МОЛЕКУЛЯРНАЯ БИОЛОГИЯ, 2012, том 46, № 4, с. 570–575

ГЕНОМИКА. ТРАНСКРИПТОМИКА

УДК 577.21:575.113.082.261

MOLECULAR CLONING AND CHARACTERIZATION OF A NEW cDNA SEQUENCE ENCODING A VENOM PEPTIDE FROM THE CENTIPEDE *Scolopendra subspinipes* MUTILANS

© 2012 Wanhong Liu[1]*, Feng Luo[2]*, Jing He[1, 3], Zhijian Cao[4], Lixia Miao[1]**

[1]School of Medicine, Wuhan University, Wuhan, 430071, China
[2]School of Medicine, Jianghan University, Wuhan, 430056, China
[3]School of Life Science, Jianghan University, Wuhan, 430056, China
[4]College of Life Science, Wuhan University, Wuhan, 430072, China

Received April 15,2011
Accepted for publication July 11, 2011

Many studies have been performed on venomous peptides derived from animals. However, little of this research has focused on peptides from centipede venoms. Here, a venom gland cDNA library was successfully constructed for the centipede *Scolopendra subspinipes* mutilans. A new cDNA encoding the precursor of a venom peptide, named SsmTx, was cloned from the venomous gland cDNA library of the centipede *S. subspinipes* mutilans. The full-length SsmTx cDNA sequence is 465 nt, including a 249 nt ORF, a 45 nt 5' UTR and a 171 nt 3' UTR. There is a signal tail AATAAA 31 nt upstream of the poly (A) tail. The precursor nucleotide sequence of SsmTx encodes a signal peptide of 25 residues and a mature peptide of 57 residues, which is bridged by two pairs of disulfide bonds. SsmTx displays a unique cysteine motif that is completely different from that of other venomous animal toxins. This is the first reported cDNA sequence encoding a venom peptide from the centipede *S. subspinipes* mutilans.

***Keywords:* *Scolopendra subspinipes* mutilans, cDNA, venom peptide, cysteine motif.**

INTRODUCTION

Centipedes are elongated and dorso-ventrally flattened arthropods that belong to the class Chilopoda [1]. They are comprised by approximately 3300 species belonging to five orders. Centipedes are known to exist on all continents except Antarctica, with the greatest diversity occurring in the tropics and warm temperate regions [2]. The centipede body is divided into a head and trunk. Each trunk segment carries one pair of ambulatory legs except for the hindmost segment, where these structures are modified for mechanical defense and/or for sensory purposes. Another pair of modified legs, the forcipules, is located on the post-cephalic segment [1]. The trochanteroprefemur part of the forcipule contains a venomous gland, which is connected with the exterior through an orifice in the tip of this structure. The venom is used to both subjugate prey and defend against predators [1, 3, 4].

Centipede venom is a complex mixture containing proteases, hyaluronidases, carboxypeptidases, histamine, serotonin, lipids, lipoproteins, esterases, polysaccharides, phospholipases, and alkaline phosphatases [2, 5, 6]. Recently, Peng et al. [7] purified and characterized two novel antimicrobial peptides named scolopin-1 and scolopin-2 from the venom of the centipede *Scolopendra subspinipes* mutilans. Generally, centipede venom is not lethal to humans. The most common clinical signs of centipede venom poisoning are pain, local erythema, and local edema in some cases [8]. Envenomations caused by centipedes are mild and associated with spontaneous healing without complications. Treatment of their stings is supportive [1, 2].

The venom of centipede species is not easily obtained in large quantities, which makes complete sequencing of individual toxins rather difficult. Rates et al. [1] reported N-terminal amino acid sequences and mass spectral data from a venom extraction from the Brazilian centipedes *S. viridicornis* nigra and *S. angulata.* The screening of venomous gland cDNA libraries from centipedes would be of great interest because the complete sequences of the venom peptides could be more easily obtained using this strategy [9–12]. Cloning of toxin genes, in addition to heterologous expression, might be a fundamental tool for obtaining sufficient amounts of material to fully characterize the biological activity of centipede venom peptides. The pathophysiological mechanisms of envenomation also have not yet been elucidated, and such studies are rare in the lit-

* These authors contributed equally to this work.
** E-mail: lixiamczj@sina.com

erature [1], likely because it is difficult to obtain enough venom to perform the tests.

In this investigation, a venom gland cDNA library was successfully constructed from the centipede *S. subspinipes* mutilans. Some cDNAs encoding the precursors of venom peptides were cloned from the venomous gland cDNA library, of which one encodes a new venom peptide with a 25-amino acid residue signal peptide and a 57-amino acid residue mature peptide bridged by two pairs of disulfide bridges. This is the first report of the cloning and characterization of a cDNA encoding a venom peptide precursor from the venom of the centipede *S. subspinipes* mutilans.

EXPERIMENTAL

RNA extraction. The centipede *S. subspinipes* mutilans was collected from Xiangfan, Hubei province, China. The venom glands connected to the first pair forceps of the centipedes were stimulated using a 3 V alterative current. After 48 h, the venom glands were separated and ground in liquid nitrogen. RNA was extracted using the Trizol Reagent ("Gibco"), and mRNA was purified using a PolyA Tract mRNA Kit according to the manual ("Promega").

cDNA library construction and characterization. A cDNA library was constructed from the total RNA extracted from 10 pairs of venomous glands of *S. subspinipes* mutilans. Approximately 5 μg of mRNA were used for synthesis of the double strand cDNA. The cDNA library was constructed using the SMART^{TM} cDNA Library Construction Kit according to the user manual ("Invitrogen"). The recombinant plasmids were transformed into *E. coli* DH5a. The quality of the cDNA library, including the portion of positive clones and the size of the inserted fragment, was characterized by colony PCR using universal M13 primers.

cDNA library screening. PCR was used to screen the venomous gland cDNA library of the centipede *S. subspinipes* mutilans. A specific primer was designed and synthesized to screen for full-length homologs of the Scolopendra 5848.48 Da toxin (SvnTx3), a venom peptide from the Brazilian centipede *S. viridicornis* nigra [1], in the venom gland cDNA library of *S. subspinipes* mutilans. The sequences of the specific forward and reverse primers were 5'-GCCGCATTTACGGGTGGAGAT-3' and 5'-GAGCGGCCGCCCT$_{15}$-3', corresponding to the primer sequence of the first cDNA synthesis and the coding region of the Scolopendra 5848.48 Da toxin AAFTGGD amino acid sequence (position: 2632), respectively.

Sequence analysis. Selected positive clones were sequenced using universal M13 primers. Sequence analysis was performed using Generunr and BioEdi software. Sequence analysis was carried out with

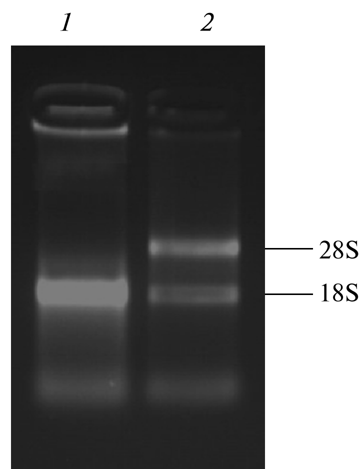

Fig. 1. Agarose gel electrophoresis analysis of the total RNA from the venom gland of the centipede *Scolopendra subspinipes* mutilans. *1* – The centpetide *S. subspinipes* mutilans venom gland; *2* – the cultured HeLa cells.

BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignment of the SsmTx peptide and homologous peptides was carried out using the DNAMAN software.

RESULTS AND DISCUSSION

Hidden break of rRNA from the centipede Scolopendra subspinipes mutilans venom gland

The total RNA extracted from the venomous gland of the centipede *S. subspinipes* mutilans was diluted in double distilled water and electrophoresed in agar gel using a 5 V/cm current for 45 min. As shown in fig. 1, the venom gland total RNA of the centipede *S. subspinipes* mutilans only has a band for 18S rRNA, but not a band for 28S rRNA, which was eliminated by the degradation of the total RNA. The venom gland total RNA agar gel electrophoresis characteristics of the centipede *S. subspinipes* mutilans RNA are completely different from those of tissues or cells of eukaryotic organisms. Generally, agar gel electrophoresis of total RNA from eukaryotic organism tissues or cells shows two bright 28S rRNA and 18S rRNA bands, and the brightness ratio of 28S rRNA to 18S rRNA bands is approximately 2.

Generally, in agar gel electrophoresis the 28S rRNA shows a single band at approximately 4.4 kb. But in several arthropods and other low-grade animals also exists the phenomenon of the 28S rRNA hidden break [13–15]. It was reported that there is no 28S rRNA band in the rRNA of the scorpion venomous gland, instead a weak band shorter than 28S rRNA but longer than the 18S rRNA was found [16]. Total RNA from both the centipede *S. subspinipes* mutilans venom gland and scorpion venom glands had no 28S rRNA

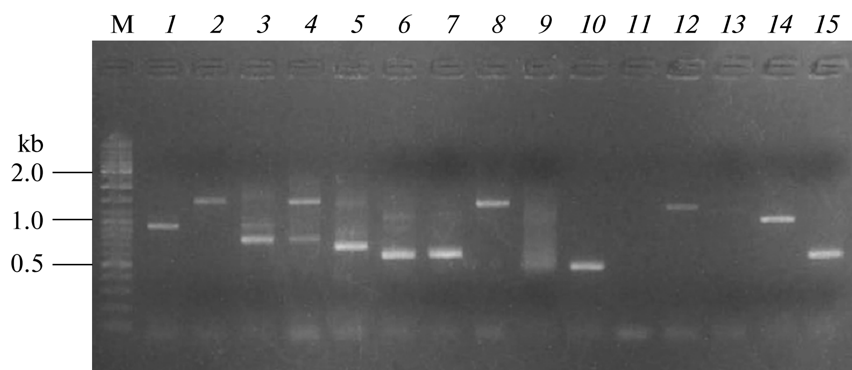

Fig. 2. Agarose gel electrophoresis analysis of inserted cDNA fragments from the constructed venom gland cDNA library of the centipede *Scolopendra subspinipes* mutilans. M DNA molecular weight marker; *1–15* – different colonies of the constructed venom gland cDNA library of the centipede *S. subspinipes* mutilans.

band during agar gel electrophoresis. Thus possibly both the centipede *S. subspinipes* mutilans venom gland and scorpion venom gland total RNAs had 28S rRNA hidden break. It remains to be extensively investigated whether all venomous animal glands share the 28S rRNA hidden break, and the physiological function of this phenomenon deserves to be clarified.

The quality of the centipede venom gland cDNA library

The non-amplified cDNA library from the centipede *S. subspinipes* mutilans venom gland yielded a titer of 1.5×10^6 cfu/mL. Using cDNA plasmids obtained from the venom gland cDNA library of *S. subspinipes* mutilans as templates, one hundred clones were randomly analyzed by colony PCR using M13 primers. The M13 primer oligonucleotides permitted a random screening of the cDNA library. Fig. 2 displays a representative result of PCR detection of the quality of the centipede *S. subspinipes* mutilans venom gland cDNA library. Based on the results of 100 colony PCRs, the recombination efficiency of the constructed venom gland cDNA library of *S. subspinipes* mutilans was more than 93%, and the size of the inserted cDNA fragments ranged from 0.3 to 1.5 kb. Together, these data indicated that our venomous gland cDNA library was successfully constructed from the Chinese centipede *S. subspinipes* mutilans. Although many venomous gland cDNA libraries from animals (e.g., scorpion, spider, snake), have been constructed to reveal and elucidate the diversity of venom peptides, as far as we know, the venomous gland cDNA library of *S. subspinipes* mutilans is the first reported cDNA library from centipede venom glands. At present, extracting crude peptides and cloning cDNA sequences that encode venom peptides from venomous animals are two important and complementary methods for unraveling the puzzle of animal venom peptide components at the protein and transcript levels [1, 6, 7, 12, 17]. In this study, we successfully constructed a venom gland cDNA library from the centipede *S. subspinipes* mutilans, which lays the foundation for research into the structure and function of the venom peptides.

Precursor nucleotide sequence analysis of SsmTx

We used 500 random colonies from the centipede *S. subspinipes* mutilans venom gland cDNA library to screen cDNA sequences encoding venom peptides homologous to the SvnTx3 venom peptide from the Brazilian centipede *S. viridicornis* nigra by colony PCR. Six positive colonies were obtained. Sequence analysis and multiple alignments revealed that one of the six positive colonies encodes a new venom peptide with homology to the SvnTx3 venom peptide, named SsmTx (*S. subspinipes* **m**utilans **Tox**in). The full-length SsmTx cDNA sequence is 429 nucleotides (nt) in length, including a 246 nt open reading frame (ORF), 48 nt 5' untranslated region (5'UTR) and a 135 nt 3' untranslated region (3'UTR) (fig. 3). The A + T content of the 3' UTR (88.1%) is higher than those of the 5' UTR and ORF (66.7 and 67.9%, respectively). Additionally, the non-coding regions of the SsmTx cDNA sequence have the following structural features: (a) the flanking nucleotides of the start codon ATG are AAA (adenines), (b) the putative polyadenylation signal (AATAAA) is found 25 nucleotides upstream of the poly(A) tail, and (c) the 3' UTRs are rich in AAA or TTT motifs that may be involved in controlling mRNA stability [18]. As far as we know, SsmTx is the first reported cDNA sequence encoding a venom peptide from the venom of the centipede *S. subspinipes* mutilans.

```
ggggagaaaagaaaataatcggaaagacatattccaattgtaagaagaATGAAAATGGAGAAAAAAATT
                                                M  K  M  E  K  K  I
ATTTTCCTGTGTTTTTTTGTTTCTCTTTTGACTCTTCCTGAATTTATTTCGTCTCAAGTTCTCGTAAAA
 I  F  L  C  F  F  V  S  L  L  T  L  P  E  F  I  S  S  Q  V  L  V  K
GATGACGTTCCGTTTAAAGAAAAAAGATTTCCTGATAGAGGTGAATGTATTAAGGCTTGTGCAGCCAAA
 D  D  V  P  F  K  E  K  R  F  P  D  R  G  E  C  I  K  A  C  A  A  K
TTTACGGATGGAGAAGAAGGCATAATAAAAGATGTAGAACCGGTTTTTTATAAGTGTATTTGCTGGTAT
 F  T  D  G  E  E  G  I  I  K  D  V  E  P  V  F  Y  K  C  I  C  W  Y
TATGTGATATTACTCTACTAGttaaatcataaaaatctaaaatcaactgatgagaagatttaaaaaata
 Y  V  I  L  L  Y  @
attcaaatagtattatttatgtatttaaaatatatgttgtaattgaaattataaaaaataaaaattcaa
ttataatttaaaaat
```

Fig. 3. The precursor nucleotide sequence and the deduced amino acid sequence of SsmTx. 5' and 3' UTR sequences are indicated in lowercase letters. The predicted amino acid sequence is shown below the corresponding nucleotide sequence. Mature peptide residues are underlined. The cysteine residues of the mature peptide are highlighted in bold. The polyadenylation signal "aataaa" is double-underlined.

```
SvnTx3  EQ.LIKKDVKHKDK.FAKKSECVRAAAAFTGGDKS..   34
SsmTx   .QVLVKDDVPFKEKRFPDRGECIKACAAKFTDGEEGII   37

SvnTx3  ....................                  34
SsmTx   KDVEPVFYKCICWYYVILLY                  57
```

Fig. 4. Sequence alignment of SsmTx and the *Scolopendra* 5848.48 Da venom peptides (SvnTx3). Sequence alignment was carried out using the DNAMAN software. Identical residues are indicated by a black background.

Amino acid sequence analysis of SsmTx

The peptide deduced from the SsmTx cDNA sequence is 82 amino acid residues long. We attempted to find SsmTx homologs in the GenBank using the BlastP method (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The results showed that the full-length SsmTx only has some similarity with the Scolopendra 5848.48 Da venom peptide from the Brazilian centipede *S. viridicornis* nigra. The N-terminal amino acid sequence of SsmTx (QVLVKDDVPFKEKRFPDRGECIKACAAKFTDG) shares a high homology (54.8% identical amino acid residues) with the partialy sequenced N-terminal amino acid sequence of the Scolopendra 5848.48 Da venom peptide (fig. 4) [1]. Therefore, the localization of the cleavage site between the signal peptide and mature peptide in the SsmTx precursor was predicted using neural networks (NN) and hidden Markov models (HMM) methods with SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Both methods showed that the probability of the presence of a signal peptide in the SsmTx precursor is high with the most likely cleavage site to be positioned between the 25 and 26 **S-Q** (fig. 5). Moreover, multiple sequence alignment of SsmTx and the Scolopendra 5848.48 Da venom peptide further confirmed that the cleavage site between the signal peptide and mature peptide of the SsmTx precursor is located between positions 25 and 26 **S-Q.** All together, these results concluded that the peptide deduced from the SsmTx cDNA sequence is 82 amino acid residues long, including a signal peptide of 25 amino acid residues and a mature peptide of 57 amino acid residues (fig. 3).

Amino acid sequence analysis of the SsmTx precursor showed that the mature SsmTx toxin contains four cysteine residues, what possibly suggests the presence of two disulfide bridges. The cysteine pattern of SsmTx is **Cys-X-X-X-Cys...Cys-X-Cys** (X: any amino acid residue). Most of the scorpion venom derived toxins have a cysteine-stabilized αβ (CSα) motif (**Cys...Cys-X-X-X-Cys...Cys...Cys-X-Cys**) (X: any amino acid residue) [9, 19]. Conotoxins of the O-superfamily and some groups of spider neurotoxins possess an inhibitor cysteine knot (ICK) (**Cys...Cys...Cys-Cys...Cys...Cys**) [20, 21]. From the organization and structure of the cysteine pattern, the centipede venom peptide SsmTx possibly provides a new cysteine pat-

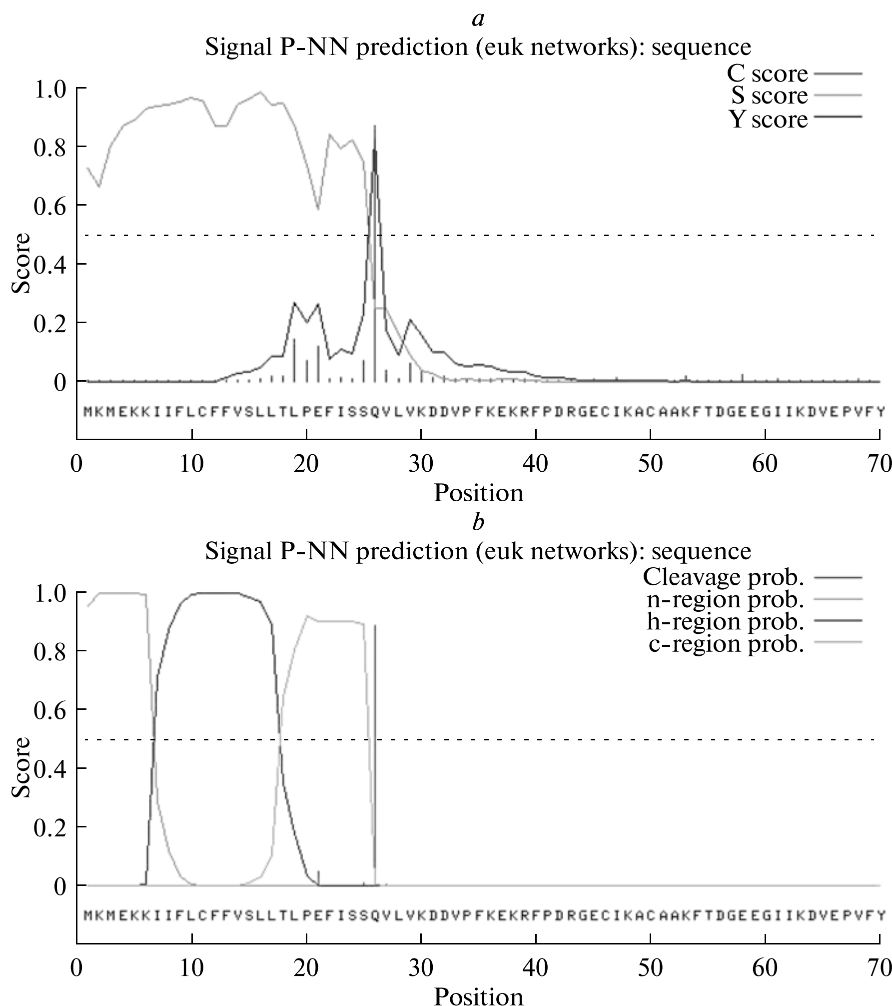

Fig. 5. The predicted cleavage site of the SsmTx signal peptide using the SignalP 3.0 Server. *a* – The cleavage site between the signal peptide and the mature peptide of SsmTx predicted using neural networks (NN). *b* – The cleavage site between the signal peptide and the mature peptide of SsmTx predicted using the hidden Markov models (HMM).

tern motif. It is of great interest to clarify the biological function of the centipede venom peptide SsmTx.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Natural Sciences Foundation of China (No. 30800571 and 30970145), the Fundamental Research Funds for the Central Universities (No. 3081002), and the Scientific Research Foundation for the Returned Overseas Scholars by the Ministry of Education of China.

REFERENCES

1. Rates B., Bemquerer M.P., Richardson M., Borges M.H., Morales R.A., de Lima M.E., Pimenta A.M. 2007. Venomic analyses of *Scolopendra viridicornis nigra* and *Scolopendra angulata* (Centipede, Scolopendromorpha): shedding light on venoms from a neglected group. *Toxicon*. **49**, 810–826.
2. Medeiros C.R., Susaki T.T., Knysak I., Cardoso J.L., Malaque C.M., Fan H.W., Santoro M.L., Franca F.O., Barbaro K.C. 2008. Epidemiologic and clinical survey of victims of centipede stings admitted to Hospital Vital Brazil (Sao Paulo, Brazil). *Toxicon*. **52**, 606–610.
3. Rouvin B., Seck M., N'Diaye M., Diatta B., Saissy J.M. 2008. Centipede bite in a woman with heterozygotous sickle cell disease. *Med. Trop. (Mars)*. **68**, 647–648.
4. Remington C.L. 1950. The bite and habits of a giant centipede (*Scolopendra subspinipes*) in the Philippine Islands. *Am. J. Trop. Med. Hyg*. **30**, 453–455.
5. Malta M.B., Lira M.S., Soares S.L., Rocha G.C., Knysak I., Martins R., Guizze S.P., Santoro M.L., Barbaro K.C. 2008. Toxic activities of Brazilian centipede venoms. *Toxicon*. **52**, 255–263.
6. Gonzalez-Morales L., Diego-Garcia E., Segovia L., Gutierrez Mdel C., Possani L.D. 2009. Venom from the centipede *Scolopendra viridis* Say: purification, gene

cloning and phylogenetic analysis of a phospholipase A2. *Toxicon*. **54**, 815.

7. Peng K., Kong Y., Zhai L., Wu X., Jia P., Liu J., Yu H. 2010. Two novel antimicrobial peptides from centipede venoms. *Toxicon*. **55**, 274–279.
8. Balit C.R., Harvey M.S., Waldock J.M., Isbister G.K. 2004. Prospective study of centipede bites in Australia. *J. Toxicol. Clin. Toxicol.* **42**, 41–48.
9. Ma Y., Zhao R., He Y., Li S., Liu J., Wu Y., Cao Z., Li W. 2009. Transcriptome analysis of the venom gland of the scorpion *Scorpiops jendeki*: implication for the evolution of the scorpion venom arsenal. *BMC Genomics*. **10**, e290.
10. Schwartz E.F., Diego-Garcia E., Rodriguez de la Vega R.C., Possani L.D. 2007. Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones). *BMC Genomics*. **8**, e119.
11. Jiang L., Peng L., Chen J., Zhang Y., Xiong X., Liang S. 2008. Molecular diversification based on analysis of expressed sequence tags from the venom glands of the Chinese bird spider *Ornithoctonus huwena*. *Toxicon*. **51**, 1479–1489.
12. Fernandes-Pedrosa Mde F., Junqueira-de-Azevedo Ide L., Goncalves-de-Andrade R.M., Kobashi L.S., Almeida D.D., Ho P.L., Tambourgi D.V. 2008. Transcriptome analysis of *Loxosceles laeta* (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. *BMC Genomics*. **9**, e279.
13. Melen G.J., Pesce C.G., Rossi M.S., Kornblihtt A.R. 1999. Novel processing in a mammalian nuclear 28S pre-rRNA: tissue-specific elimination of an 'intron' bearing a hidden break site. *EMBO. J.* **18**, 3107–3118.
14. Fujiwara H., Ishikawa H. 1986. Molecular mechanism of introduction of the hidden break into the 28S rRNA of insects: implication based on structural studies. *Nucl. Acids Res.* **14**, 6393–6401.
15. Winnebeck E.C., Millar C.D., Warman G.R. 2010. Why does insect RNA look degraded? *J. Insect Sci.* **10**,159.
16. Ma Y. 2009. Isolation, characterization and molecular evolution of scorpion toxin genes. *PhD thesis*, Wuhan, China: Wuhan University.
17. Wenhua R., Shuangquan Z., Daxiang S., Kaiya Z., Guang Y. 2006. Induction, purification and characterization of an antibacterial peptide scolopendrin I from the venom of centipede *Scolopendra subspinipes* mutilans. *Indian J. Biochem. Biophys.* **43**, 8893.
18. 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.
19. Rodriguez de la Vega R.C., Possani L.D. 2004. Current views on scorpion toxins specific for K^+-channels. *Toxicon*. **43**, 865–875.
20. Heinemann S.H., Leipold E. 2007. Conotoxins of the O-superfamily affecting voltage-gated sodium channels. *Cell. Mol. Life Sci.* **64**, 1329–1340.
21. Grishin E. 1999. Polypeptide neurotoxins from spider venoms. *Eur. J. Biochem*. **264**, 276–280.