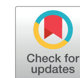
Toxicon 143 (2018) 1–19

Contents lists available at ScienceDirect

Toxicon

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

De novo transcriptomic analysis of the venomous glands from the scorpion *Heterometrus spinifer* revealed unique and extremely high diversity of the venom peptides



Yuchen Deng [a, 1], Jiawei Gu [b, 1], Zhenpeng Yan [a, 1], Mengdi Wang [c], Chaoqun Ma [d], Jiafeng Zhang [b], Guixian Jiang [b], Mengxiao Ge [b], Shuogui Xu [e], Zheng Xu [f, **], Liang Xiao [d, *]

[a] *Naval Clinical Medicine, Faculty of Naval Medicine, Second Military Medical University, Shanghai, 200433, China*
[b] *Clinical Medicine, Second Military Medical University, Shanghai, 200433, China*
[c] *Nursing School, Second Military Medical University, Shanghai, 200433, China*
[d] *Department of Marine Biotechnology, Faculty of Naval Medicine, Second Military Medical University, Shanghai, 200433, China*
[e] *Emergency Department of Changhai Hospital, Second Military Medical University, Shanghai, 200433, China*
[f] *Division of Scientific Research Administration, Second Military Medical University, Shanghai, 200433, China*

ARTICLE INFO

Article history:
Received 29 August 2017
Received in revised form
22 December 2017
Accepted 25 December 2017
Available online 2 January 2018

Keywords:
Scorpion
Heterometrus spinifer
Gland
Transcriptomics
Peptides
Toxins

ABSTRACT

Scorpion, as an ancient species, has been widely used on dozens of human diseases in traditional Chinese Medicine. Although the scorpion venom from the Buthidae family with the potent toxicity attracts more interests, toxins from the non-Buthidae family draw great attention as well because of its abundance and complexity even without harm to mammals. Moreover, several toxic components of scorpion venom have been identified as valuable scaffolds for the drug design and development. Using the Next Generation Sequencing (NGS) technique, here we reported the transcriptome of the venomous glands of *Heterometrus spinifer*, a non-Buthidae scorpion that only a few toxic and complete components have been identified known-to-date. The total mRNA extracted from the venomous glands of *H. spinifer* was subjected to illumina sequencing with a strategy of *de novo* assembly, and a total of 54 189 transcripts were unigenes from a total of 88 311 600 determined reads. We annotated 18 567 (34.26%) unigenes from NR database, 12 258 (22.62%) from SWISSPROT database, 11 161 (20.60%) from GO database, 10 159 (18.75%) from COG database and 5059 (9.34%) from KEGG database, respectively. 2843 unigenes were further selected against the toxin-related sub-database of SWISSPROT. After removing the redundancy, 13 common toxin-related subfamilies with 62 unigenes were manually confirmed, including 8 K-toxins, 1 calcin, 3 Imperatoxin I-like, 2 La1-like, 1 scorpin-like, 3 antimicrobial peptides, two types of protease inhibitors such as 8 Kunitz-type protease inhibitors and 3 Ascaris-type protease inhibitors, and 33 proteases including 16 serine proteinases, 7 phospholipases, 5 metalloproteases, 3 hyaluronidases and 2 phosphatases. Our report is the first transcriptomic analyses of venomous glands from the scorpion *H. spinifer*, serving as a public information platform for the development of novel bio-therapeutics.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Scorpion, an ancient living fossil without significant morphological changes, is one of the oldest animal species in the world (Waddington et al., 2015). The scorpions adapt to all kinds of climates and terrains (Ochoa et al., 2011; Prendini, 2005; Santibáñez-López et al., 2014; Santibanez-Lopez et al., 2015) whereas most of them distribute in the subtropical area and gradually decrease to the equator or the poles (Keirans and Pound, 2003). There are almost 20 families, 208 genera and 2231 species of scorpions on earth (Santibanez-Lopez et al., 2015). All scorpions are poisonous and 30–40 species are able to kill people with their venoms stored in the telson (Isbister and Bawaskar, 2014; Isbister et al., 2014). Scorpion venoms are a mixture of all components including salts, nucleotides, amino acids, enzymes, proteins and polypeptides.

* Corresponding author.
** Corresponding author.
E-mail addresses: steven.xz@163.com (Z. Xu), hormat830713@hotmail.com (L. Xiao).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.toxicon.2017.12.051
0041-0101/© 2018 Elsevier Ltd. All rights reserved.

Although proteins were reported as the main constituent of scorpion venom, the polypeptides that are in general 10–120 amino acid residues long are the most effective (Ortiz et al., 2015). They can be grouped into two major classes: disulfide-bridged peptides (DBPs) and non-disulfide-bridged peptides (NDBPs)(Zeng et al., 2005). The DBPs is usually composed of 28–120 amino acid residues with cross-linked by 3–5 disulfides (Santibáñez-López and Possani, 2015), mostly acting on the ion channels including Na^{+}, K^{+} and Ca^{2+} channels (Quintero-Hernández et al., 2013). Due to their high specificity and potent activity, a set of the DBPs like calcin have been proved invaluable in unraveling the structures and functions of various ion channels and served as templates for the development of novel therapeutics treating ion channels-associated diseases (Ortiz et al., 2015). The NDBPs have recently attracted great interest since the diversity of the NDBPs is even more than that of the DBPs.

As early as 1967, Rochat et al. have successfully purified and functionally characterized the neurotoxins from the venom of Buthidae scorpions using liquid chromatography (Rochat et al., 1967, 1970). However, the purification and characterization of individual components as well as their low venom yields have traditionally been a rate-limiting step in the process of realizing their biomedical potentials (Terrat et al., 2012). Until now, approximately 3000 different DBPs and 200 different NDBPs have been identified, which is estimated to be only ~2.0% of 200 000 peptides in total from the venomous glands of distinct scorpion species (Zeng et al., 2005). Therefore, molecular dissection of the scorpion toxins remains to be a high priority task.

Modern approaches for high-throughput screening for biomedically useful or promising molecules is significantly increasing the importance of characterizations of venomous components (Gonzalez-Morales et al., 2014; Rong et al., 2015; Terrat et al., 2012). Since from 1980s, the combined application of liquid chromatography and mass spectrometry (MS) greatly accelerated the identification of venomous constituents, thereby allowing the analyses of venomous proteomics. The proteomics by LC-MS/MS is able to identify ~5000 proteins in the venom that is greatly affected by the abundance of proteins. In comparison, transcriptomic analysis of venomous glands using the Next Generation Sequencing (NGS) like illumine sequencing platform (SOLEXA) is able to provide as many as ~50 000 transcripts that is far more than that of proteomics by LC-MS/MS. It provides full or part length sequences of transcripts that can be utilized as the invaluable database aligned by the short peptides with LC-MS/MS, therefore greatly improving the efficiency and decreasing the cost of the protein identification (Angeloni et al., 2011; Cahais et al., 2012; Grabherr et al., 2011; Li et al., 2012).

In this study, illumina NGS was performed for transcriptomic analysis of venomous glands of the scorpion *Heterometrus spinifer*, in which only a few components have been identified and most of them are unknown. The volume of transcriptomic data was about 10 GBs (Gigabytes). 54 189 transcripts were successfully generated. We annotated 18 567 (34.26%) transcripts from NR database, and 2843 transcripts were further selected against the toxin-related sub-database of SWISSPROT. After removing the redundancy, a total of 62 toxin-related transcripts were finally confirmed, including 8 K-toxins, 1 calcin, 3 Imperatoxin I-like, 2 La1-like, 1 scorpin-like, 3 antimicrobial, two types of protease inhibitors such as 8 Kunitz-type protease inhibitors and 3 Ascaris-type protease inhibitors, and 33 proteases including 16 serine proteinases, 7 phospholipases, 5 metalloproteases, 3 hyaluronidases and 2 phosphatases. To the best of our knowledge, this is the first report of transcriptomic analyses of venomous glands of *H. spinifer*. Results released can be served as a public information platform for the development of novel bio-therapeutics.

2. Material and methods

2.1. Sample collection and identification

The *H. spinifer* specimens were obtained from Xizang province, in southwest of China. The obtained organisms were kept in the foam box and fed with crickets.

2.2. RNA extraction and test

The total RNA was extracted from one “replete” venomous gland of the scorpion *H. spinifer* that was frozen flashly (immediately frozen in liquid nitrogen) with the RNA purification Kit from AMBION's company according to manufacturer's instructions. A NanoDrop 2000 Spectrophotometer was used to test the purity of the extracted RNA, whereas the concentration, RNA integrity number (RIN) and the ratio 28S/18S RNA were obtained with a 2100 Bioanalyzer (Agilent Technologies).

2.3. cDNA library construction and sequencing

The total RNA sample was digested using DNase to remove remaining DNA, then it was purified to enrich the poly-A containing mRNA using oligo-dT-attached magnetic beads. After that, the mRNA was fragmented randomly using divalent cation during high temperature (94 °C). Using the fragmented mRNA as the templates, the first cDNA strand was produced by reverse transcription using random hexamer primers. DNA polymerase I, dNTPs and RNase H were immediately added into the mix to degrade the RNA template followed by synthesizing the double-stranded (ds) cDNA. The ds cDNA were then purified using AMPure XP beads, to remove dNTP, random hexamer primers, salt ions and other impurities from the reaction mix. The end of ds cDNA were repaired to convert the overhang into blunt ends by phosphorylation using an End Repair (ERP) Mix. The 3′ terminals of ds cDNA were adenylated using the Klenowexo to avoid wrong ligation between strands during the adapter ligating reaction. Adapters were ligated by complementary between “A” overhang of ds cDNA and “T” overhang of adapters using a T4 DNA ligase. Then, the fragments were selected according to size using AMPure XP beads. Finally the ds cDNA with the adapters were enriched and amplified using bridge PCR with adapter-specific primers. The quality of the cDNA construction was confirmed using the Agilent 2100 Bioanalyzer (Agilent Technologies). After that, the purified and eligible cDNA library was sequenced using Illumina HiSeq™ 2500 following manufacturers' instructions.

2.4. Quality control and assembly

The raw sequencing intensities were transformed as the raw reads, which were saved with fastq format including sequence information and the Phred's quality scores. The FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to access the quality of the raw reads and the NGS QC Toolkit v2.3.3 software (http://59.163.192.90:8080/ngsqctoolkit/) was applied to remove the fragments of low quality (Quality threshold is 20). 500 000 reads were taken out randomly from the raw reads and the best aligned sequences (E value $< 10^{-10}$ and coverage > 80%) against the nucleotide sequence database (NT) (ftp://ftp.ncbi.nih.gov/blast/db) were tested if the sample was polluted. After that, *De novo* assembly was performed with the software Trinity to generate the transcripts that were further clustered to obtain unigenes by using the software TGICL. All assembled unigenes were annotated against NR database (ftp://ftp.ncbi.nih.gov/blast/db) using the BLASTx (search protein databases

Download English Version:

https://daneshyari.com/en/article/8394693

Download Persian Version:

https://daneshyari.com/article/8394693

Daneshyari.com