Developmental and Comparative Immunology 35 (2011) 1128–1134

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

The defensin gene family expansion in the tick *Ixodes scapularis*

Yanbing Wang, Shunyi Zhu*

Group of Animal Innate Immunity, State Key Laboratory of Integrated Management of Pest Insects & Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, PR China

ARTICLE INFO

Article history:
Received 23 November 2010
Received in revised form 17 March 2011
Accepted 17 March 2011
Available online 20 April 2011

Keywords:
Innate immunity
Antimicrobial peptide
Scapularisin
Gene duplication
Positive selection
Arachnida

ABSTRACT

Ixodid ticks transmit a variety of pathogens by blood feeding. Here, we report computational identification of two multigene families of defensin-like peptides (DLPs) in the *Ixodes scapularis* genome, one corresponding to scapularisin and the other named scasin. Members in the scapularisin family share high sequence similarity to some antibacterial ancient invertebrate-type defensins (AITDs) isolated from primitive insects, arachnids, bivalvia, and fungi whereas scasins represent a novel family of DLPs identified by their overall acidic molecular surface and low sequence similarity to any known defensins. Codon-substitution models support neutral evolution in scapularisins but strong positive selection signal was found throughout the molecules of scasins. The synthetic γ-core region of scapularisin-20 exhibits a wide-spectrum of antimicrobial activity at micromolar concentrations. The finding of extensive gene expansion of DLPs in a vector arachnida may be valuable in the understanding of its role in pathogen transmission.

1. Introduction

Defensins with the cysteine-stabilized α-helix and β-sheet (CSαβ) structural motif are pivotal effector elements of innate immunity that fight mainly against Gram-positive bacterial infection (Bulet and Stocklin, 2005). They are the only one class of antimicrobial peptides (AMPs) conserved across all the arthropods. In comparison with antibacterial classical insect-type defensins (CITDs), antibacterial ancient invertebrate-type defensins (AITDs) have more extensive phylogenetic distribution because they are present not only in primitive insects (e.g. the paleopteran insect *Aeschna* dragonfly), but also in arachnids (spiders, scorpions and ticks) and bivalvia (mussels), even in some fungi (Mygind et al., 2005; Rodriguez de la Vega, 2005; Zhu, 2008a). Their protective roles have been well documented by *in vivo* targeted disruption of the mosquito *Anopheles gambiae* defensin gene causing the death of the mosquitoes after Gram-positive bacterial infection (Blandin et al., 2002).

Tick is an important arthropod vector of human pathogens, such as bunyavirus that causes an emerging infectious disease in China (Stone, 2010) and is emerging as a new resource of defensins (Hill and Wikel, 2005). Some such examples include: amercin from *Amblyomma americanum*, scapularisin from *Ixodes scapularis*, and longicin from *Haemaphysalis longicornis* (Hynes et al., 2005; Todd et al., 2007; Zhou et al., 2007). All these defensins belong to AITDs and have a shorter n-loop than CITDs.

The *I. scapularis* Genome Project (IGP) offers a possibility to identify new defensin genes, which will help improve our understanding of tick immune system. In this work, we recognized two distinct multigene families of defensin-like peptides (DLPs) from the *I. scapularis* genome by using comparative genomics approaches: scapularisins are classified as AITDs, and scasins are distantly related to AITDs. Codon-substitution models detected strong positive selection signals in the scasin family. An antibacterial domain derived from scapularisin-20 was also identified.

2. Materials and methods

2.1. Database search

Search strategy used here is described in Fig. S1 (provided as supplementary material). Firstly, some representatives of known defensins from diverse organisms were used as queries to perform TBLASTN search of the *I. scapularis* genomic sequences (http://www.ncbi.nlm.nih.gov/) under default parameters. Retrieved sequences were filtered by the existence of signal peptides (SPs) (http://www.cbs.dtu.dk/services/SignalP/) and the CSαβ motif. Subsequently, the protein sequences downloaded from the GenBank database were applied to perform ScanProsite over the defensin motif at ExPASy (http://www.expasy.org/tools/scanprosite/), which were also filtered by SPs. To confirm the complete suit of defensins excavated, these putative defensins were again employed for

* Corresponding author at: 1 Beichen West Road, Chaoyang District, Beijing 100101, PR China. Tel.: +86 010 64807112.
E-mail address: Zhusy@ioz.ac.cn (S. Zhu).

0145-305X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dci.2011.03.030

new rounds of TBLASTN and BLASTP search until no new hits appeared.

2.2. Phylogenetic analysis

Multiple sequence alignment of proteins was carried out using the CLUSTAL W program (http://www2.ebi.ac.uk/clustalw/) and further adjusted by hand with reference to the cysteine residue position. Mature peptides (MPs) of aligned sequences were employed to make phylogenetic analysis by MEGA 3.1 using the neighbor-joining method with *p* distances (proportion of differences) (http://www.megasoftware.net/).

2.3. Homology structural modeling of DLPs

Structures were modeled in meta-server (http://bioserv.cbs.cnrs.fr/HTML_BIO/frame_home.html) by comparative modeling (Pons and Labesse, 2009). Experimental structures of two AITDs (fungal plectasin (1ZFU) and mussel Mgd-1 (1FJN)) were selected as templates for modeling all the structures of DLPs described here. Models were evaluated by the Verify3D. Electrostatic potentials were calculated applying the "simplecharge: red −1.8 to white 0.0 to blue 1.8" command and mapped on the model structure surface by SPDBV v4.0.1 (Guex and Peitsch, 1997).

2.4. Maximum likelihood analysis

SPs were removed and only MPs left for maximum likelihood analysis, where five *I. scapularis* sequences (scasin-8/9/10/11/16) as well as the defensins from other organisms (Figs. 1 and 2) were excluded in this study due to the presence of large indels in their C-loop regions and/or low sequence similarity. Nucleotides of 16 scasins were attached as Appendix 1 in supplementary material. For further statistical analysis, we constructed a phylogenetic tree by the neighbor-joining method in MEGA 3.1 based on nucleotide sequences of MPs aligned with CLUSTAL W and further refined manually based on the aligned protein sequences. Codon-substitution models were used to obtain the nonsynonymous-to-synonymous rate ratio ($\omega = dN/dS$) using the CODEML program of the PAML software package (http://abacus.gene.ucl.ac.uk/software/paml.html). The presence of a positively selected rate class is detected by comparing the likelihood of a neutral model with that of a selection model (Yang, 2005; Yang et al., 2000). The Model 0 (M0) assumes one ω for all sites. Two pairs of models make two likelihood ratio tests (LRTs) by M1a (nearly neutral model) against M2a (positive selection model) and M7 (β distribution model) against M8 (β and ω model). Two LRTs that calculated $2\Delta l$ (twice the log likelihood difference) were compared against χ^2 with $df = 2$ with critical values 5.99 at 5% and 9.21 at 1% significance levels. Upon detection of positively selected signals, posterior probabilities for the positive selection class under models M2a and M8 were calculated by the Bayes empirical Bayes (BEB) approach (Yang et al., 2005).

2.5. Chemical synthesis

An amidated peptide corresponding to the C-terminal 14 residues of scapularisin-20 (CGGFLKKTCICVMK) was chemically synthesized by Xi'an Huachen Bio-Technology Co., Ltd. (Xi'an, China). The peptide has >95% purity.

2.6. Antibacterial assays

Inhibition zone assays and the calculation of lethal concentration (C_L) were performed according to Hultmark's method (Hultmark, 1998). Microorganisms used include: Gram-positive bacteria: *Micrococcus luteus*, *Bacillus megaterium*, *Bacillus* sp. DM-1 and *Bacillus subtilis*; Gram-negative bacteria: *Agrobacterium tumefaciens*, *Pseudomonas putida*, *Pseudomonas aeruginosa* and *Stenotrophomonus* sp. YC-1.

In the bactericidal assay, bacterial suspension and the peptide of one-fold C_L were added to a 96-well plate with continuous shaking at 37 °C. Every half an hour, the optical density at 595 nm (OD_{595}) was measured with a microplate reader.

2.7. Hemolytic assay

Hemolytic activity of the peptide against fresh mouse blood was assayed according to the standard method (Tossi, 1997). Absorbance was measured at 570 nm. A_{blank} and A_{prep} were evaluated in the absence or presence of the peptide. 100% hemolysis (A_{tot}) was obtained in the presence of 1% Triton X-100. The percentage of hemolysis is determined as $(A_{pep} - A_{blank})/(A_{tot} - A_{blank}) \times 100$.

3. Results and discussion

3.1. Search for new I. scapularis DLPs

To identify a complete suite of DLPs in *I. scapularis*, we undertook two main steps to search the genome database (Fig. S1). Firstly, some known defensins from ticks and other diverse organisms were used as queries to identify orthologues in *I. scapularis* by TBLASTN. Several examples used here include scapularisin, longicin, amercin, drosomycin, plectasin, and MGD-1 (Charlet et al., 1996; Fehlbaum et al., 1994; Hynes et al., 2005; Mygind et al., 2005; Todd et al., 2007; Zhou et al., 2007). Secondly, pattern search of the CSαβ motif (Zhu, 2008a) by ScanProsite was used to filter protein sequences. This strategy allowed us to identify peptides with a typical CysXaaXaaXaaCys and CysXaaCys motif, where Xaa is any amino acid. As a result, we recognized 25 DLPs with significant sequence similarity to scapularisin (scapularisin-1-25, following the name of the first defensin discovered in *I. scapularis* (Hynes et al., 2005) and 21 DLPs (herein called scasin-1-21) (Table 1 and Table S2). A close comparison of sequences of nucleotides and amino acids revealed that scapularisin-1, -5 and -14 all are encoded by different genes but their mature peptides are identical. In this case, we termed scapularisin-1, -5, -14 as scapularisin-1^1, -5^1, -14^1 and scapularisin-1^2, -5^2, -14^2, respectively, to indicate their differences at the nucleotide level. For these highly similar sequences, polymorphism within species rather than gene duplication is possible given several ticks have been used in preparing genomic DNA for sequencing (Dr. Jason Meyer, personal communication).

To confirm the reliability of our computational prediction, we also searched the *I. scapularis* EST database, which identified six transcripts corresponding to the predicted DLPs (scapularisin-1, 5, 6, 7, 16 and scasin-13). The absence of transcripts for other predicted peptides could be due to their expression depending upon suitable microbial challenges (Ferrandon et al., 2007; Tian et al., 2008), or alternatively, being associated with different developmental stages of *I. scapularis*, as observed in the *Drosophila melanogaster* antifungal drosomycin (Tian et al., 2008), although pseudogene or wrong prediction resulted from computational methods cannot be completely excluded. Of the 46 peptides described here, only 11 were previously annotated as putative defensins and other 35 are firstly identified here.

In addition, due to some genomic sequences unassembled completely, full-length sequences of 11 peptides (scapularisin-26 to -36) can not be predicted (Fig. S4). For example, scapularisins-26 to -28 only have SPs; scapularisins-29 to -31 have incomplete MPs; scapularisins-32 has a stop codon that breaks the CSαβ motif;

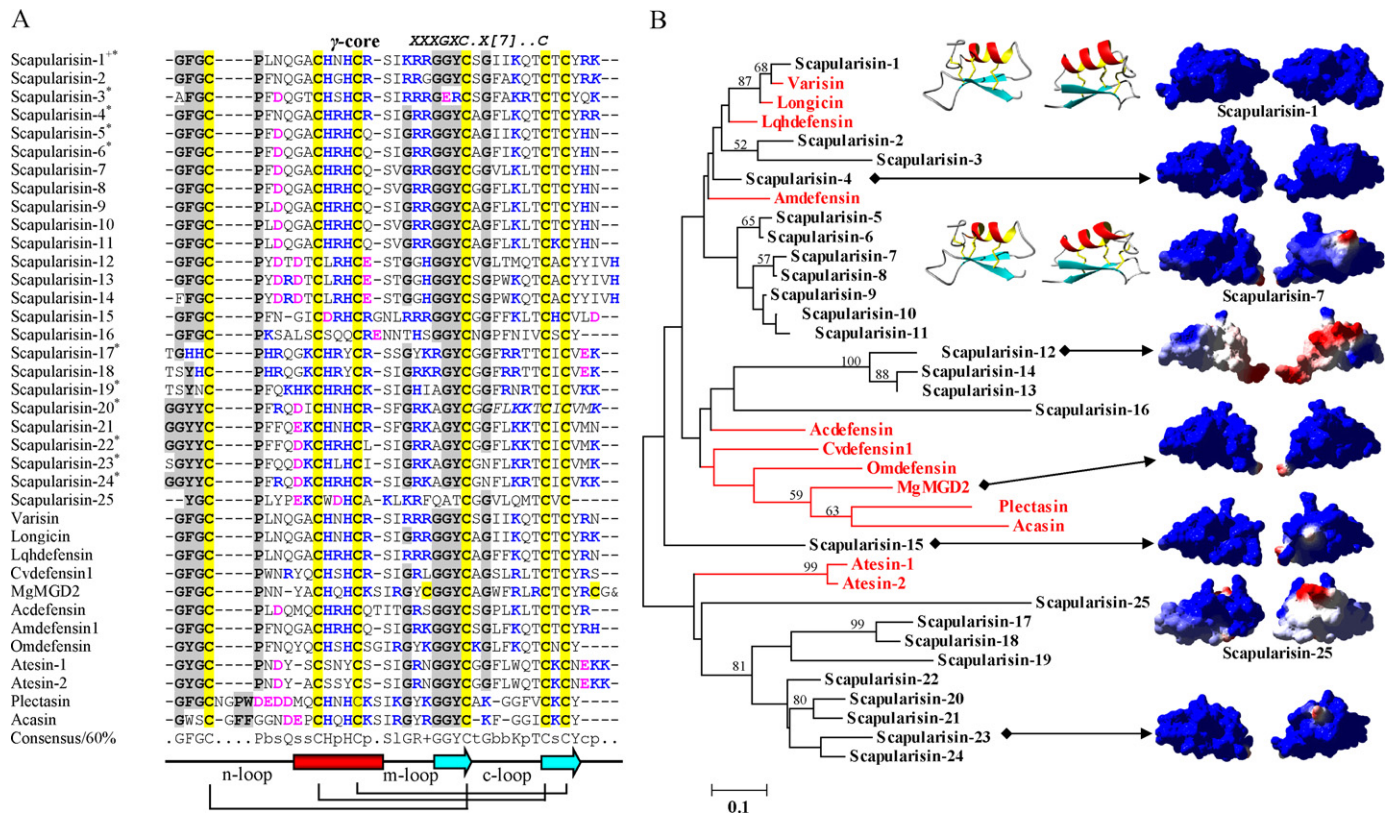

Fig. 1. The scapularisin family of defensins. (A) Multiple sequence alignment of scapularisins and known AITDs. The γ-core is highlighted on the top of sequences. Secondary structures are denoted in cartoon (cylinder: α-helix; arrow: β-strand). The names of scapularisins with complete furin motif (RVRR) were added a *. Negatively charged residues are boldfaced in pink and positively charged amino acids in blue. Cysteines are shaded in yellow, and other relatively conservative amino acids shaded in brown. Consensus was predicted by CHROMA and symbols –, *, l, +, t, a, c, s, p, b, and h respectively represent negative, Ser/Thr, aliphatic, positive, tiny, aromatic, charged, small, polar, big, and hydrophobic. The synthetic fragment is italicized. (B) Phylogenetic tree. Bootstrap values more than 50% are indicated here. Red branches represent known AITDs from other organisms. Blue and red areas indicate positive and negative charges of the molecular surfaces, respectively. Ribbon view of scapularisin-1 and scapularisin-7 are also shown here.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

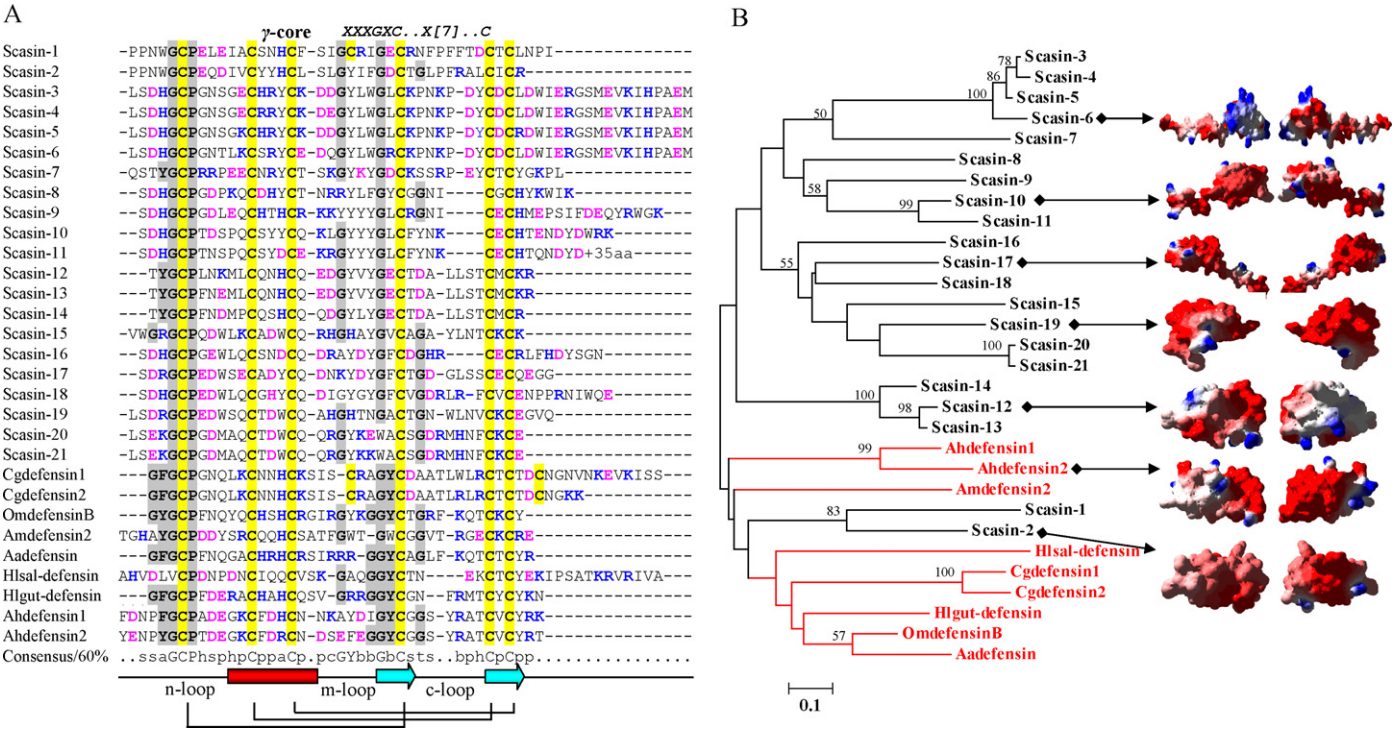

Fig. 2. The scasin family of DLPs. (A) Multiple sequence alignment of scasins and known AITDs; (B) Phylogenetic tree. Color codes and symbol notes used here are the same with those in Fig. 1.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1
Characteristics of scapularisins and scasins.

Name	Accession no. (position)	Mature peptide			
		Size	MW (Da)	NC	γ-Core
Scapularisin-1[+*]	1. EEC08934 [Scapularisin-1[1]][+] 2. ABJB010413606(7870–10124) [Scapularisin-1[2]]	38	4217.90	+7.2	14
Scapularisin-2	ABJB010277423 (696–809)	38	4115.68	+6.2	14
Scapularisin-3[*]	EEC13914	38	4354.95	+6.2	14[b]
Scapularisin-4[*]	ABJB010377501 (3878–6451)	38	4268.91	+7.2	14
Scapularisin-5[+*]	1.[FL] EEC08933 [Scapularisin-5[1]] 2.[MP] ABJB010532949 (799–912)[+] [Scapularisin-5[2]]	38	4146.69	+3.4	14
Scapularisin-6[a+*]	EEC08935 (AAV74387, Scapularisin)	38	4180.71	+3.4	14
Scapularisin-7[+]	ABJB011101692 (1114–1227)	38	4089.64	+3.4	14
Scapularisin-8	ABJB011096562 (55–168)	38	4137.68	+3.4	14
Scapularisin-9	ABJB011065840 (1439–1552)	38	4131.72	+3.4	14
Scapularisin-10	ABJB010379933 (274–387)	38	4117.69	+3.4	14
Scapularisin-11	ABJB011080676 (886–999)	38	4144.76	+4.4	14
Scapularisin-12	ABJB010674815 (106–225)	40	4320.85	+1.6	14
Scapularisin-13	ABJB010563488 (671–790)	40	4429.93	+0.4	14
Scapularisin-14	1. ABJB010360737 (251–370) [Scapularisin-14[1]] 2. ABJB010822197 (776–895) [Scapularisin-14[2]]	40	4520.05	+0.4	14
Scapularisin-15	ABJB010719289 (789–902)	38	4236.91	+4.2	14
Scapularisin-16[+]	EEC17916	38	4046.52	+1	14
Scapularisin-17[*]	ABJB011061725 (123–969)	39	4530.19	+8.7	14
Scapularisin-18	ABJB010535335 (402–518)	39	4605.34	+9.4	14
Scapularisin-19[*]	EEC01374	39	4483.26	+8.7	14
Scapularisin-20[*]	EEC17844	39	4449.25	+6.2	14
Scapularisin-21	ABJB010873090 (296–412)	39	4455.21	+5.2	14
Scapularisin-22[*]	EEC03289	39	4448.31	+6.2	14
Scapularisin-23[*]	ABJB010061346 (1299–2127)	39	4473.31	+5.2	14
Scapularisin-24[*]	ABJB010520452 (1884–2707)	39	4554.38	+9.2	14
Scapularisin-25	ABJB010066055 (8885–8986)	34	3853.59	+2.0	14[b]
Scasin-1	EEC18782	41	4609.29	−2.1	14
Scasin-2	ABJB010977893 (2001–2117)	38	4328.04	−1.0	14
Scasin-3	ABJB010442534 (17–181)	55	6274.05	−3.5	14
Scasin-4	ABJB010586641 (275–439)	55	6307.12	−2.8	14
Scasin-5	ABJB010739942 (1029–1193)	55	6316.13	−0.5	14
Scasin-6	ABJB010832625 (295–459)	55	6350.19	−1.8	14
Scasin-7	ABJB010767832 (3614–3739)	42	4879.47	3.7	14
Scasin-8	ABJB010495451 (739–855)	39	4497.06	2.4	11
Scasin-9	ABJB010461586 (5334–5477)	37	5758.43	−0.3	11
Scasin-10	ABJB010760265 (3940–4068)	43	5185.66	−1.8	11
Scasin-11	ABJB010086782 (662–886)	75	8263.04	−2.8	11
Scasin-12	ABJB010586341 (187–297)	37	4192.82	−1.0	14
Scasin-13[+]	ABJB010272194 (18165–18275)	37	4227.78	−3.0	14
Scasin-14	ABJB010907808 (7433–7540)	36	4055.55	−3.0	14
Scasin-15	ABJB010783265 (709–822)	38	4285.93	3.2	14
Scasin-16	ABJB011065775 (1923–2048)	42	4861.16	−4.5	11
Scasin-17	ABJB010453558 (876–995)	40	4394.56	−8.3	14
Scasin-18	ABJB010836816 (9493–9630)	46	5367.93	−3.8	14
Scasin-19	ABJB011050833 (1211–1333)	41	4510.91	−2.8	14
Scasin-20	ABJB010726488 (57–173)	39	4503.05	−1.0	14[b]
Scasin-21	ABJB010067897 (766–882)	39	4502.11	1.0	14[b]

Note: MW: molecular weight; NC: net charge, estimated at pH 7.0 using protein calculation V3.3 (http://www.scripps.edu/~cdputnam/protcalc.html); [+]: transcripts recorded in GenBank, [a]: proteins currently annotated in reference, [b]: without a GXC motif and [*]: proteins contain a complete furin motif (RVRR).

scapularisins-33 to -35 have no intact CSαβ motif; scapularisin-36 has only half sequence of the CSαβ motif.

3.2. Characteristics of *I. scapularis* DLPs

Most of the *I. scapularis* DLPs described here contain a complete γ-core (Zhu, 2008a) of 11 or 14 amino acids with a GXC motif (Fig. 1A, Table 1). Similar to insect defensins and plectasin, most of scapularisins have a conserved precursor organization comprising SP, propeptide (PP) and MP (Fig. S2). PPs are rich in acidic amino acids and will be cut off during post-translational processing for the release of MPs. Most of MPs are composed of 38-40 residues and always start with GFG, GYG or GYY, a typical sequence characteristics of AITDs. Scapularisins share similar sequences with up to 40% identity among family members. Overall, these peptides carry net positive charges with considerable variations from +0.4 to +9.4. From the multiple sequence alignment, a consensus for scapularisins can be drawn as: $CysProXaa_{4-5}CysXaa_3CysXaa_{9-10}CysXaa_{5-7}CysXaa_1Cys$. One additional sequence feature for most of members in the scapularisin family is that they contain a glycine-rich motif (e.g. $GlyXaa_2GlyGlyXaa$) between the third and fourth cysteine. Cluster analysis provides evidence supporting an orthologous relationship between scapularisins and AITDs (Fig. 1B, Table S1).

Scasins share low sequence similarity to scapularisins and other AITDs (Fig. 2, Fig. S3), but all these peptides have six conserved cys-

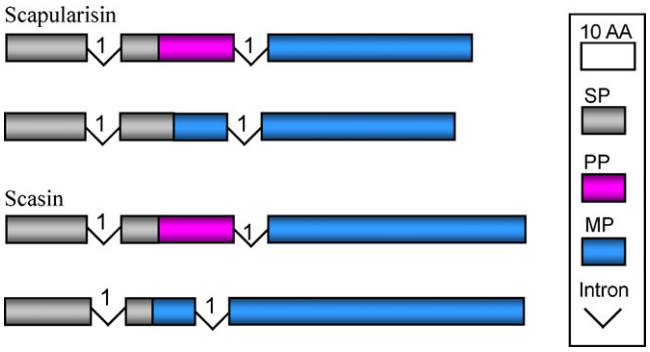

Fig. 3. Gene structures of DLPs. SP: signal peptide; PP: propeptide; MP: mature peptide. Numbers above the junctions correspond to the phase of introns.

teines and most members in this family contain a conserved proline in the n-loop and two glycines in their m-loop, as seen in scapularisins. Relative to AITDs, scasins have two distinct features: (1) All the members are rich in acidic residues and some contain a long C-terminus; (2) The m-loop has a high content of aromatic residues. Most scapularisins possess the typical furin cleavage motif (RVRR) for post-translational processing, whereas for the scasin family, it is not sure whether they have a complete furin motif because of incomplete genome assembling. In several scasins, a "R" can be found to precede the mature peptides.

It is also worth mentioning that the assembly errors by residual polymorphism and/or sequencing error should not be excluded for several highly similar sequences because several ticks have been used in preparing genomic DNA for sequencing (Jason Meyer, personal communication). However, 13 scapularisins from the first family all have "full-length" cds and thus in this case the assembly errors should not cause any problems (Jason Meyer, personal communication).

3.3. Conserved structures of I. scapularis DLPs

Homology modeling approach was employed to infer the structural conservation of the newly discovered DLPs. As expected, all these DLPs modeled adopt a common CSαβ fold due to the presence of a conserved cysteine framework (Figs. 1 and 2). Most of DLPs possess a typical amphipathic architecture in their γ-core regions (Yount and Yeaman, 2004), identified by at least one positively charged residue surrounded by several hydrophobic amino acids with high exposure on the molecular surface (data not shown). Although some scasins have a longer C-terminus and shorter c-loop, homology modeling produced satisfactory models in their core regions. Our structural analysis supports that the long C-terminus of scasin-11 separated from the core CSαβ could fold into an α-helical conformation.

3.4. Gene structures of I. scapularis DLPs

Genes encoding *I. scapularis* DLPs exhibit two classes of different exon–intron structures (Fig. 3). The first class contains two phase I introns (i.e. it splits a codon after the first base), one towards the 3′ end of the SP-coding region and the other at the border between the PP- and MP-coding regions. Genes encoding scapularisin-1 to -24 belong to the first class. The second phase I intron of the second class was found toward the 5′ end of the MP-coding region and splits the MP of the precursor without PP. Although the gene structures of scasins were mixed by these two classes, the first phase I intron is conserved throughout these two defensin families, indicating that they both are evolutionarily related.

Table 2
Maximum likelihood estimates of parameters for scapularisins.

Model	l	LRT	Parameters	Positively selected sites
M0	−1307.3176		$\omega = 0.2037$	None
M1a	−1258.9859		$p_0 = 0.4764$, $\omega_0 = 0.0602$ $p_1 = 0.5236$, $\omega_1 = 1.0000$	Not allowed
M2a	−1258.9859	0	$p_0 = 0.4764$, $\omega_0 = 0.0602$ $p_1 = 0.3101$, $\omega_1 = 1.0000$ $p_2 = 0.2135$, $\omega_2 = 1.0000$	None
M7	−1244.2678		$p = 0.3688$, $q = 0.7461$	Not allowed
M8	−1244.2678	0	$p = 0.3688$, $q = 0.7462$ $p_0 = 1.0000$, $\omega = 1.0000$ $p_1 = 0.0000$	None

3.5. Positive selection of I. scapularis DLPs

To investigate whether positive selection has prompted accelerated evolution of two multigenes of *I. scapularis* DLPs following gene duplication, we employed codon-substitution models to estimate the ratio of nonsynonymous-to-synonymous substitution ($dN/dS = \omega$). The results show that the scapularisin family could be evolving neutrally because no positive selection signals were detected by models M2a and M8 (Table 2). However, considering the importance of these AITDs in the immune system, we assume that the duplicated scapularisins are likely subfunctionalized and may have overlapping antimicrobial activities (Scannell et al., 2006), which is needed to be further investigated in the future.

Different from scapularisins, strong positive selection appears to have driven the evolution of the scasin family (Table 3). In the M0 model, the total tree length was 12.01869. Average dS across the tree was 0.129514, and dN across the tree was 0.167303 (ω, $dN/dS = 0.77411$). Sequences used in the PAML analysis and the phylogeny with dN and dS on each branch have been provided in Fig. S5.

Table 3
Maximum likelihood estimates of parameters and sites inferred to be under positive selection for scasins.

Model	l	LRT	Parameters	Positively selected sites
M0	−1247.3705		$\omega = 0.7741$	None
M1a	−1171.4138		$p_0 = 0.3195$, $\omega_0 = 0.0328$ $p_1 = 0.6805$, $\omega_1 = 1.0000$	Not allowed
M2a	−1162.7083	17.411	$p_0 = 0.2722$, $\omega 0 = 0.0261$ $p_1 = 0.2432$, $\omega_1 = 1.0000$ $p_2 = 0.4846$, $\omega_2 = 2.9939$	10E*, 11I**, 12A*, 14S**, 19S**, 20I**, 23R**, 30F**, 31P *, 39L **
M7	−1168.3322		$p = 0.0444$, $q = 0.0165$	Not allowed
M8	−1160.6947	15.275	$p = 0.1374$, $q = 0.2484$ $p_0 = 0.4865$, $\omega = 2.5953$ $p_1 = 0.5135$	8E*, 10E*, 11I**, 12A*, 14S**, 19S**, 20I**, 23R*, 26E*, 30F**, 3 1P**, 33F*,34T*, 37T*, 39L**

Note: Positively selected sites identified by the Bayes empirical Bayes (BEB) method under models M2a and M8 with posterior probabilities (p)≧0.95 and those with (p)≧0.99 are indicated by * and **. Amino acids underlined exclusively belong to M8 model.

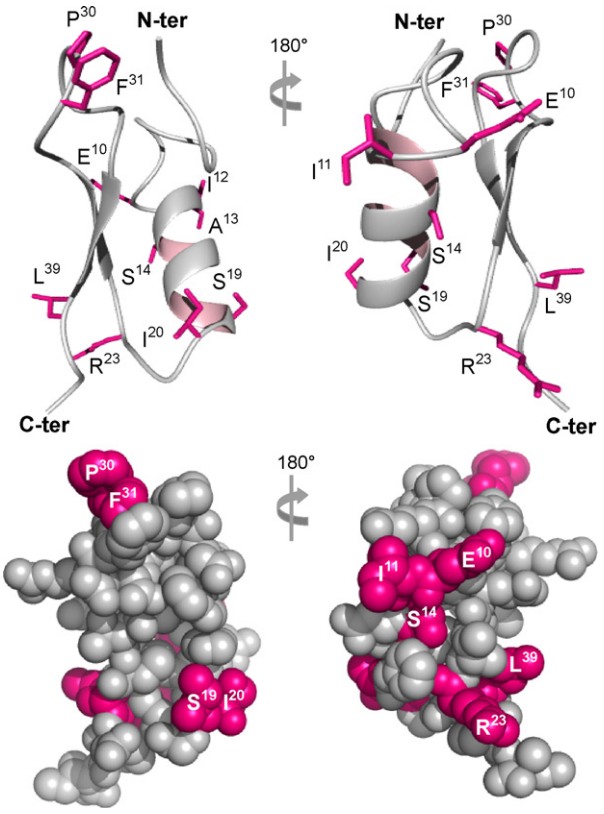

Fig. 4. Mapping positively selected sites onto the scasin-6 structure. Side chains of positively selected amino acids are shadowed in hot pink.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Although M1a does not allow for sites with $\omega > 1$, the selective model M2a with an additional site class estimated the ω ratio to be 2.99, LRT is 14.32, which is much greater than the critical value from a χ^2 distribution with $df = 2$. This suggests that M2a fits the data better than M1a and thus indicates the existence of positively selected sites with $\omega > 1$. Additional LRT performed by comparing the $2\Delta l$ values between M7 and M8 led to consistent results. M2a and M8 identified 10 and 15 positively selected sites, respectively, which are distributed throughout the whole molecules. The commonly identified sites by these two models include E^{10}, A^{12}, S^{14}, S^{19}, I^{20}, R^{23}, F^{30}, P^{31}, T^{37}, and L^{39} (residues numbered according to scasin-1). Mapping these positive selection sites on the structure of scasin-1 reveals that most of them are clustered in one side of the molecular surface (Fig. 4). To rule out possible impact of gene conversion on the PAML results, we analyzed the sequences by GENECONV (http://www.math.wustl.edu/~sawyer/geneconv/#progfiles) and found no signal in favor of such events occurring ($p = 0.1584$, >0.05).

Table 4
Lethal concentration (C_L) of S-20(26–39) against different microbial strains.

Strain of bacteria	C_L(μM)	
Gram⁻ bacteria		
Agrobacterium tumefaciens	31.16	($y = 0.3322x + 0.1501$, $R^2 = 0.9717$)
Pseudomonas putida	16.66	($y = 1.4159x + 0.1333$, $R^2 = 0.9683$)
Pseudomonas aeruginosa	14.80	($y = 1.4117x + 0.2075$, $R^2 = 0.9989$)
Stenotrophomonas sp. YC-1	33.06	($y = 0.7972x + 0.0367$, $R^2 = 0.9478$)
Gram⁺ bacteria		
Bacillus sp. DM-1	13.26	($y = 2.0553x - 0.0645$, $R^2 = 0.9916$)
Bacillus subtilis	20.97	($y = 1.2456x + 0.0631$, $R^2 = 0.9586$)
Bacillus megaterium	10.71	($y = 2.6740x + 0.4170$, $R^2 = 0.9701$)
Micrococcus luteus	17.22	($y = 1.3820x + 0.1250$, $R^2 = 0.9975$)

It is known that genes involved in the immune defense system typically show a faster rate of amino acid substitutions than other genes (Sackton et al., 2007). Such accelerated evolution has been observed in several insect-derived defensins (Tennessen, 2005; Viljakainen and Pamilo, 2008; Zhu, 2008b). In scasins, strong positive selection appears to act on the whole molecule, an evolutionary phenomenon initially observed in *Conus* conotoxins (Duda and Palumbi, 1998). Rapid evolution in these toxins is considered as substantial variation in venom effectiveness against particular prey types. In addition, a logistic model correlating the substitution rates of each residue with their surface accessibility indicates that the probability of natural substitutions occurring in the fully exposed residue is 2.6–3.5 times greater than that of substitutions occurring in buried residues (Kini and Chan, 1999). A similar case is also present in scasins, in which most of positively selected sites were found to have more than 30% accessibility. These observations therefore provide evidence for possible involvement of this unusual DLPs family in the tick immune system. Our opinion is further supported by the finding that some anionic peptides constitute part of the innate immune system (Fales-Williams et al., 2002; Harris et al., 2009).

3.6. Biological activity of scapularisin-20(26–39)

Previous studies have shown that the functional region of insect defensins is primarily located in the C-terminal β-sheet domain, called the 'γ-core motif' (Yount and Yeaman, 2004; Zhu, 2008a). Several synthetic γ-core-derived peptides from tenecin 1, mussel defensin, and navidefensin2-2 were found to possess a wide spectrum of antimicrobial activity (Gao and Zhu, 2010; Lee et al., 1998; Romestand et al., 2003). To identify such functional motif in these newly discovered DLPs, we synthesized a fragment corresponding to the γ-core of scapularisin-20 (named scapularisin-20(26–39) and evaluated its antibacterial activity. Overall, this peptide shows slightly more potency to Gram-positive bacteria (C_L values between 10 to 20 μM) (Table 4) than to Gram-negative bacteria (C_L values between 15 to 33 μM), consistent with the functional feature of

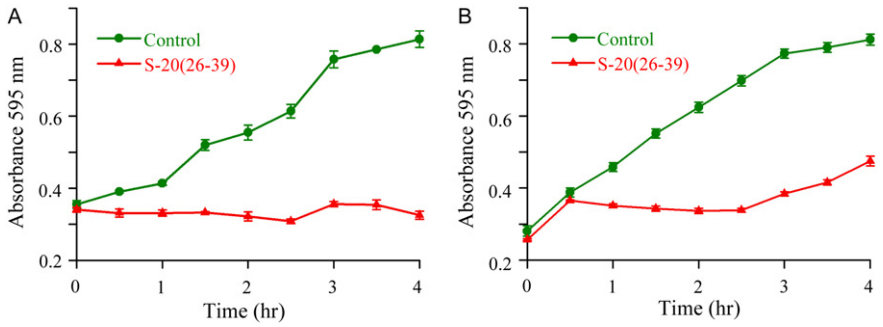

Fig. 5. Kinetics of bacterial killing of scapularisin-20(26–39). *B. megaterium* (A) and *P. aeruginosa* (B) were respectively incubated with scapularisin-20(26–39) at $1 \times C_L$ at 37 °C.

insect defensins (Gao and Zhu, 2010). Subsequently, we assayed bactericidal dynamic of scapularisin-20(26–39) on *B. megaterium* and *P. aeruginosa*, which confirmed time-dependent bactericidal feature of this peptide (Fig. 5). Scapularisin-20(26–39) exhibited nonhemolytic toxicity on fresh mouse blood cells even at relatively high concentration (400 μM) (data not shown), indicating its potential for the design of anti-infective drugs.

4. Conclusion

Bioinformatics-based approaches have succeeded in the discovery of immune-related genes in several eukaryotic model organisms (Li et al., 2008; Neupert et al., 2009). By using integrated computational approaches, we have recently established the first AMP repertoire of a parasitic wasp (Tian et al., 2010). The work presented here is the first report that describes two large DLP multigene families in a member of the subphylum Chelicerata. Given the importance of the black-legged tick *I. scapularis* (Acari: Ixodidae) as an important vector of microbial pathogens, our discovery is of value in enhancing the understanding of tick-host-pathogen interactions and developing anti-infective agents.

Acknowledgements

We thank *Ixodes scapularis* Genome Project (IGP) for providing sequence data for public use and Prof. Chuanling Qiao (Institute of Zoology, Chinese Academy of Sciences, China) for microbial strains used in this study. This work was supported by National Natural Science Foundation of China Grants 30730015 and 30921006 and National Basic Research Program of China (2010CB945300).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.dci.2011.03.030.

References

Blandin, S., Moita, L.F., Kocher, T., Wilm, M., Kafatos, F., Levashina, E.A., 2002. Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the defensin gene. EMBO Rep. 3, 852–856.

Bulet, P., Stocklin, R., 2005. Insect antimicrobial peptides: structures, properties and gene regulation. Protein Pept. Lett. 12, 3–11.

Charlet, M., Chernysh, S., Philippe, H., Hetru, C., Hoffmann, J.A., Bulet, P., 1996. Innate immunity. Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. J. Biol. Chem. 271, 21808–21813.

Duda Jr., T.F., Palumbi, S.R., 1998. Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod Conus. Proc. Natl. Acad. Sci. U.S.A. 96, 6820–6823.

Fales-Williams, A.J., Brogden, K.A., Huffman, E., Gallup, J.M., Ackermann, M.R., 2002. Cellular distribution of anionic antimicrobial peptide in normal lung and during acute pulmonary inflammation. Vet. Pathol. 39, 706–711.

Fehlbaum, P., Bulet, P., Michaut, L., Lagueux, M., Broekaert, W.F., Hetru, C., Hoffmann, J.A., 1994. Insect immunity. Septic injury of *Drosophila* induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. J. Biol. Chem. 269, 33159–33163.

Ferrandon, D., Imler, J.L., Hetru, C., Hoffmann, J.A., 2007. The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. Nat. Rev. Immunol. 7, 862–874.

Gao, B., Zhu, S., 2010. Identification and characterization of the parasitic wasp *Nasonia defensins*: positive selection targeting the functional region? Dev. Comp. Immunol. 34, 659–668.

Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723.

Harris, F., Dennison, S.R., Phoenix, D.A., 2009. Anionic antimicrobial peptides from eukaryotic organisms. Curr. Protein Pept. Sci. 10, 585–606.

Hill, C.A., Wikel, S.K., 2005. The *Ixodes scapularis* Genome Project: an opportunity for advancing tick research. Trends Parasitol. 21, 151–153.

Hultmark, D., 1998. Quantification of antimicrobial activity, using the inhibition-zone assay. In: Wiesner, A., Dumphy, A.G., Marmaras, V.J., Morishima, I., Sugumaran, M., Yamakawa, M. (Eds.), Techniques in Insect Immunology. SOS Publications Fair, pp. 7–103.

Hynes, W.L., Ceraul, S.M., Todd, S.M., Seguin, K.C., Sonenshine, D.E., 2005. A defensin-like gene expressed in the black-legged tick, *Ixodes scapularis*. Med. Vet. Entomol. 19, 339–344.

Kini, R.M., Chan, Y.M., 1999. Accelerated evolution and molecular surface of venom phospholipase A2 enzymes. J. Mol. Evol. 48, 125–132.

Lee, K.H., Hong, S.Y., Oh, J.E., Kwon, M., Yoon, J.H., Lee, J., Lee, B.L., Moon, H.M., 1998. Identification and characterization of the antimicrobial peptide corresponding to C-terminal beta-sheet domain of tenecin 1, an antibacterial protein of larvae of *Tenebriomolitor*. Biochem. J. 334, 99–105.

Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., Williamson, M., Arakane, Y., Verleyen, P., Schoofs, L., Schachtner, J., Grimmelikhuijzen, C.J., Park, Y., 2008. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res. 18, 113–122.

Mygind, P.H., Fischer, R.L., Schnorr, K.M., Hansen, M.T., Sonksen, C.P., Ludvigsen, S., Raventos, D., Buskov, S., Christensen, B., De Maria, L., Taboureau, O., Yaver, D., Elvig-Jorgensen, S.G., Sorensen, M.V., Christensen, B.E., Kjaerulff, S., Frimodt-Moller, N., Lehrer, R.I., Zasloff, M., Kristensen, H.H., 2005. Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. Nature 437, 975–980.

Neupert, S., Russell, W.K., Predel, R., Russell, D.H., Strey, O.F., Teel, P.D., Nachman, R.J., 2009. The neuropeptidomics of *Ixodes scapularis* synganglion. J. Proteomics 72, 1040–1045.

Pons, J.L., Labesse, G., 2009. @TOME-2: a new pipeline for comparative modeling of protein–ligand complexes. Nucleic Acids Res. 37, W485–W491.

Rodriguez de la Vega, R.C., Possani, L.D., 2005. On the evolution of invertebrate defensins. Trends Genet. 21, 330–332.

Romestand, B., Molina, F., Richard, V., Roch, P., Granier, C., 2003. Key role of the loop connecting the two beta strands of mussel defensin in its antimicrobial activity. Eur. J. Biochem. 270, 2805–2813.

Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D., Clark, A.G., 2007. Dynamic evolution of the innate immune system in *Drosophila*. Nat. Genet. 39, 1461–1468.

Scannell, D.R., Byrne, K.P., Gordon, J.L., Wong, S., Wolfe, K.H., 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 440, 341–345.

Stone, R., 2010. Rival teams identify a virus behind deaths in central China. Science 330, 20–21.

Tennessen, J.A., 2005. Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection. J. Evol. Biol. 18, 1387–1394.

Tian, C., Gao, B., Rodriguez Mdel, C., Lanz-Mendoza, H., Ma, B., Zhu, S., 2008. Gene expression, antiparasitic activity, and functional evolution of the drosomycin family. Mol. Immunol. 45, 3909–3916.

Tian, C., Gao, B., Fang, Q., Ye, G., Zhu, S., 2010. Antimicrobial peptide-like genes in *Nasonia vitripennis*: a genomic perspective. BMC Genomics 11, 187.

Todd, S.M., Sonenshine, D.E., Hynes, W.L., 2007. Tissue and life-stage distribution of a defensin gene in the Lone Star tick *Amblyomma americanum*. Med. Vet. Entomol. 21, 141–147.

Tossi, A., Scocchi, M., Zanetii, M., Gennaro, R., Storici, P., Romeo, D., 1997. An approach combining rapid cDNA amplification and chemical synthesis or the identification of novel, cathelicidin-derived, antimicrobial peptides. In: Shafer, W.M. (Ed.), Antibacterial Peptide Protocols, vol. 78. Humana Press, Tokowa, NJ, USA, pp. 133–155.

Viljakainen, L., Pamilo, P., 2008. Selection on an antimicrobial peptide defensin in ants. J. Mol. Evol. 67, 643–652.

Yang, Z., 2005. The power of phylogenetic comparison in revealing protein function. Proc. Natl. Acad. Sci. U.S.A. 102, 3179–3180.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Yang, Z., Wong, W.S., Nielsen, R., 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22, 1107–1118.

Yount, N.Y., Yeaman, M.R., 2004. Multidimensional signatures in antimicrobial peptides. Proc. Natl. Acad. Sci. U.S.A. 101, 7363–7368.

Zhou, J., Liao, M., Ueda, M., Gong, H., Xuan, X., Fujisaki, K., 2007. Sequence characterization and expression patterns of two defensin-like antimicrobial peptides from the tick *Haemaphysalis longicornis*. Peptides 28, 1304–1310.

Zhu, S., 2008a. Discovery of six families of fungal defensin-like peptides provides insights into origin and evolution of the CSαβ defensins. Mol. Immunol. 45, 828–838.

Zhu, S., 2008b. Positive selection targeting the cathelin-like domain of the antimicrobial cathelicidin family. Cell. Mol. Life Sci. 65, 1285–1294.