



Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (Hymenoptera: Formicidae): The first transcriptome analysis of an ant species

Wafa Bouzid^a, Christophe Klopp^b, Marion Verdenaud^c, Frédéric Ducancel^c, Angélique Vétillard^{a,*}

^a Venoms and Biological Activities Laboratory, EA 4357, PRES-Université de Toulouse, Jean-François Champollion University Center, Albi, France

^b Plateforme bioinformatique Genotoul, UR875 Biométrie et Intelligence Artificielle, INRA, 31326 Castanet-Tolosan, France

^c CEA, iBiTec-S, Department of Pharmacology and Immunoanalysis, Laboratory of Antibody Engineering for Health, Gif-sur-Yvette F-91191, France

ARTICLE INFO

Article history:

Received 6 December 2012

Received in revised form 5 March 2013

Accepted 11 March 2013

Available online 11 April 2013

Keywords:

Tetramorium bicarinatum

Ant

Venom glands

Allergens

ESTs library

Hymenoptera

ABSTRACT

Animal venoms are complex mixtures containing a range of bioactive elements with potential pharmacological and therapeutic use. Even though ants account among the most diverse zoological group, little information is available regarding their venom composition. To initiate the characterization of the transcriptomic venom gland expression of the ant species *Tetramorium bicarinatum*, 400 randomly selected clones from cDNA library were sequenced and a total of 364 high quality expressed sequence tags (ESTs) were generated. Based on the results of BLAST searches, these sequences were clustered and assembled into 83 contigs (22 multiple sequences) and 61 singletons. About 74% (267) of the contigs matched BLASTx hits with an interesting diversity together with an unusual abundance of cellular transcripts related to gene expression regulation (29% of the total library) reflecting the specialization of this tissue. About eighteen per cent of the ESTs were categorized as Hymenoptera venom compounds, the major part represented by allergens (62% of the total venom compounds). In addition, a high number of sequences (26%) had no similarity to any known sequences. This study provides a first insight of the gene expression scenario of the venom gland of *T. bicarinatum* which might contribute to acquiring a more comprehensive view on the origin and functional diversity of venom proteins among ants and more broadly among Hymenopteran insects.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

An extremely diverse range of animals have evolved their venoms for hunting, defense efficiency or competitor dissuasion. Some of the animal toxins have become an interesting source of pharmacological and therapeutic

molecules due to their high specificity and potency to act on a number of molecular targets (King, 2011).

Despite the obvious concern that these venom molecules represent, their investigation from animals remains under-exploited. Thus, in arthropods (Hoffman, 2010) more than 98% of their venoms components are completely unknown (Quintero-Hernández et al., 2011). Moreover, studies of venom composition have mainly focused on biochemical characterization of specific toxins rather than genomic or transcriptomic investigation. Nevertheless, transcriptomic approaches, have been revealed more effective in unraveling

* Corresponding author. Tel.: +33 563486440; fax: +33 563486432.

E-mail addresses: avetillard@yahoo.fr, angelique.vetillard@univ-jfc.fr (A. Vétillard).

the complexity of some animal venom components (Liang, 2004; Ruiming et al., 2010). Indeed, they can reflect the biological processes inside the venom gland cells, but also provide clues to the research of evolutionary path leading to toxin diversification. Because venom peptides are synthesized in specialized toxin-producing tissues, the cloning and sequencing of molecules from venom-cells cDNA libraries have become one of the methods of choice in the discovery of novel venom peptides: hundreds of venom peptides have thus been discovered. However, the sequencing of the whole toxin-encoding cDNAs has been progressively substituted by partial cDNA sequencing generating the so-called expressed sequence tags (ESTs), a rapid and reliable method that provides an informative overview of major transcripts in specific tissues (Adams et al., 1991). In venomous animals, ESTs are currently used for the investigation of the organization of the non-mature proteins and post-translational modifications and/or the discovery of several new peptides that have not been revealed by direct analyses from crude venom (Wagstaff and Harrison, 2006; Fry et al., 2010; Ruiming et al., 2010). More recently, cDNA deep sequencing technologies, such as pyrosequencing using 454 Life Sciences technology have provided a more comprehensive characterization of precursors encoding animal venom peptides and have overcome some difficulties regarding the availability in limited amounts of venoms from small or rare species (Stöcklin and Vorherr, 2011).

The studied species here: *Tetramorium bicarinatum*, is one of the world's most broadly distributed ant species. It was selected because of the biological diversity of its ecosystem and the potential of its venom toward human beings and vertebrates in general, suggesting the presence of new bioactive substances. However, obtaining reasonable quantities of ant venom severely hampers investigation of these components (Lima and Brochetto-Braga, 2003). Indeed, when ants produce about 10 µg of dry venom or less per individual, spiders, scorpions, and snakes produce 1–300 mg. Nevertheless, the venom ants of the genus *Solenopsis* and *Myrmecia* have been studied by proteomic approaches, revealing several allergenic and venom related compounds (Hoffman et al., 1990; Wiese et al., 2006; Dos Santos Pinto et al., 2012).

As far as we are aware, no studies have attempted to examine the transcriptomic pattern of venom glands from ants. In order to investigate a first insight on the expression profile in the venom gland of the ant species *T. bicarinatum* (Myrmicinae), we generated an enriched *T. bicarinatum* cDNA library by the use of the suppression subtractive hybridization (SSH) technique. This approach, with the enrichment with specific expressed transcripts, limits the redundancy with non-interesting information such as housekeeping genes (Baek et al., 2009). The results of the sequencing and annotation of 364 high quality ESTs are presented and discussed.

2. Material and methods

2.1. Ants and rearing conditions

Polygynous colonies of the species *T. bicarinatum* were collected from Brazil (Itabuna, Bahia). In the laboratory,

colonies were maintained under homogeneous conditions (a temperature of 25°–28 °C). They were reared in artificial nests consisting of a plastic box (covered with a black lid) connected to a foraging arena. Ants were fed on mealworm larvae and diluted honey twice a week.

Two hundred ant workers were anesthetized with CO₂ and kept on ice during dissection. The abdominal organs were gently pulled out with tweezers. The venom glands/reservoirs and stinging apparatus (Fig. 1) were retrieved from other tissues (stomach, alimentary canal, etc.). The venom reservoirs were discarded from the pair of glands as they have been shown to contain only myosin and muscle-specific proteins (Baek et al., 2009). Dissected glands were flash frozen in liquid nitrogen and stored at –80 °C until processed. In order to achieve the suppression subtractive hybridization, the remaining bodies of the dissected ant workers were pulled in a separate tube and immediately stored in liquid nitrogen, with prior retrieval of stomach and alimentary canal, in order to avoid contaminating the ESTs library with unnecessary possible micro-organism genes.

2.2. Total RNA extraction and construction of subtractive venom glands library

Total RNAs (tRNAs) from venom glands were isolated with RNeasy Micro Kit (Qiagen, France) including an on column DNase digestion whereas total RNAs from ant body carcasses were extracted using 400 µl of TRI reagent (Sigma) according to the manufacturer's protocol. Generation of cDNA was carried out using the SMARTer™ Pico PCR cDNA synthesis kit (Clontech, Mountain View, CA, USA) following the protocol instructed by the manufacturer. Synthesis of the first strand cDNA was performed with a modified oligo-dT primers (provided in the kit) according to the manufacturer's protocol from a starting material of 9 ng/µl of venoms glands and 20 ng/µl for ant carcass tRNA samples. The resulting first strand cDNAs from each RNA sample were then amplified with the Advantage2 PCR kit (Clontech) using 24 polymerase chain reaction (PCR) cycles. The resulting double-stranded cDNAs were used as template for library preparation. They were digested with *Rsa*I (Clontech) to obtain shorter and blunt-ended fragments. In order to achieve the suppressive subtractive hybridization, 300 ng/µl of cDNAs from venom glands, designated as “tester”, were divided into two aliquots that were subsequently ligated

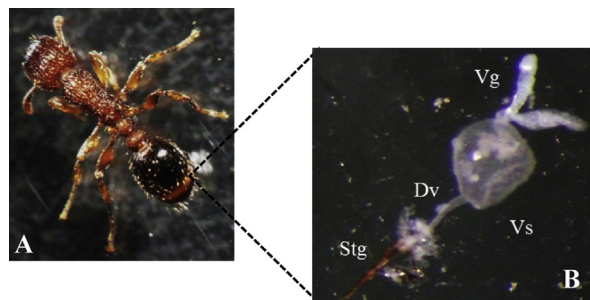


Fig. 1. A: *Tetramorium bicarinatum* worker reared in our laboratory seen under binocular microscope (X6). B: Venom sac (Vs), venom glands (Vg), stinging apparatus (Stg) and ductus venatus (Dv) dissected from worker of *T. bicarinatum*.

with Adaptor 1R or Adaptor 2R (provided in the kit). The obtained two different adaptor-ligated tester cDNAs pools were separately denatured at 98 °C for 90 s and then hybridized at 68 °C for 8 h with an excess of ant carcass cDNAs, called “driver”. These two primary hybridization samples were mixed together without denaturing, and then fresh denatured driver cDNAs from ant carcass was added for the second enrichment hybridizing at 68 °C for 16 h. After that, two suppressive PCRs were performed to selectively amplify the differential transcripts (the primers and the program parameters used for the two PCRs are supplied by the manufacturer). The products of the second PCR (nested PCR) were inserted into pGEM® T-Easy Vector (Promega, Madison, WI, United States) and cloned in High Efficiency JM109 Competent Cells (Promega).

2.3. Sequence data analysis

A total of 400 clones were picked out randomly and sequenced with ABI Prism BigDye Terminator v3.1 by cycle sequencing (Applied Biosystems, Foster City CA) with an SP6 Primer (Plateau de Génomique GeT-Purpan, Toulouse, France). The sequence processing chained three steps: sequences cleaning, clustering and annotation. In the first step, the base calling of the ESTs sequences was performed using Phred (version: 0.020425.c <http://www.phrap.com/phred/>). The produced reads have been cleaned of low complexity areas and polyA tails using seqclean (<http://compbio.dfci.harvard.edu/tgi/software/>) and then trimmed from vector sequences using crossmatch (www.phrap.org/) with parameter -minmatch 7 -minscore 15 and an in-house python script. Low quality zones were cleaned using another in-house python script based on a sliding windows strategy (window size: 20 bp, threshold quality: 20). In the second step, remaining high-quality ESTs were assembled into contigs using CAP3 (<http://genome.cshlp.org/content/9/9/868.full>) with -p 75 -s 500 parameters. In the third step, contigs (cluster of two or more sequences) and singletons (unique sequence) were annotated. To assign annotation to the assembled ESTs, they were searched against the non-redundant (nr) protein database for homologous comparison using BLASTx algorithm. In order to investigate ESTs that have not matched any annotation in GenBank, ORFs were predicted using framedp program (Gouzy et al., 2009) with the Insecta gene model extracted from UniProtKB (Magrane and Consortium, 2011) as protein reference database and getorf script from the EMBOSS software suite (Rice et al., 2000) with a minimal ORF length set at 200 nucleotides. The signal peptide was predicted with the signalP 4.0 program (<http://www.cbs.dtu.dk/services/SignalP/>) (Petersen et al., 2011). Cysteine patterns were identified from an in-house perl script, on the predicted mature peptide if a signal sequence was found or on the whole sequence otherwise.

3. Results

3.1. Analysis of cDNA quality and sequence clustering

A cDNA library specific to *T. bicarinatum* venom glands was generated using the SSH technique. A total of 400 ESTs

independent clones were obtained, among which 364 sequences were of high quality and were processed for annotation (Table 1). After vector, primers and adaptors sequences trimming, the ESTs encompassed a range of 264–2220 bp long (Table 2). The retained ESTs formed 83 clusters comprising 22 contigs and 61 singletons. The average length of sequences forming each contig is informed in Table 2. The number of ESTs in each contig ranged from 2 to 55. About 55% (12 contigs) of the 22 multiple sequences contained 2 to 9 ESTs and 4 multiple sequences contained 55, 39, 33 and 30 ESTs.

ESTs successfully identified using BLASTx, were gathered into protein families involved in cellular processes (38%), Hymenoptera venom compounds (18%) and toxin-like venom candidates (5%). In addition, about 13% of the ESTs matched hypothetical proteins with unknown function and 26% had no blast hits (Fig. 2). EST sequences were submitted to the ESTs database of GenBank (<http://www.ncbi.nlm.nih.gov/projects/dbEST/>) under accession numbers JZ168417–JZ168724.

3.2. Clusters related to cellular functions

Clusters encoding proteins with common cellular functions showed high representation of ESTs matching the Histone-lysine N-methyltransferases (about 74%) among the housekeeping encoding ESTs, followed by transcripts matching transposases (16%) (Fig. 3). The remaining 10% common transcripts were represented by two or one singlet and contained proteins related to general metabolism (nucleoside diphosphate sugar epimerase), cellular regulation (Zinc finger BED protein), DNA damage repair (Breast cancer suppressor protein), feeding physiological process (enterin neuropeptide) and protection (cuticular protein). Details on the length and the matched organisms are provided in Table 2. Noteworthy, the entries contig21, contig2, contig20, contig11, S154, S389 and S182 (Table 2) matched separately the Histone-lysine N-methyltransferase and were gathered under the largest (74%) transcript group (Fig. 3). However no evidence whether they covered different regions that possibly belong to the same gene or represent different isoforms are provided.

3.3. Clusters related to Hymenoptera venom compounds

Possible toxins and venom related compounds were found in this group (Fig. 4) where allergen peptides seem to have a major place (Fig. 5).

Table 1
Summary of ESTs from *T. bicarinatum* venom gland cDNA library.

Total number of input sequences	400
Trimmed sequences	364
(Average length, bp)	(886)
Contigs	83
Multiple sequences	22
2–9 ESTs	12
10–19 ESTs	4
20–29 ESTs	2
≥30 ESTs	4
Singletons	61
BLAST- hits	43
No BLAST- hits	40

Table 2

Annotation of the identified ESTs in *T. bicarinatum* venom glands. Numbers between brackets indicate the average length of the sequences forming each contig in bp.

Contig/singlet number	No of ESTs	Length (bp)	Best matched protein	Matched species and maximum identity	Genbank	E-value	Accession number
<i>Housekeeping transcripts</i>							
Contig 21 (553)	55	1131	Histone-lysine N-methyltransferase	<i>Acromyrmex echinator</i> [ant] (42%)	EG161648.1	2e-45	JZ168597–JZ168650
Contig 15 (741)	33	794	Histone-lysine N-methyltransferase	<i>Camponotus floridanus</i> [ant] (47%)	EFN70690.1	5e-29	JZ168558–JZ168588
Contig 2 (459)	23	1099	Transposase	<i>Agrilus planipennis</i> [beetle] (39%)	ADB28039.1	7e-16	JZ168440–JZ168459
Contig 20 (775)	7	795	Histone-lysine N-methyltransferase	<i>Camponotus floridanus</i> [ant] (59%)	EFN70690.1	5e-29	JZ168595–JZ168596
Contig 11 (684)	6	1121	Histone-lysine N-methyltransferase	<i>Camponotus floridanus</i> [ant] (74%)	EFN65776.1	2e-37	JZ168541
Contig 5 (204)	3	323	Breast Cancer Suppressor Protein	<i>Camponotus floridanus</i> [ant] (31%)	EFN68364.1	6e-04	JZ168490–JZ168491
Contig 10 (331)	3	382	Enterin neuropeptide	<i>Acromyrmex echinator</i> [ant] (40%)	EG168727.1	0.009	JZ168538–JZ168540
Contig 7 (354)	2	389	Zinc finger BED domain-containing protein 1-like	<i>Acyrtosiphon pisum</i> [aphid] (30%)	XP_003240281.1	5e-04	JZ168507–JZ168508
Vg 389	1	347	Histone-lysine N-methyltransferase	<i>Camponotus floridanus</i> [ant] (46%)	EFN65776.1	4e-21	JZ168572
Vg 105	1	778	Predicted nucleoside-diphosphate sugar epimerases	<i>Anopheles gambiae</i> [mosquito] (36%)	XP_001688591.1	3e-04	JZ168700
Vg 182	1	552	Histone-lysine N-methyltransferase	<i>Camponotus floridanus</i> [ant] (69%)	EFN65776.1	4e-25	JZ168708
Vg 254	1	1134	Cuticular protein	<i>Papilio polytes</i> [butterfly] (47%)	BAM18958.1	6e-04	JZ168706
Vg 361	1	1789	Cytochrome <i>b</i>	<i>Camponotus crozieri</i> [ant] (32%)	ABI97056.1	3e-04	JZ168702
Vg 285	1	445	GF14688	<i>Drosophila ananassae</i> [Drosophila] (39%)	XP_001961936.1	2e-04	JZ168689
Vg 121	1	517	Odorant-binding protein precursor	<i>Solenopsis invicta</i> [ant]	ADX94412.1	8e-47	JZ168690
Vg 393	1	1313	Protein involved in post-transcriptional gene expression processes	<i>Drosophila yakuba</i> [Drosophila] (37%)	XP_002098871.1	9e-04	JZ168701
<i>Transcripts specific to Hymenoptera venoms</i>							
Contig 9 (481)	18	520	Pilosulin 5-like	<i>Myrmecia banksi</i> [ant] (61%)	A9CM07.1	8e-06	JZ168521–JZ168537
Contig 6 (389)	17	800	Insect cytokine precursor uENF2	<i>Neogurelca himachala sangaica</i> [moth] (11%)	BAJ21218.1	3e-06	JZ168492–JZ168506
Contig 12 (456)	11	543	Pilosulin 5-like	<i>Myrmecia banksi</i> [ant] (56%)	A9CM07.1	3e-06	JZ168542–JZ168552
Contig 14 (739)	4	1067	Venom allergen 3 Sol i 3	<i>Solenopsis invicta</i> [ant] (58%)	P35778.2	5e-96	JZ168554–JZ168557
Contig 4 (182)	3	275	Secapin-like	<i>Apis florea</i> [honeybee] (31%)	XP_003689823.1	3e-07	JZ168487–JZ168489
Contig 16 (396)	3	402	pilosulin 3a-like	<i>Myrmecia pilosula</i> [ant] (58%)	2206305A	4e-06	JZ168589–JZ168591
Contig 13 (429)	2	441	pilosulin 3a-like	<i>Myrmecia pilosula</i> [ant]	Q26464.1	1e-04	JZ168553
Contig 17 (348)	2	399	Secapin preproprotein	<i>Apis mellifera</i> [honeybee] (64%)	NP_001011618.1	8e-05	JZ168592–JZ168593
Vg 181	1	456	pilosulin1-like	<i>Myrmecia pilosula</i> [ant]	AAB50883.1	4e-04	JZ168710
Vg 251	1	277	Secapin	<i>Polistes hebraeus</i> [wasp] (59%)	Q7YWA6.1	3e-05	JZ168705
Vg 252	1	264	Venom allergen 3 Sol i 3	<i>Solenopsis invicta</i> [ant] (45%)	AAB36116.1	3e-17	JZ168704
Vg 241	1	595	Venom allergen 3 Sol i 3	<i>Solenopsis invicta</i> [ant] (49%)	P35778.2	3e-42	JZ168698
Vg 19	1	421	pilosulin 3a-like	<i>Myrmecia pilosula</i> [ant] (60%)	2206305A	7e-04	JZ168713
Vg 381	1	2084	Predicted A disintegrin and metalloproteinase	<i>Acyrtosiphon pisum</i> [aphid] (42%)	XP_001947361.2	0.011	JZ168725
<i>Venom toxin candidates</i>							
Contig 8 (371)	14	743	Predicted ATP-dependent zinc metalloprotease	<i>Bombus impatiens</i> [bumblebee] (48%)	XP_003485086.1	0.005	JZ168509–JZ168520
Vg 180	1	681	Mitogen-activated protein kinase	<i>Camponotus floridanus</i> [ant] (42%)	EFN62860.1	0.022	JZ168709

(continued on next page)

Table 2 (continued)

Contig/singlet number	No of ESTs	Length (bp)	Best matched protein	Matched species and maximum identity	Genbank	E-value	Accession number
Vg 233	1	1198	Sialidase	<i>Acromyrmex echinator</i> [ant] (47%)	EGI66499.1	0.001	JZ168683
Vg 385	1	2080	Prophenoloxidase	<i>Neobellieria bullata</i> [flesh fly] (42%)	AAD45527.1	0.016	JZ168694
Vg 355	1	2556	Trypsin-like serine protease	<i>Drosophila willistoni</i> [Drosophila]	XP_002068835.1	0.02	JZ168693
Vg 352	1	372	Prolyl 4-hydroxylase alpha-related protein	<i>Drosophila melanogaster</i> [Drosophila] (45%)	AAM18063.1	0.002	JZ168720
<i>Proteins with unknown functions</i>							
Contig 22 (338)	39	390	Hypothetical protein	<i>Harpegnathos saltator</i> [ant] (46%)	EFN76725.1	7e-06	JZ168651–JZ168687
Contig 19 (293)	2	328	Hypothetical protein	<i>Danaus plexippus</i> [butterfly] (50%)	EHJ78437.1	9e-04	JZ168594
Vg 154	1	689	Putative uncharacterized protein	<i>Camponotus floridanus</i> [ant] (60%)	EFN74761.1	1e-36	JZ168714
Vg 345	1	624	Hypothetical protein	<i>Aedes aegypti</i> [mosquito] (30%)	XP_001660252.1	7e-04	JZ168724
Vg 346	1	1115	Putative uncharacterized protein	<i>Camponotus floridanus</i> [ant] (59%)	EFN74761.1	4e-34	JZ168688
Vg 7	1	253	Hypothetical protein	<i>Culex quinquefasciatus</i> [mosquito] (32%)	XP_001851371.1	8e-04	JZ168696
Vg 162	1	372	Hypothetical protein	<i>Culex quinquefasciatus</i> [mosquito] (25%)	XP_001868571.1	4e-03	JZ168712
<i>No hits</i>							
Contig 3 (469)	30	476	–	–	–	–	JZ168460–JZ168486
Contig 1 (273)	25	521	–	–	–	–	JZ168417–JZ168439
Contig18 (971)	2	971	–	–	–	–	JZ168691–JZ168692

3.3.1. Allergenic/defense peptides

Since venoms of Hymenoptera are common causes of allergic and serious anaphylactic reactions, we have further investigated the amino-acid sequence of all ESTs sequences of *T. bicarinatum* identified as allergens (Table 3) and compare it to known hymenopteran allergen peptides retrieved from GenBank. Multiple alignments were done with ClustalW2 at the EMBL-EBI website (<http://www.ebi.ac.uk/Tools/clustalw2/>).

One transcript (S181) results in a 456 bp sequence matching strongly with allergenic polypeptide pilosulin 1.

Contigs16 and 13 (3 and 2 ESTs, respectively) and one singleton (S19) matched pilosulin 3a with 402, 441 and 421 bp length respectively. Two contigs (contig 9 and 12, 35 ESTs) of 520 and 543 bp respectively were identified as pilosulin 5-like peptides (Tables 2 and 3). Contig 14 (7 ESTs, 1067 bp) and the singletons S241 (595 bp) and S252 (264 bp) were found similar to the venom allergen 3 Soil i 3 peptide. Deduced amino-acid sequences from *T. bicarinatum* identified as pilosulin-like peptides ranged from 46 to 79 amino acid residue lengths (Fig. 6). Identified amino acids from *T. bicarinatum* soil i3 were much longer with 237 residues (Fig. 7).

3.3.2. Paralytic peptides

Transcripts encoding for paralytic-like peptides generated in the present ESTs library were exclusively constituted of insect cytokine precursor uENF2 (Table 2). Transcripts were clustered in a unique contig (contig 6) of 800 bp represented by a relative high number of transcripts (20 ESTs). These sequences matched the paralytic peptide of the moth species *Neogurelca himachala sangaica* with a confident E-value of 3,00 E-06.

3.3.3. Other peptides

The remaining Hymenoptera venom compounds were represented by transcripts matching a group of secapins: secapin-like (3ESTs), secapin preprotein (2 ESTs) and secapin (1 ESTs). One transcript (S381) of 2084 bp was identified as disintegrin and metalloproteinase (Table 2).

3.4. *T. bicarinatum* venom toxin-like candidates

One contig (14 ESTs) of 743 bp was identified as homologous to the sequence of an ATP-dependant zinc

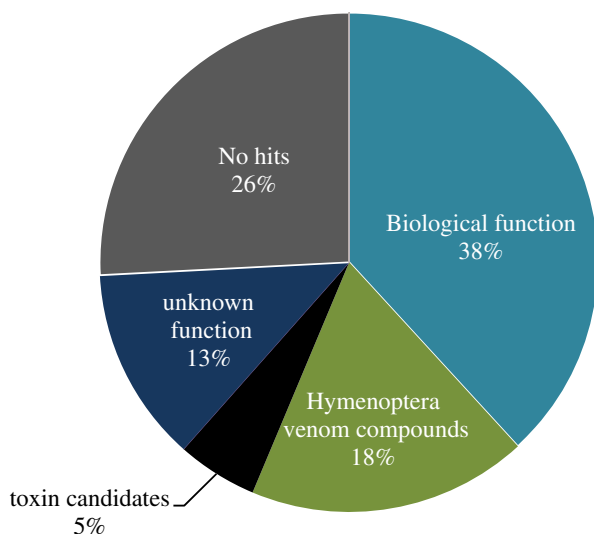


Fig. 2. General overview of the transcripts identified in *T. bicarinatum* venom glands.

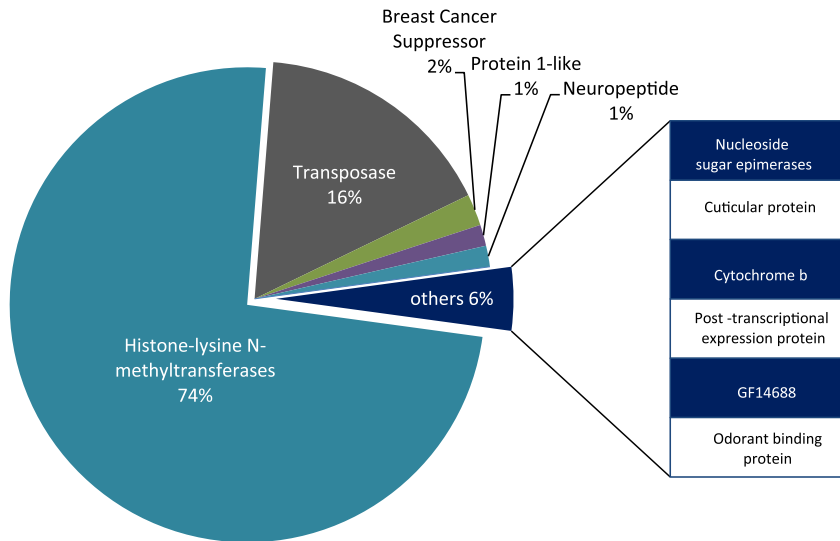


Fig. 3. Potential biological function of transcripts expressed in *T. bicarinatum* venom glands.

metalloprotease (Table 3). Sequences of this contig shared 30% of maximum identity with metalloproteases of the bumblebee species *Bombus impatiens*. A set of other transcripts were represented by single ESTs showing similarities with different enzymes namely mitogen activated protein kinase, sialidase, prophenoloxidase, trypsin-like serine protease and Prolyl 4-hydroxylase alpha-related proteins (Table 2). The corresponding singletons generated long transcripts of 681, 1198, 2080, 2556 and 372 bp respectively (Table 2).

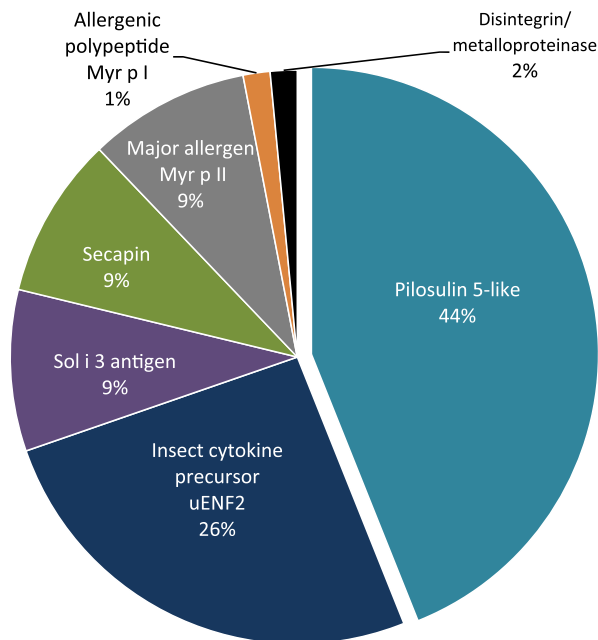


Fig. 4. Proportions and details of potential ESTs toxins and possible toxins expressed in *T. bicarinatum* venom glands.

3.5. No hits

Three contigs (n° 3, 1 and 18) of 30, 25 and 2 ESTs respectively, have not found homologous sequences in all the requested databases (Table 2). In addition 37 singletons have also failed to match any similar sequence (these are not shown in Table 2).

4. Discussion

Investigations on Hymenoptera venoms reported that peptide components comprise up to 70% of their freeze-dried weight (Palma and Brochetto-Braga, 1993). Most of these peptides account for cell lysis, hemolysis, antibiosis and in a few cases, some neurotoxic peptides (Palma, 2006). Even though venom composition of Hymenoptera has been relatively well characterized, especially that of bee and wasp, knowledge on the composition and properties of ant venoms is still limited, as there is great difficulty in obtaining sufficient venom samples (Lima and Brochetto-Braga, 2003). This fact is disappointing given that ants contain some of the most specialized species among the social insects, as well as species that may be considered as the least specialized ones (Palma, 2006). The ESTs cDNA

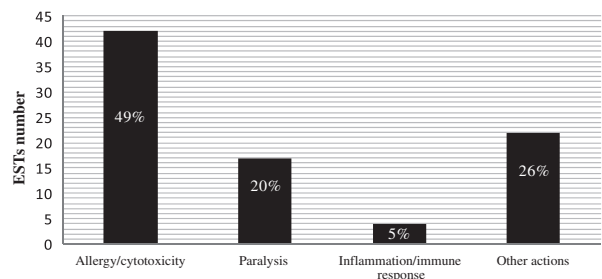


Fig. 5. General representation of Hymenoptera venom compounds and toxin-like candidates transcribed in *T. bicarinatum* venom gland cells.

Table 3

List of Hymenoptera compounds and toxin candidates identified in the present investigation of *T. bicarinatum* venom glands and their possible mechanism of actions.

Peptide categories	ESTs number
Allergy and cytotoxicity	
Pilosulin 5-like	29
Sol i 3	6
Major allergen Myr p II (pilosulin 3-like ^a)	6
Allergenic polypeptide Myr p I (pilosulin 1-like ^a)	1
Paralysis	
Insect cytokine precursor uENF2	17
Inflammation and immune response	
Disintegrin and metalloproteinase	1
Prophenoloxidase	1
Trypsin-like serine protease	1
Sialidase	1
Other actions	
Predicted metalloprotease	14
Secapin	6
Fibrinolytic enzyme 5	1
Mitogen-activated protein kinase	1

^a Myr p and their synonym pilosulins were based on Wiese et al. nomenclature (Wiese et al., 2007).

library generated in the present work shows the existence of weaponry of cell related function proteins, allergenic and paralytic peptides and different possible venom specific enzymes (Fig. 5). Given that we generated a normalized cDNA library which means that concentrations of high and low abundance cDNAs transcripts become roughly equal, we cannot state their actual respective level of transcription. Hence the first catalog of transcripts presented here, is rather a preliminary qualitative than a quantitative description of the transcriptome activity of the ant species *T. bicarinatum* venom gland cells.

According to the blast results, ESTs of *T. bicarinatum* venom glands have significantly matched cDNA from ants rather than from other Hymenoptera such as bees or wasps even though their venom composition has been well identified. This fits with the fact that the defensive weapon of Hymenoptera is different. Whereas bee and wasp venom are composed of protein and peptides, that of the well-studied ant species *Solenopsis invicta* consists of alkaloids and few soluble proteins (Santos et al., 2011). cDNA

sequences from *T. bicarinatum* obtained here share an overall average of 44% identity with available DNA sequences from related ant species (carpenter, jumper, leaf-cutter and fire ants) belonging to the same family of Formicidae and most of the time to the same sub-family of Myrmicinae (Table 2). DNA sequence divergence between *T. bicarinatum* and these ant species could easily be attributed to species distance lineage and to ecological behavior differences.

4.1. Hymenoptera compounds present in *T. bicarinatum* ESTs library

4.1.1. Allergenic/defense peptides

Allergenic peptides represent 49% of the *T. bicarinatum* venom compounds described in this work. This result fits with the fact that social insects of the order Hymenoptera cause mainly inflammatory and/or immunological reactions in their predator or prey rather than lethal effects. Indeed, investigations on the composition of Hymenoptera venoms showed a complex mixture of proteins with generally low molecular mass compounds and where proteins/peptides are recognized as important allergens (Hoffman, 2006). Several major allergens have been characterized from venom bees, vespids and ants by analytical techniques. They classically contain phospholipases, antigen 5, hyaluronidases and serine proteases with a molecular weight ranging from 10 to 50 KDa (Santos et al., 2011).

4.1.1.1. Pilosulin-like sequences. In our study, different pilosulin-like peptides were identified (Table 2) where pilosulin 5-like is likely to be the major expressed defense peptide in *T. bicarinatum* (Table 3). Pilosulins are allergenic peptides with immunoglobulin-binding activities. They derived from the cDNA Myr p 1, the major expressed allergenic product being a 56-residue peptide that have been described in venoms of *Myrmecia* spp (Wu et al., 1998; Inagaki et al., 2004, 2008). Previous amino-acid alignments of pilosulin isoforms showed that pilosulins peptides share high sequence similarities in the signal peptide and the following propeptides but have substantial variations in their mature peptide domain (Inagaki et al., 2008; Gruber

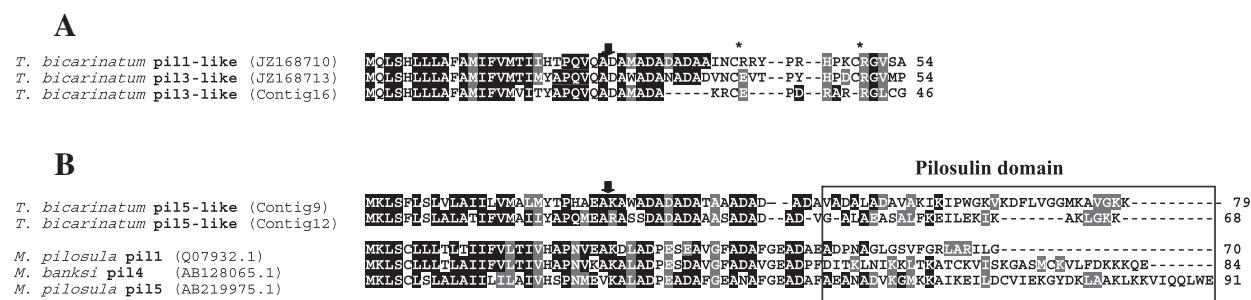


Fig. 6. Multiple alignments of the deduced amino acid sequences of the different identified pilosulin-like peptides pil1 and 3 (A) and pilosulin 5-like (B) from *T. bicarinatum* and other pilosulins from species of the genus *Myrmecia*. Gaps (–) have been introduced to improve sequence homology. The alignments were prepared with ClustalW2 and Boxshade. Identical and conserved amino acids are highlighted in black and gray respectively. The predicted mature peptides are presented in the box according to Gruber and Muttenthaler (2012) and the signal peptide cleavage site is indicated by an arrow. Cysteine residues are marked by asterisks. Pilosulin-like sequences from *T. bicarinatum* were assigned to Genbank under the accession numbers JZ168710, JZ168713, JZ168589–JZ168591 (for contig 16), JZ168521–JZ168537 (for contig 9) and JZ168542–JZ168552 (for contig 12).

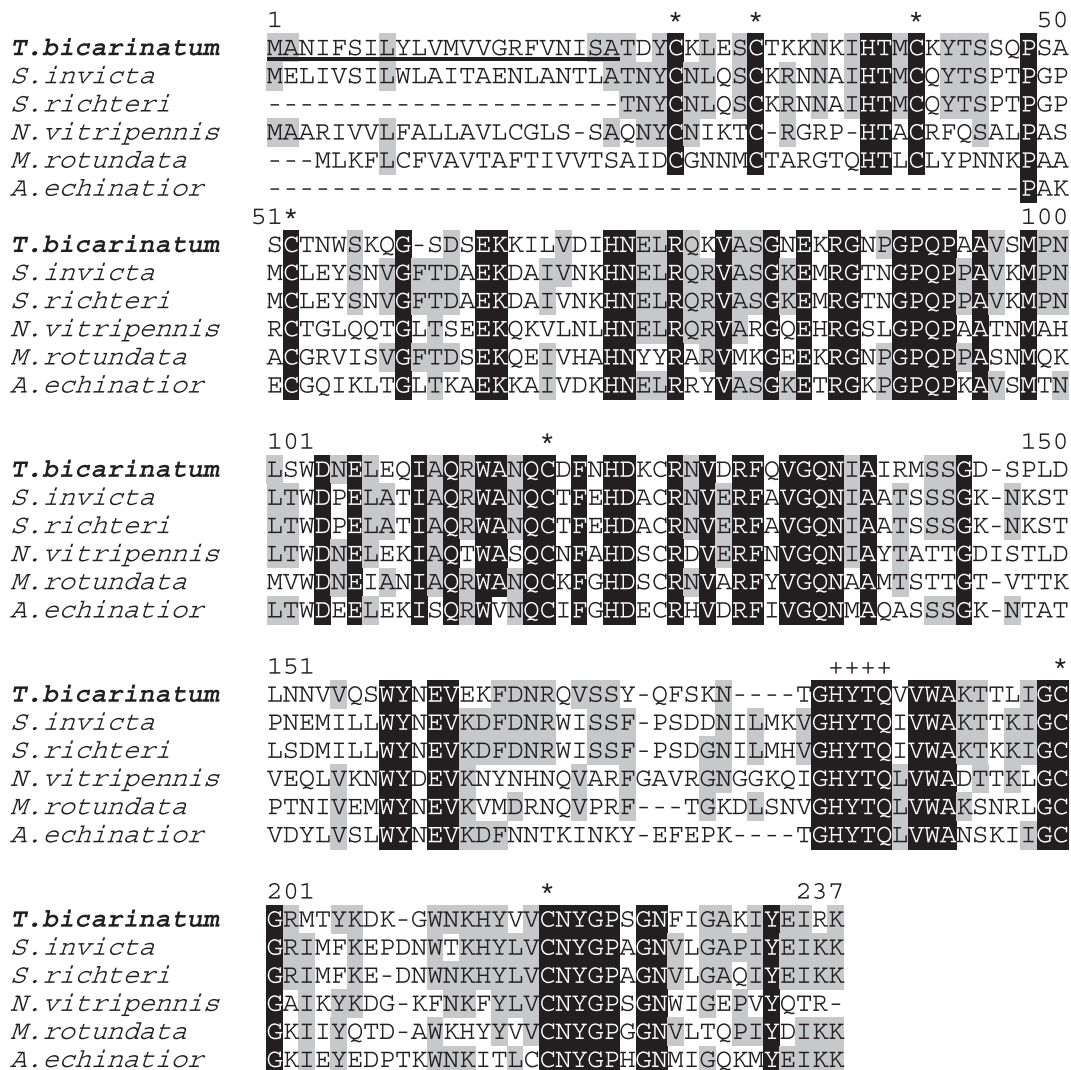


Fig. 7. Multiple alignments of the deduced amino acid sequences of venom allergen III from *T. bicarinatum*, with known sequences from different species. Identical and conserved amino acids are highlighted in black and gray respectively. Putative signal peptide is underlined and the conserved cysteine residues are indicated by asterisks. The conserved motif HYTQ is indicated by crosses. Residues are numbered according to (Fernandes-Pedrosa et al. (2008). The GenBank accession number for the aligned venom allergen III sequences are: *Solenopsis invicta* (P35778.2), *Solenopsis richteri* (P35779.2), *Nasonia vitripennis* (XP_001603551.2), *Megachile rotundata* (XP_003708569.1) and *Acromyrmex echinator* (EGL65623.1). Venom allergen III from *T. bicarinatum* (contig 14) was assigned to Genbank under the accession numbers JZ168554–JZ168557.

and Muttenthaler, 2012). Based on the alignments shown in Fig. 6, we find the same pattern *i.e.* high identities in the first 47 amino-acids (including the leader sequence) and high divergence in the predicted mature peptides. It was stated that members of certain family of bioactive peptides tend to have conserved leader sequences, as long as the bioactive peptides belong to the same family (Inagaki et al., 2008). For example, precursors belonging to bioactive peptides of the family of preprodermaseptin (antimicrobial peptides from the dermal glands of frogs) have an N-terminal preprosequence that is remarkably well conserved within and between species, while the C-terminal sequence is a hypermutable antimicrobial domain (Vanhoey et al., 2003). It was recently suggested that variation of active domain would seemingly be attributed to recombination as a

diversity-generating mechanism giving individualistic utilization of specific toxin isoforms and hence individualistic pattern of expression (Gilly et al., 2011).

In our library, it is likely that the identified pilosulin-like peptides belong to different families of peptides. Indeed, the pilosulin 5-like group (contigs 9 and 12) share more consistency with the compared *Myrmecia* spp. pilosulins in the first 47 amino acids including the leader sequence. For instance, patterns 'MKLS', 'ADA' and 'AD' (positions 1, 38 and 43, respectively) in the sequence leader and the pro-sequence were conserved. Even though, the mature peptide isoforms appear very divergent, it shows however more similarities in amino-acid composition than the other identified pilosulin like from *T. bicarinatum*. For instance, two alanine and one lysine residues (positions 47, 51 and 58

respectively) were conserved (Fig. 6B). Regarding the other identified group of pilosulin-like (bearing an 'MQLS' motif in the beginning of their signal sequences), an obvious substantial divergence was notified especially in the pro-peptide and mature domain (Fig. 6A). The formers were thus difficult to align given their short size (54 amino acids versus an average of 78 amino acids of the other pilosulin sequences). In addition, this group has unique cysteine locations (positions 41 and 50, Fig. 6A) which probably mean a different disulfide bonding. Given all together, we presume that we have a pilosulin-like group of peptides that probably belongs to a different family of biologically active *T. bicarinatum* venom compounds.

We are cautious that this assumption needs to be confirmed with further investigations from the direct peptide by Mass spectrometry and proteomic analyses. In addition, biological activity characterizations are needed to draw a conclusion regarding their status and explain their presence and role in *T. bicarinatum* with respect to its ecological features. Previous investigations of the allergenic effect of a synthetic pilosulin 5 peptide showed a significant histamine release in a dose dependent-manner but (unlike the other pilosulins) a weak antimicrobial effect against a panel of microbial species (Inagaki et al., 2008) whilst pilosulin 1 has been reported to have significant cytotoxic effects (King et al., 1998; Wu et al., 1998).

4.1.1.2. Soli i 3 allergenic-like peptides. Transcripts encoding Soli i 3 allergenic-like peptides were also identified in our library (9 ESTs). Soli i 3 is the major allergen of a series of Soli i peptides (Soli i 1–4) identified as the most frequent cause of hypersensitivity reactions to the fire ant *S. invicta* (Padavattan et al., 2008). The sol i 3, as described in wasp and some sandflies venom belongs to the cysteine-rich secretory proteins family. It was suggested that their presence in insects may bind cytokines with potential effects on the host immune response (Anderson et al., 2006).

Multiple alignments of amino acid sequence of the contig 14 with known sequences of venom allergen III from other species showed substantial similarities (Fig. 7). For example, all the cysteine residues were conserved and several motifs like GPQP, VGQN and CNYGP (positions 90, 136 and 218 respectively) were conserved. In addition, the motif HYTQ (position 186), shared by the majority of the allergen-like proteins, was also conserved. We also notified some amino-acid differences probably attributed to lineage speciation. The leader sequence seems to be very different especially with the species *Megachile rotundata* (the leaf-cutter bee) and the parasitoid wasp *Nasonia vitripennis* likely because of the phylogenetic distance. Indeed, more similarities are found with leader sequence of the ant species *S. invicta* (Fig. 7). A part from the leader sequence, we noticed specific amino-acid differences exclusively exhibited by the sol i3-like sequence identified in *T. bicarinatum*. While all compared species share asparagine (position 27) and arginine residues at positions 33 and 78, *T. bicarinatum* has lysine residues (Fig. 7). Taken together all the consistent similarities and specific differences, we are confident of the existence of sol i3 allergen-like compounds in *T. bicarinatum* venom. Special signature of sol i3-like

peptide found in the present study, could be useful for future synthesis of recombinant allergen from Soli i 3 aiming at preventing serious anaphylactic reactions.

4.1.1.3. Other allergens? In our ESTs venom gland library, we have not detected the presence of the major allergens commonly described in Hymenoptera and ants namely phospholipases and hyaluronidases (Hoffman, 2006). Even though the library provided here is far to cover all the transcripts synthesized by the venom gland cells of *T. bicarinatum*, this result could corroborate with previous work on the pavement ant *Tetramorium caespitum* where no activities of these enzymes have been detected (Von Sicard et al., 1989). Venom composition reveals significantly different protein expression across ant sub-family (Lima and Brochetto-Braga, 2003) and although ants share some common proteins in their venoms, each group appears to have a number of possibly specific components which further complicates ant venom investigations (Hoffman, 2010). For instance, Wiese et al. (2006, 2007) described significant differences between the primitive Australian ants of the genus *Myrmecia* and other ants and wasps. For instance, while the major expressed allergen in the Australian ant *Myrmecia pilosula* was Myr p 1 (Palma, 2006) that in the fire ant *S. invicta* is Soli i 2 group among which we find Soli i 1 which is a phospholipase that is similar to those found in wasp venoms and belongs to the lipoprotein lipase family (Hoffman et al., 1988). Differences in venom allergen composition through Hymenoptera groups could likely be attributed to their phylogenetic and behavior differences (e.g prey capture and feeding ways, defensive and offensive strategies, habitat, etc.).

4.1.2. Paralytic peptides

In the present library, paralytic peptides were exclusively represented by 20 ESTs being part of the insect cytokine paralytic peptide precursor uENF2 (clustered in contig 6, Table 2). Paralytic peptides are insect cytokines belonging to the ENF peptide family (Kamimura et al., 2001) that occur naturally in the hemolymph of some lepidopteran insects. They have been shown to cause rapid and rigid paralysis when injected into some lepidopteran larvae (Skinner et al., 1991). In the same way, Ishii et al. (2010) described that active paralytic peptides promoted *in vitro* and *in vivo* phagocytosis of *Staphylococcus aureus* and induced *in vivo* phosphorylation of p38 mitogen-activated protein kinase in the fat body. Intriguingly, it was shown that mRNA from silkworm encoding paralytic peptide precursor is tricistronic and encoded for two cytokine precursor-like proteins, namely uENF1 and uENF2 as in operon of prokaryote cells (Kanamori et al., 2010). However, the mechanism by which the three proteins are inter-translated remains unclear. In our study, transcripts encoding for paralytic-like peptides covered 11% of the paralytic peptide sequence from the moth *Neogurelca himachala sangaica* (*E*-value = 3,00 *E*-06). As far as we are aware, no transcript from the ENF peptide family has been described from venom of other Hymenoptera meaning that the identified sequence here would help to support gene sequence comparison for future investigations on insect

cytokine paralytic peptides and answer some unresolved points regarding the mechanism by which non monocistronic mRNA are structured in Eukaryotes.

4.1.3. Other peptides

Other transcripts in *T. bicarinatum* ESTs venom library encode different peptides and enzymes with a slight dominance of secapin/secapin-like preproteins. Secapin has been previously described as a major component in queen-bee venom after melittin (the principle allergen in bees) and minor component in worker-bee venom (Gauldie et al., 1976; Vlasak and Kreil, 1984). Preprosecapin (Secapin preprotein) is also found in higher amount in queen-bee. While the activation of the mellitin precursor is extracellular to venoms glands taking place in venom reservoirs, that of secapin-precursor is thought to be intracellular processing (Vlasak and Kreil, 1984). Investigations on physiological function of secapin are rare and its role remains unknown up to now (Palma, 2006).

One singlet was also identified (S381) in our library as disintegrin and metalloproteinase belonging to the ADAM family proteins. The presence of metalloproteinases in animal venoms has been intensively described in snakes and vipers (e.g. Boukhalfa-Abib et al., 2009; Lingott et al., 2009) showing their involvement in most of the systemic and local damaging effects of bite envenoming with the vascular endothelium, platelets and coagulation factors as major targets (Moura-da-Silva and Baldo, 2012). In social *Hymenoptera*, the occurrence of myonecrosis and skin damage has been reported in some victims of massive attacks (Stripja and Boonpuknavig, 1972). As far as we know, this is the second record about the presence of this protein in the venom of ants. The first disintegrating was recently reported in the venom of the fire ant *S. invicta* (Dos Santos Pinto et al., 2012) and its presence was validated by the use of specific metalloproteinase inhibitors through gelatin zymography. Future characterization and investigation of this protein from other *Hymenoptera* should better elucidate its specific function as venom compound and its possible potential as pharmacological/therapeutic molecule. Indeed, it was recently shown that Leucurolysin-B (a metalloproteinase class P-III, isolated from the snake *Bothrops leucurus*) has potent cytotoxic effect in a micromolar range against the tested cancer cell lines (Gabriel et al., 2012).

4.2. ESTs related to toxic functions (toxin-like venom candidates)

In this group, the most expressed transcripts matched zinc-metalloproteinase. Metalloproteinases are omnipresent and broadly involved in metabolism regulation with the ability to degrade proteins and to selectively hydrolyze specific peptide bonds (Gomis-Rüth, 2003). With cysteine- and serine-peptidase, metallo-peptidases constitute the major classes of peptidases with more consistency based on the number of their described genes (Turner and Nalivaeva, 2007). Two distinct zinc metallopeptidase families are well studied, namely M13 and M2 whose representative (neutral endopeptidase and peptidyl dipeptidase) require zinc for their activity and contain one

or two histidine residues in the zinc binding motif. Both representative have widespread tissue distribution and a large collection of possible physiological roles (Campbell et al., 1998).

The expression of metalloproteinases in venom compounds could play a major role in immune host suppression and targeting one of the major host defense cascade namely the coagulation cascade. It was previously suggested that zinc-metalloproteinases from snakes and also wasp venoms are involved in the blood clotting following envenomations which contributes to an effective spread of their venoms and gives a more digestible prey (Danneels et al., 2010). The abundance of metalloproteinases in a solitary hunting wasp of the species *Orancistrocerus drewensi* was also suggested to be related to the processing of inactive precursors of venom peptides that have to be processed into active forms (Baek et al., 2009). In line with this, the record of metalloproteinases in *T. bicarinatum* could suggest that it is a functional component of the venom probably related to stimulate precursor venom peptides in active forms and play a role in immune prey suppression. Given that no signal peptide was detected in all the metalloproteinase transcripts identified in our study (data not shown), we cannot exclude that these molecules could perform other roles beyond physiological functions.

Some proteins/enzymes that are potentially involved in targeting major host defense cascades were also revealed in our library such as sialidase, prophenoloxidase and serine protease. Phenoloxidase is a multicopper oxidase which controls the innate immune process of the melanization of pathogens and damaged tissues with the involvement of serine proteases in this phenoloxidase cascade of insects. This process is known to innately produce highly reactive and toxic quinone intermediates (Christensen et al., 2005; Danneels et al., 2010).

4.3. Clusters relevant to cellular functions

From a total of 364 ESTs that presented significant hits in the databases, 38% represent proteins involved in various cellular functions. The most dominant expressed cellular transcript found was the histone-lysine N-methyltransferase which plays a critical role in the regulation of gene transcription through histone methylation. Given the relative high length that the transcripts of histone-lysine N-methyltransferases identified in our venom gland library represent (average length 790 bp) and their different amino-acid sequence features (data not shown), we rather consider them as multigenes family. The prevalence of these molecules is consistent with previous work on spider venomous gland using ESTs showing the abundance of gene expression and translation related proteins and suggesting that this result may reflect the functional feature of this specialized tissue in the production and secretion of substances involved in feeding and predator interactions (Fernandes-Pedrosa et al., 2008).

Interestingly, we note the high presence of transposase expressed in our dbEST. As we have no evidence of the actual level of expression of this protein and for the existence of transposable elements, we cannot state their involvement in toxin gene diversity. Further investigations

on sequences repeats and inverted repeats and the genes they harbor may explain the role that this enzyme plays in the venom glands.

One EST encoding for odorant-binding protein (OBP) was also found in *T. bicarinatum* venom. This protein is called as General Odorant Binding Protein (GOBP) in insects and is involved in insect chemoreception. Even though this protein is known to interact between odorants and the elements of the sensillar lymph of insects, its physiological role in olfactory transduction remains unclear (Pelosi and Maida, 1995). In addition, the presence of this protein in venom compounds is uncommon and up to now, there's only one record of its presence in the venom of parasitoid wasp (Danneels et al., 2010). Recent studies based on search for OBPs in the genome of the parasitoid wasp *N. vitripennis* revealed 90 different sequences encoding putative OBPs and demonstrated the extremely dynamic evolution of the insect OBP family (Vieira et al., 2012).

4.4. No hits

Clusters that were not related to any known sequences in databases account for 26% of the total ESTs library. This high level is not surprising, since there is a limited annotated sequence from other ant species and families. The abundance of these ESTs (Table 1) and the presence of putative signal peptide and S–S bridge arrangement (data not shown) suggest that they may encode venom compounds and indicate that number of new toxin-candidates and proteins are still to be discovered from this ant venom.

5. Conclusion

By using Expressed Sequencing Tag technique, we described specific precursors encoding peptides and proteins expressed in the venom gland cells of the Brazilian terrestrial ant *T. bicarinatum*. In addition, we also expected the cellular transcripts that may represent a general panorama of the physiological events taking place in the venom glands, providing a survey of the gene expression from this specialized secretory tissue. Although there are already available sequences from other ants in the databases, mainly from *S. invicta* and *M. pilolusa*, this is the first time that ant venom gland transcripts are described. The generation of a more extensive sequence data and the characterization of this venom with proteomic approaches will enable a more confident overview of the toxin expression profile and enable comparative studies with other venom profiles from other Hymenoptera and arthropods in general.

Acknowledgments

Funding provided by the Centre Universitaire de Formation et de Recherche Jean-François Champollion; Wafa Bouzid is funded by a grant from the Région Midi-Pyrénées. We are grateful to J. Orivel for providing and identification of ant specimens and David J. Read (MRC, Toxicology unit, Leicester, UK) for proofreading the manuscript.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

References

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651–1656.
- Anderson, J.M., Oliveira, F., Kamhawi, S., Mans, B.J., Reynoso, D., Seitz, A.E., Lawyer, P., Garfield, M., Pham, M., Valenzuela, J.G., 2006. Comparative salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. *BMC Genomics* 7, 52.
- Baek, J.H., Woo, T.H., Kim, C.B., Park, J.H., Kim, H., Lee, S., Lee, S.H., 2009. Differential gene expression profiles in the venom gland/sac of *Orancistrocerus drewseni* (Hymenoptera: Eumenidae). *Arch. Insect Biochem. Physiol.* 71, 205–222.
- Boukhalfa-Abib, H., Meksem, A., Laraba-Djebbari, F., 2009. Purification and biochemical characterization of a novel hemorrhagic metalloproteinase from horned viper (*Cerastes cerastes*) venom. *Comp. Biochem. Physiol. C: Pharmacol. Toxicol.* 150, 285–290.
- Campbell, D.J., Anastasopoulos, F., Duncan, A.M., James, G.M., Kladis, A., Briscoe, T.A., 1998. Effects of neutral endopeptidase inhibition and combined angiotensin converting enzyme and neutral endopeptidase inhibition on angiotensin and bradykinin peptides in rats. *J. Pharmacol. Exp. Ther.* 287, 567–577.
- Christensen, B.M., Li, J., Chen, C.-C., Nappi, A.J., 2005. Melanization immune responses in mosquito vectors. *Trends Parasitol.* 21, 192–199.
- Danneels, E.L., Rivers, D.B., De Graaf, D.C., 2010. Venom proteins of the parasitoid wasp *Nasonia vitripennis*: recent discovery of an untapped pharmacopee. *Toxins* 2, 494–516.
- Dos Santos Pinto, J.R.A., Fox, E.G.P., Saidenberg, D.M., Santos, L.D., Da Silva Menegasso, A.R., Costa-Manso, E., Machado, E.A., Bueno, O.C., Palma, M.S., 2012. Proteomic view of the venom from the fire ant *Solenopsis invicta* Buren. *J. Proteome Res.* 11, 4643–4653.
- Fernandes-Pedrosa, M. de F., Junqueira-de-Azevedo, I. de L.M., Gonçalves-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, 279.
- Fry, B.G., Winter, K., Norman, J.A., Roelants, K., Nabuurs, R.J.A., Van Osch, M.J.P., Teeuwisse, W.M., Van der Weerd, L., McNaughtan, J.E., Kwok, H.F., Scheib, H., Greisman, L., Kochva, E., Miller, L.J., Gao, F., Karas, J., Scanlon, D., Lin, F., Kuruppu, S., Shaw, C., Wong, L., Hodgson, W.C., 2010. Functional and structural diversification of the Anguimorpha lizard venom system. *Mol. Cell. Proteomics* 9, 2369–2390.
- Gabriel, L., Sanchez, E., Silva, S., Santos, R., 2012. Tumor cytotoxicity of leucurolysin-B, a P-III snake venom metalloproteinase from *Bothrops leucurus*. *J. Venom Anim. Toxin. Incl. Trop. Dis.* 18, 24–33.
- Gauldie, J., Hanson, J.M., Rumjanek, F.D., Shipolini, R.A., Vernon, C.A., 1976. The peptide components of bee venom. *Eur. J. Biochem.* 61, 369–376.
- Gilly, W.F., Richmond, T.A., Duda, T.F., Elliger, C., Lebaric, Z., Schulz, J., Bingham, J.P., Sweedler, J.V., 2011. A diverse family of novel peptide toxins from an unusual cone snail, *Conus californicus*. *J. Exp. Biol.* 214, 147–161.
- Gomis-Rüth, F.X., 2003. Structural aspects of the metzincin clan of metalloendopeptidases. *Mol. Biotechnol.* 24, 157–202.
- Gouzy, J., Carrere, S., Schiex, T., 2009. FrameDP: sensitive peptide detection on noisy matured sequences. *Bioinformatics* 25, 670–671.
- Gruber, C.W., Muttenthaler, M., 2012. Discovery of defense- and neuro-peptides in social ants by genome-mining. *PLoS One* 7, e32559.
- Hoffman, D.R., Dove, D.E., Moffitt, J.E., Stafford, C.T., 1988. Allergens in Hymenoptera venom. XXI. Cross-reactivity and multiple reactivity between fire ant venom and bee and wasp venoms. *J. Allergy Clin. Immunol.* 82, 828–834.
- Hoffman, D.R., Smith, A.M., Schmidt, M., Moffitt, J.E., Guralnick, M., 1990. Allergens in Hymenoptera venom. XXII. Comparison of venoms from two species of imported fire ants, *Solenopsis invicta* and *richteri*. *J. Allergy Clin. Immunol.* 85, 988–996.
- Hoffman, D.R., 2006. Hymenoptera venom allergens. *Clin. Rev. Allergy Immunol.* 30, 109–128.
- Hoffman, D.R., 2010. Ant venoms. *Curr. Opin. Allergy Clin. Immunol.* 10, 342–346.

- Inagaki, H., Akagi, M., Imai, H.T., Taylor, R.W., Kubo, T., 2004. Molecular cloning and biological characterization of novel antimicrobial peptides, pilosulin 3 and pilosulin 4, from a species of the Australian ant genus *Myrmecia*. Arch. Biochem. Biophys. 428, 170–178.
- Inagaki, H., Akagi, M., Imai, H.T., Taylor, R.W., Wiese, M.D., Davies, N.W., Kubo, T., 2008. Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (Jack Jumper Ant). Arch. Biochem. Biophys. 477, 411–416.
- Ishii, K., Hamamoto, H., Kamimura, M., Nakamura, Y., Noda, H., Imamura, K., Mita, K., Sekimizu, K., 2010. Insect cytokine paralytic peptide (PP) induces cellular and humoral immune responses in the silkworm *Bombyx mori*. J. Biol. Chem. 285, 28635–28642.
- Kamimura, M., Nakahara, Y., Kanamori, Y., Tsuzuki, S., Hayakawa, Y., Kiuchi, M., 2001. Molecular cloning of silkworm paralytic peptide and its developmental regulation. Biochem. Biophys. Res. Commun. 286, 67–73.
- Kanamori, Y., Hayakawa, Y., Matsumoto, H., Yasukochi, Y., Shimura, S., Nakahara, Y., Kiuchi, M., Kamimura, M., 2010. A eukaryotic (insect) tricistronic mRNA encodes three proteins selected by context-dependent scanning. J. Biol. Chem. 285, 36933–36944.
- King, M.A., Wu, Q.X., Donovan, G.R., Baldo, B.A., 1998. Flow cytometric analysis of cell killing by the jumper ant venom peptide pilosulin 1. Cytometry 32, 268–273.
- King, G.F., 2011. Venoms as a platform for human drugs: translating toxins into therapeutics. Expert Opin. Biol. Ther. 11, 1469–1484.
- Liang, S., 2004. An overview of peptide toxins from the venom of the Chinese bird spider *Selenocosmia huwena* Wang [= *Ornithoctonus huwena* (Wang)]. Toxicon 43, 575–585.
- Lima, P.R. de, Brochetto-Braga, M.R., 2003. Hymenoptera venom review focusing on *Apis mellifera*. J. Venom Anim. Toxin. Incl. Trop. Dis. 9, 149–162.
- Lingott, T., Schleberger, C., Gutiérrez, J.M., Merfort, I., 2009. High-resolution crystal structure of the snake venom metalloproteinase BaP1 complexed with a peptidomimetic: insight into inhibitor binding. Biochemistry 48, 6166–6174.
- Magrane, M., Consortium, U., 2011. UniProt Knowledgebase: a hub of integrated protein data. Database (Oxford) 2011, bar009.
- Moura-da-Silva, A.M., Baldo, C., 2012. Jararhagin, a hemorrhagic snake venom metalloproteinase from *Bothrops jararaca*. Toxicon 60, 280–289.
- Padavattan, S., Schmidt, M., Hoffman, D.R., Marković-Housley, Z., 2008. Crystal structure of the major allergen from fire ant venom, Sol i 3. J. Mol. Biol. 383, 178–185.
- Palma, M.S., Brochetto-Braga, M.R., 1993. Biochemical variability between venoms from different honey-bee (*Apis mellifera*) races. Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 106, 423–427.
- Palma, M.S., 2006. Insect venom peptides. In: Abba, J.K. (Ed.), Handbook of Biologically Active Peptides, pp. 409–416.
- Pelosi, P., Maida, R., 1995. Odorant-binding proteins in insects. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 111, 503–514.
- Petersen, T.N., Brunak, S., Von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.
- Quintero-Hernández, V., Ortiz, E., Rendón-Anaya, M., Schwartz, E.F., Becerril, B., Corzo, G., Possani, L.D., 2011. Scorpion and spider venom peptides: gene cloning and peptide expression. Toxicon 58, 644–663.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBoss: the European molecular biology open software suite. Trends Genet. 16, 276–277.
- Ruiming, Z., Yibao, M., Yawen, H., Zhiyong, D., Yingliang, W., Zhijian, C., Wenxin, L., 2010. Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. BMC Genomics 11, 452.
- Santos, L.D., Pieroni, M., Menegasso, A.R.S., Pinto, J.R.A.S., Palma, M.S., 2011. A new scenario of bioprospecting of Hymenoptera venoms through proteomic approach. J. Venom Anim. Toxin. Incl. Trop. Dis. 17, 249–263.
- Skinner, W.S., Dennis, P.A., Li, J.P., Summerfelt, R.M., Carney, R.L., Quistad, G.B., 1991. Isolation and identification of paralytic peptides from hemolymph of the lepidopteran insects *Manduca sexta*, *Spodoptera exigua*, and *Heliothis virescens*. J. Biol. Chem. 266, 12873–12877.
- Stöcklin, R., Vorherr, T., 2011. Future Perspectives of Venoms for Drug Discovery. Pharmanufacturing: the International Peptide Review, pp. 21–24.
- Stripja, V., Boonpuknavig, V., 1972. Renal failure and myonecrosis following wasp sting. Lancet 1, 749–750.
- Turner, A.J., Nalivaeva, N.N., 2007. New insights into the roles of metalloproteinases in neurodegeneration and neuroprotection. Int. Rev. Neurobiol. 82, 113–135.
- Vanhoye, D., Bruston, F., Nicolas, P., Amiche, M., 2003. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. Eur. J. Biochem. 270, 2068–2081.
- Vieira, F.G., Forêt, S., He, X., Rozas, J., Field, L.M., Zhou, J.-J., 2012. Unique features of odorant-binding proteins of the parasitoid wasp *Nasonia vitripennis* revealed by genome annotation and comparative analyses. PLoS One 7, e43034.
- Vlasak, R., Kreil, G., 1984. Nucleotide sequence of cloned cDNAs coding for preprosecapin, a major product of queen-bee venom glands. Eur. J. Biochem. 145, 279–282.
- Von Sicard, N.A., Candy, D.J., Anderson, M., 1989. The biochemical composition of venom from the pavement ant (*Tetramorium caespitum* L.). Toxicon 27, 1127–1133.
- Wagstaff, S.C., Harrison, R.A., 2006. Venom gland EST analysis of the saw-scaled viper, *Echis ocellatus*, reveals novel alpha9beta1 integrin-binding motifs in venom metalloproteinases and a new group of putative toxins, renin-like aspartic proteases. Gene 377, 21–32.
- Wiese, M.D., Chataway, T.K., Davies, N.W., Milne, R.W., Brown, S.G.A., Gai, W.P., Heddle, R.J., 2006. Proteomic analysis of *Myrmecia pilosula* (jack jumper) ant venom. Toxicon 47, 208–217.
- Wiese, M.D., Brown, S.G.A., Chataway, T.K., Davies, N.W., Milne, R.W., Aulfrey, S.J., Heddle, R.J., 2007. *Myrmecia pilosula* (Jack Jumper) ant venom: identification of allergens and revised nomenclature. Allergy 62, 437–443.
- Wu, Q.X., King, M.A., Donovan, G.R., Alewood, D., Alewood, P., Sawyer, W.H., Baldo, B.A., 1998. Cytotoxicity of pilosulin 1, a peptide from the venom of the jumper ant *Myrmecia pilosula*. Biochim. Biophys. Acta 1425, 74–80.