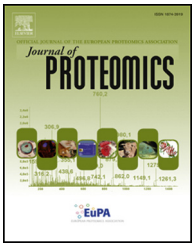
JOURNAL OF PROTEOMICS XX (2013) XXX–XXX

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot



Molecular diversity of Chaerilidae venom peptides reveals the dynamic evolution of scorpion venom components from Buthidae to non-Buthidae

Yawen He[1], *Ruiming Zhao*[1], *Zhiyong Di*, *Zhongjie Li*, *Xiaobo Xu*, *Wei Hong*, *Yingliang Wu*, *Huabin Zhao*, *Wenxin Li*, *Zhijian Cao*[*]

State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, PR China

ARTICLE INFO

Article history:
Received 21 February 2013
Accepted 9 June 2013

Keywords:
Scorpion venom arsenal
Transcriptome analysis
Molecular diversity
Molecular recruitment
Peptidome evolution

ABSTRACT

The scorpion family Chaerilidae is phylogenetically differentiated from Buthidae. Their venom components are not known, and the evolution of the venom components is not well understood. Here, we performed a transcriptome analysis of the venom glands from two scorpion species, *Chaerilus tricostatus* and *Chaerilus tryznai*. Fourteen types of venom peptides were discovered from two species, 10 of which were shared by both *C. tricostatus* and *C. tryznai*. Notably, the venom components of Chaerilidae were also found to contain four toxin types (NaTx, β-KTx, Scamp and bpp-like peptides), previously considered to be specific to Buthidae. Moreover, cytolytic peptides were the most abundant toxin type in *C. tricostatus*, *C. tryznai* and the family Euscorpiidae. Furthermore, 39 and 35 novel atypical venom molecules were identified from *C. tricostatus* and *C. tryznai*, respectively. Finally, the evolutionary analysis showed that the NaTx, β-KTx, and bpp-like toxin types were recruited into the venom before the lineage split between Buthidae and non-Buthidae families. This study provides an integrated understanding of the venom components of the scorpion family Chaerilidae. The family Chaerilidae has a specific venom arsenal that is intermediate between Buthidae and non-Buthidae, which suggests the dynamic evolution of scorpion venom components from Buthidae to non-Buthidae species.

Biological Significance

This work gave a first overview of the venom components of Chaerilidae scorpions, and discovered large numbers of new toxin molecules, which significantly enriches the molecular diversity of scorpion venom peptides/proteins components. Based on phylogenetic analysis we speculated that the NaTx, β-KTx and bpp-like toxin type genes were recruited into venom before the lineage split between Buthidae and non-Buthidae. By Comparing the toxin types and abundance of the Buthidae, Chaerilidae and non-Buthidae families, we found that the family Chaerilidae has a specific venom arsenal that is intermediate Buthidae and non-Buthidae, which suggests the dynamic evolution of scorpion venom components from Buthidae to non-Buthidae species.

[*] *Corresponding author*. Tel.: +86 27 68752831; fax: +86 27 68756746.
E-mail address: zjcao@whu.edu.cn (Z. Cao).
[1] These authors contribute to this work equally.

1874-3919/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jprot.2013.06.007

1. Introduction

Scorpions are among the most ancient animals on Earth. Scorpions use their venom glands for defense and immobilizing prey, which has been important for their successful survival for more than 400 million years [1]. Scorpion venom contains a large variety of biologically active components that could potentially be a rich source for drug design and development [2]. It has been assumed that approximately hundreds of thousands of distinct polypeptides are present in the approximately 2000 known scorpion species in the world [3]. However, only approximately 1% of these polypeptides are currently known, and the remaining part has an important potential for research and development.

Proteomic and transcriptomic approaches can be used to profile the venom toxin composition and discover that the general compositions of scorpion venoms vary among different families, genera, species and individuals [4–7]. Previously, most studies focused on the venoms of the family Buthidae due to their medical importance. However, the toxin types and abundances from non-Buthidae families are distinct from those of the Buthidae family. Recently, increased attention has been paid to the non-Buthidae families because several classes of venom peptides and proteins from non-Buthidae scorpions were shown to possess unique primary sequences and biological activities [8–10].

Transcriptome analysis of the venom gland is a powerful approach in obtaining large novel toxin molecules and can provide insight into the venom components of a scorpion species [11]. Comparative venom gland transcriptome analysis was used to reveal the intraspecific diversity of venom components [7]. Furthermore, some recent investigations comparing the venom gland transcriptome of different scorpion families and an evolutionary analysis of toxin sequences have provided insight into the evolution of the scorpion venom arsenal [12]. However, the evolutionary clues of some toxin types are not clearly understood because of the absence of transcriptomic data on different scorpion lineages, particularly non-Buthidae scorpions.

The extant scorpions can be phylogenetically divided into 13 families based on a cladistic morphological analysis [13,14]. The family Chaerilidae which distributes in the South and Southeast of Asia is phylogenetically and morphologically differentiated from the family Buthidae. The developmental processes of Chaerilidae scorpions have been reported to be more similar to those of Buthidae than those of non-Buthidae scorpions [15]. Whether the venom components of Chaerilidae scorpions tend to be similar to those of Buthidae or non-Buthidae scorpions is worthy of investigation. However, nothing is currently known regarding Chaerilidae venom components.

In this study, an EST approach was used to review the venom gland transcriptome of two Chaerilidae scorpions, *Chaerilus tricostatus* and *Chaerilus tryznai*. We identified a total of 14 known toxin types in the venom of these two scorpion species, 10 of which were shared by both species. Additionally, 39 and 35 atypical toxin molecules were discovered from *C. tricostatus* and *C. tryznai*, respectively. This study provides the first profile of the venom components from the family Chaerilidae. A comparison of the venom components from the Buthidae and non-Buthidae families showed that Chaerilidae has a specific venom arsenal that is intermediate between the Buthidae and non-Buthidae families. A phylogenetic analysis suggested that three toxins, NaTx, β-KTx and bpp-like peptides, were recruited into the venom before the lineage split between Buthidae and non-Buthidae (except for Pseudochactidae). Our study reveals the evolutionary track of the scorpion venom arsenal from the Buthidae to non-Buthidae families.

2. Materials and methods

2.1. cDNA library construction

Eight specimens of *C. tricostatus* were collected from Motuo County, Tibet, and 14 specimens of *C. tryznai* were collected from Bomi County, Tibet. The scorpion venom glands were extracted by electrical stimulation 2 days before RNA isolation [16]. Total RNA was extracted with TRIZOL Reagent (Invitrogen, Carlsbad, CA, USA), and then a FastTrack 2.0 mRNA Isolation Kit (Invitrogen) was used to purify the mRNA. The cDNA libraries were constructed from 0.1 μg of mRNA using a Creator SMART cDNA Library Construction Kit (Clontech Laboratories, Palo Alto, CA). The cDNA inserts were digested by restriction enzymes *Sfi*IA and *Sfi*IB and were then directionally cloned into the plasmids pDNR-LIB. The recombinant plasmids were transformed into electrocompetent *Escherichia coli* DH5α (Invitrogen).

2.2. Sequencing

The two cDNA libraries from the venom glands of the Chaerilidae scorpions *C. tricostatus* and *C. tryznai* were plated onto a Luria Broth (LB) culture medium containing 30 μg/ml of chloramphenicol and then cultured overnight. Random colonies were selected to obtain an unbiased overview of the venom gland transcriptome. The plasmid DNA was then isolated using the alkaline lysis method after the overnight culture. The purified plasmids were single-pass sequenced on ABI 3730 automated sequencers (Applied Biosystems, Foster City, CA, USA).

2.3. Bioinformatics analysis

The trace files of sequenced clones were subjected to the Phred program, in which the cutoff Phred score was established at 40 [17]. High-quality sequences were processed on the EGassembler website, http://egassembler.hgc.jp/, with the default parameters [18]. Adaptor and vector sequences were removed using Cross_Match software. We discarded sequences shorter than 100 bp after removing the PolyA tail. The obtained sequences were deposited into the dbEST and assembled into clusters using the program CAP3.

Each cluster was annotated by searching for it against the SWISS-PROT (http://www.expasy.org/tools/blast/) and GenBank NCBI (http://www.ncbi.nlm.nih.gov/blast) databases using BLAST algorithms with an e-value cut-off established at $<e^{-4}$. After the BLAST search, the unmatched clusters were further identified for open reading frames using the ORF Finder (http://www.ncbi.nlm.nih.gov/projects/gorf/). Considering the diversity of scorpion toxins, clusters putatively encoding venom peptides were re-examined manually to determine the individual different

Please cite this article as: He Y, et al, Molecular diversity of Chaerilidae venom peptides reveals the dynamic evolution of scorpion venom components from Buthidae to non-Buthidae, Journal of Proteomics (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.007