under a particular treatment or developmental stage. Integrating transcriptomic and genomic data from a species will allow the most complete dataset as it will give information on expression patterns of these genes and also the detection of potential pseudogenes in the genome. Overall, increased species sampling across more treatments and developmental stages will allow for a more complete story of the evolutionary history of the expansion of TIR-domain-containing proteins in corals.

From the data presented here, it can be concluded that cnidarians contain a variety of TIR-domain-containing proteins, and that there is increased complexity of TIR-only proteins in coral groups, particularly in the Acroporids and Pocilloporids. This pattern cannot be explained by symbiotic state or colonial lifestyle, but does align with the anthozoan phylogeny. One hypothesis to explain the complexity of TIR-only proteins in some corals is that the expanded repertoire allows them increased diversity in their immune response to manage the large variety of microbes in their environment and within the coral holobiont. If TIR-domain-containing proteins are involved in the initial detection of microbes, having a more complex array of receptors available on the surface of the cell could allow for fine scale differentiation between beneficial and pathogenic microbes in the absence of acquired immunity. A testable hypothesis is that the diversity of the resident microbiome is directly correlated with the complexity of the TIR-domain-containing protein repertoire in cnidarians. As TIR-domain complexity represents only one aspect of coral immunity, this hypothesis could also be strengthened by determining whether these patterns hold true for other innate immunity proteins, such as NOD-like receptors or TNF-domain-containing proteins. This would provide a more holistic view of anthozoan immune-competing proteins. This work also provides information about the evolution of TIR-domain-containing proteins in cnidarians and other invertebrates. With the exception of MyD88 sequences, most proteins show the greatest sequence similarity to metazoan TIR-only proteins or TLRs, which suggests that IL-1R family members emerged after the divergence of the cnidarians. Lastly, this work also revealed that cnidarians have a diverse repertoire of TIR-2-domain-containing proteins, and further work is needed to determine the function and evolutionary history of these uncharacterized proteins. Overall, the characterization of the TIR-domain-containing protein repertoire in anthozoans lays the groundwork for functional work to elucidate the role of these proteins in immunity, symbiosis, development, or a combination of these processes since they are not mutually exclusive. Overall this will allow for a more complete understanding of the function of this domain in anthozoans.

Acknowledgements

We thank Eli Meyer for the reassembly of the A. pallida transcriptome and the Matz lab for providing the A. millepora transcriptome with reduced redundancy and updated annotations (http://www.bio.utexas.edu/research/matz_lab/matzlab/Data.html). We would also like to extend a special thanks to Nate Kirk and the rest of the Weis lab for feedback during the writing process.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.dci.2014.06.002.

References