Characterisation of Ant Venom Peptides and Proteins

Samira Ryma Aili

BMedSc (Hons)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy from:

School of Life Sciences,

University of Technology Sydney

2018

Neurotoxin Research Group: UTS

Certificate of Authorship and Originality

I, Samira Ryma Aili, certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program Scholarship.

Signature of Student

Production Note:
Signature removed prior to publication.

Date: 16th April 2018

Table of Contents

Acknowledgements

First and foremost I would like to thank my supervisor Graham Nicholson. I couldn't have chosen a better supervisor to do my PhD with. I am forever grateful for all you've done for me and your patience with my constant emails/ questions and visits to your office. I appreciate all the extra weekend/ plane time/ after-hours work you've dedicated to helping me. I chose you as a supervisor because you came across as the most organised of my lecturers – a fact that proved to be very true and has inspired me to follow in your footsteps. I have learnt so much in our 4/5 years together about travelling, organising my diary, my mac, communicating, fixing figures and of course science-ing. I hope more students will benefit from your wisdom and you continue being the greatest supervisor one could hope for. Your constant support when stuff goes wrong and attitude of "if it was easy everyone would be doing it" and reassurance definitely helped ease my PhD journey.

To Matt, my other supervisor who is also always there for me and is always willing to answer my questions after initially saying no. Your knowledge about everything is a constant inspiration. Despite being initially terrified of you, it wasn't long till I realised I was very wrong. Your help in the first couple of years the lab was invaluable, especially when Graham is away. I'm glad I was a member of the proteomics family whose support is part of the reason I stayed on for a PhD. You have taught me so much about science but also a lot about other random things such as metal bands, bikes and cars.

I would like to thank Dr Axel Touchard, as without him this project would not have taken place. The number of ants you've collected and glands you've dissected for me is innumerable. Thank you to Dr Pierre Escoubas as well, for his contribution in collections and feedback throughout the project. A big thanks also goes to Hadrien L., Frederic P., Alain D., and Jerome., who've all played a role in the various ant collections.

Thank you Regan Hayward! A life saver with all your help in analysing the transcriptomics and excel tricks. I would also like to thank A/ Prof Garry Myers for your guidance with the transcriptomics work.

I'd like to thank Prof Kini for giving me the opportunity to go over to Singapore and funding the part of the transcriptomics project. Special thanks to Mrinalini for introducing me to the land of transcriptomes as well and her feedback and guidance.

Another big thank you goes to Sandy P. Gonzalez who has been supporting me in many aspects of the last stages of my PhD thesis stress and helping me with my transcriptomics data analysis and interpretation as well.

I would like to thank Dr. Jaye Liu and Prof. Liz Harry from the iThree institute, UTS for their help in the antibacterial activity assays and for providing me with the bacteria.

To the proteomics crew (Jacqueline, Ben, Iain, Marcelo, Krish, Kate, Michael, Ronnie, and Jerran) your support and friendship is invaluable. You adopted me as a part of your group, seeing as I was the only person in my lab group. Our Lorne adventures will be forever remembered and cherished. A special thank you to Mike, my metabolic buddy and computer shortcut wiz, for all your help and friendship particularly in the last week with teaching approaching. As well as to Jacqueline, aka my breakfast and gym partner, Thursday traditions need to be me maintained forever! Your blonde moments never cease to entertain me; it's been great sharing your crochet and knitting hobbies.

I'd also like to thank all my other UTS friends Brendan, Isa, Megan, Chris, Louis, Emma, Dan and Louis. In particular, thank you to Emma my sprout buddy for always being there for me and helping me format this thesis the day after you got back from your holidays. Your strength and commitment is always an inspiration and I'm glad I've had the privilege of knowing and befriending you.

I'd also like to thank the markers for taking the time to review this thesis and the Australian Government Postgraduate Award Scholarship I received throughout my PhD candidature.

Finally, I would like to thank my Family. My mum Ghania, dad Arezki, and sisters (Sadia, Hanna, Massilia, Zahia and Rama). Your support and understanding of my long hours spent away from home is truly appreciated. Without you I wouldn't be here and I hope to one day be able to explain what I actually do "all day at uni". Particular thanks to my parents for sacrificing a lot to give us a good education. I hope I made you proud.

List of Publications

1) **Aili, S. R.,** Touchard, A., Escoubas, P., Padula, M. P., Orivel, J., Dejean, A., & Nicholson, G. M. (2014). Diversity of peptide toxins from stinging ant venoms. *Toxicon*, *92*, 166-178.

2) Touchard, A., **Aili, S. R.,** Fox, E. G. P., Escoubas, P., Orivel, J., Nicholson, G. M., & Dejean, A. (2016). The biochemical toxin arsenal from ant venoms. *Toxins*, *8*(1), 30.

3) Touchard, A., Koh, J., **Aili, S. R.,** Dejean, A., Nicholson, G. M., Orivel, J., & Escoubas, P. (2015). The complexity and structural diversity of ant venom peptidomes is revealed by mass spectrometry profiling. *Rapid Communications in Mass Spectrometry*, *29*(5), 385-396.

4) **Aili, S. R.,** Touchard, A., Koh, J. M., Dejean, A., Orivel, J., Padula, M. P., & Nicholson, G. M. (2016). Comparisons of protein and peptide complexity in poneroid and formicoid ant venoms. *Journal of proteome research*, *15*(9), 3039-3054.

5) **Aili, S. R.,** Touchard, A., Petitclerc, F., Dejean, A., Orivel, J., Padula, M. P., & Nicholson, G. M. (2017). Combined peptidomic and proteomic analysis of electrically stimulated and manually dissected venom from the South American bullet ant *Paraponera clavata*. *Journal of proteome research*, *16*(3), 1339-1351.

Conference Proceedings

Venoms to Drugs Conference, Noosa, QLD, Australia 2017
Oral Presentation
"An integrated proteomic and transcriptomic analysis of venom gland toxins from the bullet ant *Paraponera clavata*"

TEDx Youth Sydney 2017
Oral Presentation, fast ideas session
Ant venom to the rescue

AMP Amplify Ignite PhD pitch competition 2017
Oral presentation
"Ants: the solution to world hunger"

XII Congress of the Pan American Section of the International Society on Toxinology, Miami, USA 2016
Oral Presentation
"Ant venom as a novel source of bioinsecticide leads"

Sydney Protein Group Thompson Prize finalist, Sydney 2016
Oral Presentation
"Ant venom as a source of bioinsecticides"

The 21st Lorne Proteomics Symposium, Lorne, Victoria 2016
Poster presentation
"Peptidomic and proteomic comparison of electrically stimulated and manually dissected venom of the bullet ant *Paraponera clavata*"

University of Technology Sydney 3-minute thesis competition 2015
Oral presentation
"Ant venom derived insecticides"

The 18th world congress of the International Society on Toxinology, Oxford, United Kingdom 2015
Oral presentation
"Ant venom as a source of bioinsecticide and antimicrobial drug leads"

The 20th Lorne Proteomics Symposium, Lorne, Victoria 2015
Lightning talk and oral presentation: awarded best poster prize
"Characterisation of the peptide and protein content of ant venoms for use as bioinsecticide and antimicrobial leads"

New horizons 31st Combined Health Science Conference, Kolling Institute RNSH 2014
Poster presentation
"Characterisation of the Antimicrobial and Insecticidal Peptides From Ant Venoms"

APAF 2nd Proteomics and Beyond Symposium, Sydney 2014
Oral presentation
"Characterisation of Poneroid and Formicoid ant venoms"

The 19th Lorne Proteomics Symposium, Lorne, Victoria 2014
Lightning talk and oral presentation: awarded best poster prize
"Characterising the peptide and protein diversity of Neotropical and Australian ant venoms and the identification of novel peptide toxins for bioinsecticide discovery"

New horizons 30th Combined Health Science Conference, Kolling Institute RNSH 2013
Oral presentation
"Characterising the diversity of Neotropical and Australian ant venoms: Novel peptide libraries for bioinsecticide discovery"

List of Figures

***Please note that the Figures listed above include those in chapters 4 and 7 only.**

List of Tables

***Please note that the Tables listed above include those in chapters 4 and 7 only.**

Abbreviations

1,5-DAN	1,5-Diaminonaphtalene
1D-SDS-PAGE	One dimensional - sodium dodecyl sulphate - polyacrylamide gel electrophoresis
2D-SDS-PAGE	Two dimensional - sodium dodecyl sulphate - polyacrylamide gel electrophoresis
AA	Amino acid
AChE	Acetylcholinesterase
ACN	Acetonitrile
BCA	Bicinchoninic assay
BSA	Bovine serum albumin
Bt	*Bacillus thuringiensis*
BLAST	Basic local alignment search tool
CHCA	α-Cyano-4-hydroxycinnamic acid
CID	Collision induced dissociation
cDNA	Complementary deoxyribonucleic acid
Cys	Cysteine
DDT	Dichlorodiphenyltrichloroethane
DTT	Dithiothreitol
DNA	Deoxyribonucleic acid
EPA	Environmental Protection Agency
ESI-MS	Electrospray ionisation mass spectrometry
FA	Ferulic acid
FAO	Food and Agriculture Organisation
GABA	γ-Aminobutyric acid
GM	Genetically modified
Gln	Glutamine
Glu	Glutamic Acid

IEF	Isoelectric focussing
IPG	Immobilised pH gradient
ICK	Inhibitor cysteine knot
KD_{50}	Median knockdown dose
kDa	Kilodaltons
LD_{50}	Median lethal dose
LC-MS/MS	Liquid chromatography tandem mass spectrometry
m/*z*	Mass-to-charge ratio
NanoESI-QTOF MS	Nano electrospray ionisation quadrupole time-of-flight mass spectrometry
Lys	Lysine
MALDI-TOF MS	Matrix-assisted laser-desorption ionization/time-of-flight mass spectrometry
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
NIS	Normal insect saline
PAGE	Polyacrylamide gel electrophoresis
PI	Isoelectric point
PLA_2	Phospholipase A_2
RP-HPLC	Reversed-phase high-performance liquid chromatography
RPM	Revolutions per minute
SCX	Strong cation exchange
Ser	Serine
SDS	Sodium dodecyl sulphate
TFA	Trifluoroacetic acid
WHO	World Health Organisation

Abstract

Venom peptides are currently being developed as novel therapeutics and bioinsecticides. Given that ants use their venoms for predation and defence against insects, and other organisms, they are a potential source of these peptides. Although ants represent one of the largest groups of venomous animals, little is known about their venom composition. The present study therefore investigated the peptidome, proteome and transcriptome of a range of poneroid and formicoid ant venoms.

Initial experiments sought to confirm the insecticidal and antibacterial activity of whole ant venom using house crickets and minimum inhibitory concentration assays, respectively. Several ant venoms showed significant paralytic and insecticidal activity and others showed antibacterial activity peptides confirming the utility in studying ant venoms.

Subsequent experiments investigated the difference in venom composition obtained using differing venom collection methods: manual venom gland dissection or electrical stimulation. The peptide and protein components of the bullet ant (*Paraponera clavata*) were compared and revealed numerous proteins of which 96 could be assigned a biological function, and 70% of which were common to both collection methods. However, the peptidomic analysis revealed over 300 peptides of which only 30% were common to both collection methods. Therefore, each method reveals a unique set of peptides and proteins.

The peptide components of six different ant species were also characterised. The venoms were found to contain between 132–1032 peptides, but the large number of undescribed proteins and peptides highlighted the need for a transcriptomic investigation. Accordingly, an integrated approach using a combination of shotgun proteomics in parallel with Illumina sequencing of the venom gland transcriptome was used to identify toxins in the venom of *P. clavata*. A BLASTx search of the assembled contigs revealed 354 proteins with homology to existing toxins. Alignments of some of these toxins revealed novel insights into their role in ant venom.

A Tox|Note analysis revealed several predicted novel peptide toxins, with some conforming to the conotoxin cysteine frameworks VI/VII and framework I, both of which have yielded therapeutic drug and bioinsecticide leads. The translated transcriptome was then used as a database to query the MS/MS data obtained from the shotgun experiment which identified 44 toxins. Several of these were not identified in the transcriptome BLASTx search, such as δ-

paraponeritoxin-Pc1e (formerly poneratoxin). These results reveal the advantages of combining proteomic and transcriptomic methods, and further demonstrates the richness and diversity of ant venoms as potential sources of bioactive compounds.

Chapter One:

An overview of the thesis

Chapter One: Overview of the Thesis

1.1. Introduction

Insect pests are a major threat to human health and the agricultural industry. Vector-borne diseases affect almost 500 million people, and insect pests destroy almost 20% of the world's crop yields [1-4]. Our current methods of eliminating these insects using conventional synthetic insecticides is failing, due to resistance by the insects to all classes of insecticides and the harmful effects of the chemicals themselves to both human health and the environment. This problem is projected to become worse, due to the exponential increases in the world's population that is estimated to reach 12 billion by the year 2100 [5]. A larger population means we need to find ways of feeding these people, as current crop yields will not be sufficient. A potential solution to this problem lies in insecticidal toxins from animal venoms.

Animal toxinology is an increasingly popular field of research due to the rich diversity in bioactive compounds contained in animal venoms. Venoms are rich chemical libraries injected by one animal into another for defence or predation. They are made up of a mixture of peptides, proteins, enzymes and small molecules which affect multiple molecular targets. Millions of years of selection pressure have optimised this venom arsenal to produce peptides with high potency, selectivity, and stability, making them attractive sources for investigation of natural insecticides, therapeutics or pharmacological tools. Additionally, the limited amount of research into this field raises the need for further investigations to uncover the true venom arsenal, particularly from neglected taxa such as ants. Thus far, the majority of research has focused on snakes, spiders and cone snails. Hymenopterans are a large group of arthropods (>120,000 spp.) and include bees (Apidae), wasps (Vespidae) and ants (Formicidae), however the most investigated species are bees and wasps due to their medically relevant stings. Ants are a largely overseen group of animals despite their success in all different types of environments and their possession of a venom that is used for predation, defence and communication. This thesis is an investigation of the venom gland components of several ant species to determine their true diversity and potential for future investigation as bioinsecticide or therapeutic drug candidates.

1.2. Aims of the thesis

The previous section outlines a number of key areas that require further investigation. From these, a set of aims were constructed to address within this dissertation. The overarching aim of this thesis was to investigate the potential of ant venoms as candidates for bioinsecticide and therapeutic leads. Specifically, these aims are:

- Aim one: Determine which of the investigated ant venoms possess insecticidal/ antibacterial activity
- Aim two: Determine the best collection method for ant venom by comparing the peptide and protein profiles of two different techniques (electrical stimulation and venom gland dissection)
- Aim three: Compare the venom peptidome and proteome of the two major clades of ants
- Aim four: Perform a holistic proteomic and transcriptomic investigation of the venom gland from the bullet ant *Paraponera clavata*

This thesis will present the first in-depth investigation of ant venom peptidomes, proteomes and combined proteome/ transcriptome, where we were able to link the transcripts to the expressed proteins. In addition, it describes the differences in the venom profiles of different venom collection methods. The ant venoms used in this thesis were all collected by Dr. Axel Touchard from various regions in French Guiana.

1.3. Dissertation organisation

Following this chapter, this dissertation is organised in the following manner:

- Chapters 2 and 3: Critical reviews on ant venom composition

  Both of these articles were published in the journals *Toxicon* and *Toxins*. They aim to provide background information on what is currently known about ant venom composition – particularly the peptide component.
- Chapter 4: Ant venom insecticidal and antibacterial activity

  This chapter describes the insecticidal and antibacterial activity of various ant venoms to confirm their potential as sources of bioactive peptides. This chapter aims to answer Aim 1 of this dissertation.

- Chapter 5: Ant venom collection methods

  This chapter describes, for the first time, the differences in venom profiles of ant venom collected by electrical stimulation versus ant venom collected by venom gland dissection by comparing the peptide and protein composition of the two samples. This chapter aims to answer Aim 2 of this dissertation and is published in the *Journal of Proteome Research*.
- Chapter 6: Ant venom clade diversity

  This chapter is the first in-depth investigation of ant venom peptide and protein profiles as well as the first investigation comparing ants of different clades. This chapter uncovered the huge diversity in ant venoms – addressing Aim 3 and is published in the *Journal of Proteome Research*
- Chapter 7: Proteomic/ transcriptomics investigation of *Paraponera clavata*

  This chapter provides the first combined proteome/ transcriptome investigation of ant venom gland. It focuses on the identification of the toxin components of the venom as well as on the identification of novel toxin peptides for future investigation as bioactive molecules. This chapter will be submitted for publication to the journal *Scientific Reports* and aims to address Aim 4 of this dissertation.
- Chapter 8: Concluding remarks and future directions

  This is the final chapter to the thesis and contains the concluding remarks and recommendations for future directions of the project.

- Supplementary material

  Supplementary material for Chapter 7 is provided as a USB, as much of the supplementary material is too large for printing in the dissertation.

Chapter Two

Diversity of peptide toxins in stinging ant venoms

Chapter Two: Diversity of Peptide Toxins in Stinging Ant Venoms

Compound Abstract

Prior to the publication of this chapter, no review was available summarising the known literature on ant venom peptides. This was despite the fact that there had been several investigations on various ant peptides over the years and that ants are one of the most successful species on earth. They use their venom for various tasks such as defence and predation. This chapter aimed to shed light on the role of ant venom and the known ant venom peptides based on previously published investigations. This chapter summarises all the published information of ant venom peptides including their sequences, homology to other ant venom peptide and their pharmacological activity (if known). We found that ant venom peptides seem to have a diverse range of functions and effects with high levels of potency. Moreover, this review highlighted the low number of characterised peptides in relation to the total predicted number of peptides in ant venoms.

Certificate of authorship and originality

This paper was published in Toxicon journal © Elsevier Ltd. I certify that the work present in this chapter has not previously been submitted as part of the requirements for a degree. I also certify that I carried out the majority of the work presented in this paper.

- Samira Aili: wrote the majority (~70%) of the first draft and edited all subsequent drafts of the paper.
- Axel Touchard: wrote ~30% of the manuscript, proofed the full draft of the manuscript
- Matt Padula, Jerome Orivel, Alain Dejean: Proof read manuscript
- Pierre Escoubas: conceived the project idea and proof read manuscript
- Graham Nicholson: performed alignments, helped with direction of manuscript, proof-read manuscript and conceived project idea

Principal supervisor

Professor Graham Nicholson	Samira Ryma Aili
_____	_____
Signature	Signature
_____	_____
Date	Date

Toxicon 92 (2014) 166–178

Contents lists available at ScienceDirect

Toxicon

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

Review

Diversity of peptide toxins from stinging ant venoms

Samira R. Aili [a, 1], Axel Touchard [b, 1], Pierre Escoubas [c], Matthew P. Padula [a], Jérôme Orivel [b], Alain Dejean [b, d, **], Graham M. Nicholson [a, *]

[a] Neurotoxin Research Group, School of Medical & Molecular Biosciences, University of Technology Sydney, NSW 2007, Australia
[b] CNRS, UMR Écologie des Forêts de Guyane (EcoFoG), Campus Agronomique, BP 316, 97379 Kourou Cedex, France
[c] VenomeTech, 473 Route des Dolines – Villa 3, 06560 Valbonne, France
[d] Laboratoire Écologie Fonctionnelle et Environnement, Université de Toulouse, 118 Route de Narbonne, 31062 Toulouse, France

ARTICLE INFO

Article history:
Received 3 September 2014
Accepted 27 October 2014
Available online 28 October 2014

Keywords:
Ant venom
Peptides
Venom biochemistry
Disulfide linkage
Chemotaxonomy

ABSTRACT

Ants (Hymenoptera: Formicidae) represent a taxonomically diverse group of arthropods comprising nearly 13,000 extant species. Sixteen ant subfamilies have individuals that possess a stinger and use their venom for purposes such as a defence against predators, competitors and microbial pathogens, for predation, as well as for social communication. They exhibit a range of activities including antimicrobial, haemolytic, cytolytic, paralytic, insecticidal and pain-producing pharmacologies. While ant venoms are known to be rich in alkaloids and hydrocarbons, ant venoms rich in peptides are becoming more common, yet remain understudied. Recent advances in mass spectrometry techniques have begun to reveal the true complexity of ant venom peptide composition. In the few venoms explored thus far, most peptide toxins appear to occur as small polycationic linear toxins, with antibacterial properties and insecticidal activity. Unlike other venomous animals, a number of ant venoms also contain a range of homodimeric and heterodimeric peptides with one or two interchain disulfide bonds possessing pore-forming, allergenic and paralytic actions. However, ant venoms seem to have only a small number of monomeric disulfide-linked peptides. The present review details the structure and pharmacology of known ant venom peptide toxins and their potential as a source of novel bioinsecticides and therapeutic agents.

1. Stinging ant biodiversity

Hymenopterans are among the most speciose group of venomous animals. With approximately 120,000 currently described species (van Emden, 2013), they are significantly more diverse than the major venomous phyla including spiders (44,906 species), snakes (3496 species), cone snails (3253 species), sea anemones (3248 species) and scorpions (1454 species) (Fautin, 2014; Hallan, 2005; Kohn and Anderson, 2009; Platnick, 2014; Uetz and Hošek, 2014). Among the stinging aculeate Hymenoptera, ants and wasps (superfamily Vespoidea) and bees together with sphecoid wasps (superfamily Apoidea) are sister groups (Johnson et al., 2013). Ants (family Formicidae) evolved from wasp-like ancestors between 115 and 135 million years ago (Brady et al., 2006) and became a diverse taxonomical group with ~13,000 extant species belonging to 21 subfamilies (Agosti and Johnson, 2005; AntWeb, 2014). Due to their ubiquitous nature in terrestrial environments, and the fact that they constitute 15–20% of the animal biomass in tropical rainforests (Hölldobler and Wilson, 1990; Wilson, 1990), ants are arguably amongst the most abundant venomous animals.

Ants that belong to the subfamilies Formicinae, Dolichoderinae, Aneuretinae and Dorylinae lost their ability to sting during evolution (Fig. 1). Instead, they usually spray their venoms or have a residual, but non-functional, abdominal stinger. Also, it is unclear if the recently discovered subfamily Aenictogitoninae is venomous or not, as only male castes have been seen and females (workers and queens) are yet to be described (Brady et al., 2006). The remaining 16 subfamilies are all stinging ants (Fig. 1) and comprise of ~9100 extant species. This makes ants taxonomically more diverse than scorpions, snakes and cone snails. However, this biodiversity is not equally distributed within stinging ant subfamilies (Figs. 1 and 2). For example, Myrmicinae is the most speciose ant subfamily, with

* Corresponding author.
** Corresponding author. Écologie des Forêts de Guyane, Campus Agronomique, BP 316, 97379 Kourou Cedex, France.
E-mail addresses: alain.dejean@wanadoo.fr (A. Dejean), Graham.Nicholson@uts.edu.au (G.M. Nicholson).
[1] Authors contributed equally to this review.

http://dx.doi.org/10.1016/j.toxicon.2014.10.021
0041-0101/

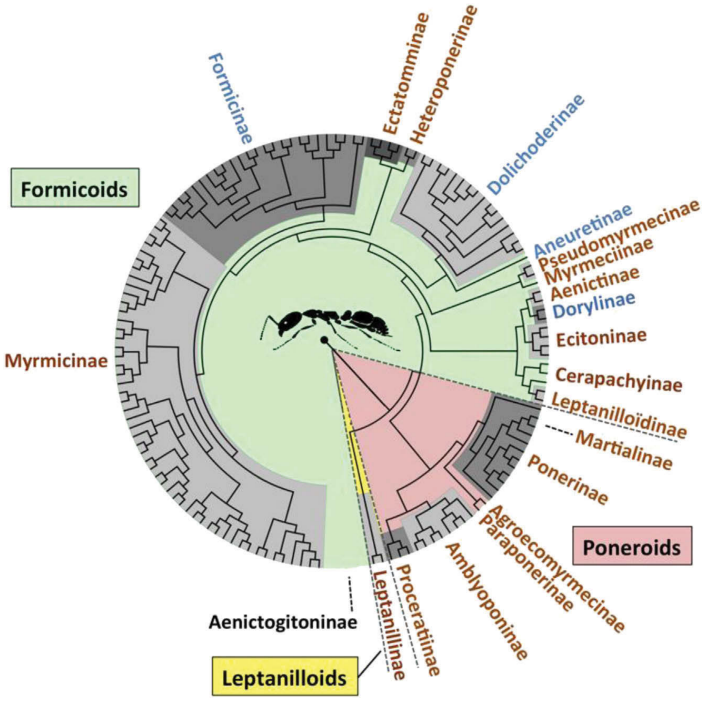


Fig. 1. Ant subfamily relationships as inferred from molecular phylogenetic studies. Phylogenetic relationships were generated from the S1573 TreeBASE data file (Moreau et al., 2006) using the FigTree v1.4.2 software package (http://tree.bio.ed.ac.uk/software/figtree/). Phylogenetic relationships for the subfamilies Aenictogitononae and Martialinae are currently unavailable. During evolution, four subfamilies lost their capacity to sting (blue text). Remaining subfamilies represent stinging ants (brown text). Ant clades are shaded green (Formicoids), red (Poneroids) and yellow (Leptanilloids). Females of subfamily Aenictogitoninae (black text) remain undiscovered and so this subfamily cannot be classified as either stinging or non-stinging. For clarification of colours in this figure, refer to the web version of this article. Note added in proof: Recently, the ant subfamilies Leptanilloïdinae, Cerapachyinae, Ecitoninae, Dorylinae, Aenictinae and Aenictogitoninae have been regrouped into one subfamily; Dorylinae (Brady, S., Fisher, B., Schultz, T., Ward, P., 2014. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol. Biol. 14, 93–106.).

~6500 extant species, with a widespread distribution throughout the world. However, ponerine ants that belong to the subfamily with the second highest number of ants, Ponerinae (~1200 species), are mainly confined to tropical rainforests (AntWeb, 2014; Johnson et al., 2013). Furthermore, the subfamilies Paraponerinae and Martialinae only contain a single ant species both of which are found in Neotropical areas. Thus, taxonomic diversity varies within each ant subfamily however there is little doubt that ant venoms likely constitute a vast source of unique bioactive toxins.

2. Ant venom functions

Ant venom is composed of a complex mixture of chemicals such as proteins, enzymes, biogenic amines, peptides, hydrocarbons, formic acid and alkaloids (Davies et al., 2004; Kem et al., 2004; Yi et al., 2003). All these compounds are produced by the venom gland, which consists of two free cylindrical elongated and convoluted tubes, linked to a venom reservoir (Ortiz and Mathias, 2006). The venom secreted by the tubular glands is stored in the reservoir, linked to the delivery apparatus and, for example, can deliver up to 130 μg of venom after each sting (Schmidt, 1990). The stinger itself is a modified ovipositor located at the distal base of the abdomen. Ants use their venom for several purposes such as a defence against predators/competitors and microbial pathogens, for predation, as well as for social communication (Orivel et al., 2001; Schmidt, 1982). Hence, ant venoms have evolved to carry out many different functions.

2.1. Offensivevenoms

Ants are one of the leading predators of invertebrates in most ecosystems (Brady et al., 2006). They have developed, through natural selection, a vast arsenal of behavioural adaptations and weapons to subdue their prey including trap-mandibles and potent venoms (Casewell et al., 2013). Ant venom has paralytic and lethal effects on many arthropods (Maschwitz et al., 1979; Orivel and Dejean, 2001) and many ants are generalist predators, preying on numerous classes of invertebrates. Nevertheless, many ants are specialised predators and only feed on a restricted group of species. Such specialized hunters prey exclusively on earthworms, isopods, centipedes, millipedes, polyxena, collembolan, termites, other ants or even spider eggs (Cerdá and Dejean, 2011). Solitary hunting is the most common hunting behaviour employed by primitive ants such as ponerines. However, many ants have also developed a cooperative hunting behaviour such as army ants exhibiting extreme group hunting behaviour.

The ecological diversity of ants is also revealed in their preference for various nesting habitats. Predatory ants are primarily ground, or litter-dwelling, predators. However, some ants have evolved predatory behaviours adapted to foraging in trees (arboreal ants) and exhibit adaptions to prevent their prey from escaping by flying away, jumping or dropping. Accordingly, venoms of solitary-foraging, arboreal predatory ants are believed to be more efficient than ground-dwelling species at rapidly immobilising prey (Orivel and Dejean, 2001). Thus, the use of venom as an offensive weapon

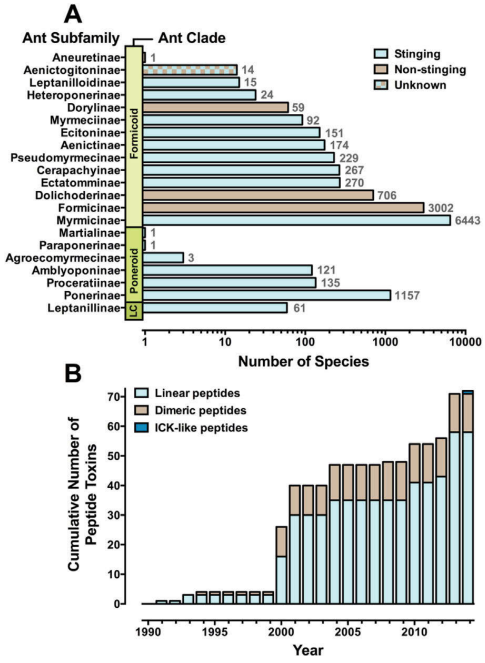


Fig. 2. (A) Species richness of ant subfamilies. Ants have been grouped according to three clades, where LC represents the single genus Leptanilloid clade. Stinging ants are represented by cyan bars and comprise around 70% of all ant species. Non-stinging ant subfamilies are depicted by brown bars. The total number of species in each subfamily is noted at right of each bar. The Aenictogitoninae subfamily is currently unclassified. Note added in proof: Recently, the ant subfamilies Leptanilloïdinae, Cerapachyinae, Ecitoninae, Dorylinae, Aenictinae and Aenictogitoninae have been regrouped into one subfamily; Dorylinae (Brady, S., Fisher, B., Schultz, T., Ward, P., 2014. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol. Biol. 14, 93–106.). (B) Cumulative total number of peptide-toxin sequences reported from ant venom studies since the first described venom peptide (poneratoxin) in 1991, showing the three main structural classes: cyan, linear peptides; brown, dimeric peptides; teal, ICK-like peptides. Ant venom peptides remain barely investigated with only 72 peptides sequenced to date. For clarification of colours in this figure, refer to the web version of this article.

is likely to be the major driver of the venom composition during evolution. This has been shown with the differing composition and toxicity of venoms from arboreal *versus* ground-dwelling species of *Pseudomyrmex* and *Pachycondyla* (Dejean et al., 2014; Orivel and Dejean, 2001; Touchard et al., 2014b). The wide ranging diet and hunting behaviours of ants are therefore likely to drive major differences among ant venom toxins.

2.2. Defensive venoms

Eusociality within hymenopteran colonies offers a range of evolutionary advantages including the capability of a mounting a collective defence against vertebrate and other arthropod predators, the ability to gather and store food and nutrients more efficiently, and to specialize in specific tasks, such as to care cooperatively for offspring (Wilson, 1971). Nevertheless, these benefits can only be realized if the colony can defend against large predators who find the large biomass of the colony a potential food source worth their effort, in contrast to preying upon solitary hymenopterans. The evolution of venom in hymenopterans therefore provided a mechanism of defence against large intelligent vertebrate predators and enabled them to develop complex societies. The combination of algesic and lethal actions of ant venom is therefore thought to be critical in the long term evolutionary success of insect stings to deter large predators (Schmidt, 2014). For example, some ant stings are known to be extremely painful for humans. These include stings by fire ants (*Solenopsis* spp.), ponerine ants (*Pachycondyla* spp.) or the bullet ant (*Paraponera clavata*). In particular, bullet ants have been classified as producing the most painful sting among all hymenoptera and the third most painful sting of all venomous animals (Schmidt et al., 1983; Starr, 1985).

It is also clear that some ants, such as the *Pogonomyrmex* group of harvester ants, have developed venoms primarily for defence against vertebrates (Schmidt and Snelling, 2009). For example, the venom of *Pogonomyrmex badius* is highly toxic towards mice, but not very toxic towards insects. Therefore, *Pogonomyrmex* ants do not appear to employ their venom to hunt, but use it exclusively as a deterrent against vertebrate predators (Schmidt and Blum, 1978a, 1978b), akin to the defensive role of bee venom against vertebrates. Some Pseudomyrmecine ants have also evolved a 'defensive venom' as part of a mutualistic relationship with myrmecophytes. Myrmecophytes are plants that provide a nesting place for a limited number of ant species, whilst the ants protect the myrmecophyte from defoliating arthropods and browsing mammals by stinging them. Natural selection has allowed ants that are known to have a painful sting to survive in such a habitat to the extent that some ants from the genera *Pseudomyrmex* and *Tetraponera* are obligate inhabitants of myrmecophytes. In some cases, ants use their venom in unusual ways. For example, *Pachycondyla tridentata* ants produce a foaming venom when disturbed and use their venom to paralyse their prey. This release of foam is a defence mechanism which is very effective against other small ants (Maschwitz et al., 1981).

It is therefore clear that ants have evolved venoms containing numerous toxins to induce pain, discomfort, paralysis and/or death in vertebrate and arthropod predators or prey. This is because protection of the nest, particularly protection of the brood and the queen, is a major concern for worker ants.

2.3. Antimicrobial properties of ant venoms

Ants are eusocial insects that typically live in colonies of relatives with a high population density. This increases the risk of introduction and spread of microbial pathogens. Consequently, ants have evolved strategies to inhibit microbial infections including the development and use of antimicrobial peptides. Firstly, predatory ant species may use their venom to inhibit internal pathogens present in captured prey that are brought back to the colony. In this way, the venom may protect the colony from infections following consumption of the prey species. In the ant venoms studied so far, this activity has been attributed to abundant linear, polycationic cytolytic peptides (see Section 3.1) that demonstrate potent antibacterial activity against both Gram-positive and Gram-negative bacteria (Cologna et al., 2013; Davies et al., 2004; Inagaki et al., 2004; Johnson et al., 2010; Kuhn-Nentwig, 2003; Mackintosh et al., 1995, 1998; Orivel et al., 2001; Rifflet et al., 2012; Viljakainen and Pamilo, 2008; von Sicard et al., 1989; Zelezetsky et al., 2005). More recently, similarity searches of ant genomes have revealed a number of tachystatins (antimicrobial chitin-binding peptides) with an inhibitor cystine knot (ICK) fold, as well as proline-rich abaecin-like, glycine-rich hymenoptaecin-like, insect defensin-like, and crustin-like antimicrobial peptides (Zhang and Zhu, 2012). These peptides may be part of the uncharacterized antimicrobial secretions from the thoracic metapleural or other glands that are spread over certain ants and the nest (Mackintosh et al., 1999; Yek and Mueller, 2011). Nevertheless, there is no evidence that these peptides are present in ant venoms.

3. Ant venom peptides

Alkaloid-rich ant venoms have been well-studied, particularly among the genera *Solenopsis* (Brand, 1978; Jones et al., 1996) and *Monomorium* (Jones et al., 1982, 2003, 2009, 1988). However, proteinaceous venoms remain highly understudied despite the fact that they appear to be very common in both the Poneroid and Formicoid clades of ant venoms. Thus, venoms from Poneroid ants have been shown to be rich in peptides especially venoms from the subfamilies Ponerinae (Cologna et al., 2013; Johnson et al., 2010; Orivel et al., 2001; Torres et al., 2014; Touchard et al., 2014a) and Paraponerinae (Piek et al., 1991a, 1991b; Rykaczewska-Czerwinska et al., 2008). Peptides have also been characterized from the venoms of Formicoid ants belonging to the subfamilies Myrmicinae (Bouzid et al., 2013; Rifflet et al., 2012), Myrmeciinae (Davies et al., 2004; Inagaki et al., 2004, 2008a; Lewis et al., 1968; Mackintosh et al., 1998; Wiese et al., 2006; Wu et al., 1998), Pseudomyrmecinae (Touchard et al., 2014b) and Ectatomminae (Arseniev et al., 1994; Nolde et al., 1995; Pluzhnikov et al., 1999).

Peptides are the dominant compounds in most animal venoms and they represent a huge source of structurally diverse and biologically active toxins with high potency and selectivity for a range of targets (King and Hardy, 2013). Despite the clear potential that ant venom peptides represent, their investigation and characterisation remains highly underexplored. To date, only 72 ant venom peptides, from 11 ant species, have been fully sequenced (Fig. 3). This is a very small number in comparison to snakes, cone snails, scorpions or spiders. For example, 922 spider peptide toxins have currently been sequenced from 86 spider species and are available in the ArachnoServer 2.0 (Herzig et al., 2011). Therefore, it has been estimated that more than 98% of arachnid venoms remain completely uncharacterized (Quintero-Hernández et al., 2011), and with ant venoms this figure would be closer to 99.9%.

Until recently, the main reason for the limited number of studies on ant venoms is the small size of ants, and hence the small yield of venom. However, advancements in analytical techniques, particularly in mass spectrometric technologies, has resulted in higher sensitivity and resolving power, allowing for a more extensive exploration of the ant venom peptidome. This review summarizes the current knowledge on the biochemical and pharmacological properties of all peptide toxins sequenced from ant venoms to date.

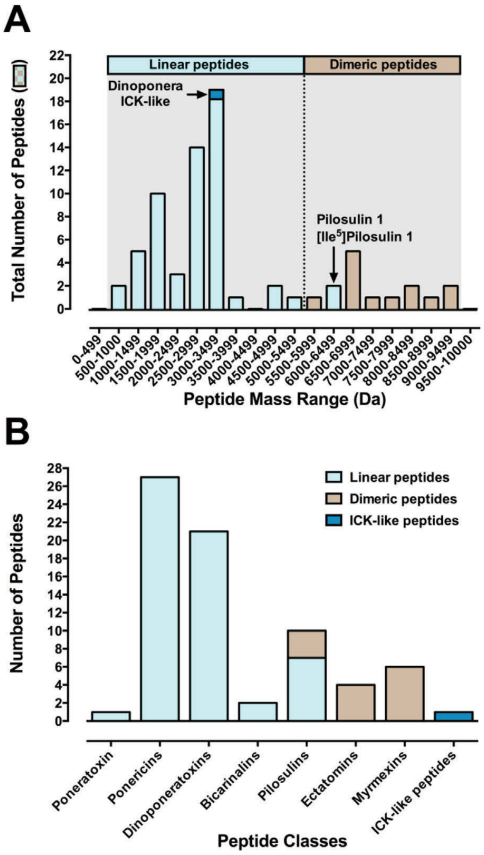


Fig. 3. (A) Bimodal mass distribution of the 72 characterized peptide toxins from ant venoms. Linear peptides range in mass from 761 to 5275 Da (except pilosulin 1, 6048 Da and [Ile^5]pilosulin 1, 6062 Da), while dimeric peptides range from 5603 to 9419 Da. (B) Ant peptide toxin classes. In both panels: cyan, linear peptides; beige, dimeric peptides; teal, ICK-like peptides. For clarification of colours in this figure, refer to the web version of this article.

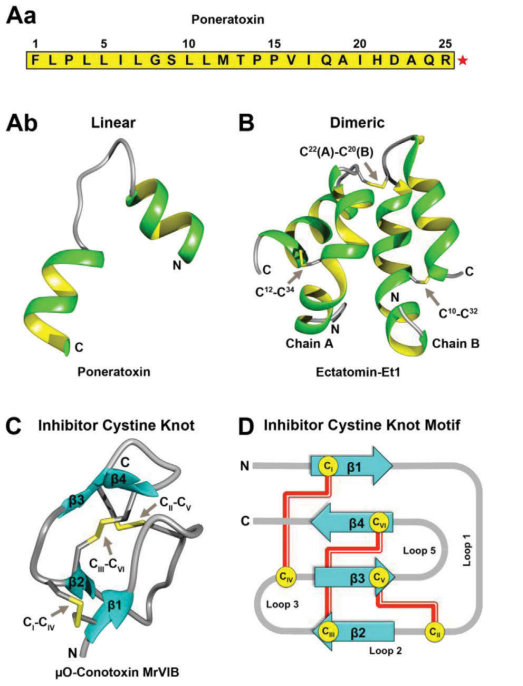


Fig. 4. Structures of ant peptide toxins. (Aa) Poneratoxin is a 2754.60 Da linear peptide (UniProtKB Accession POTX_PARCV) with no sequence homology to other peptides. The red star represents C-terminal amidation. (Ab) NMR structure of poneratoxin (PDB Accession 1G92) shows it comprises of two α-helices. (B) NMR structure of ectatomin (PDB Accession 1ECI), a heterodimeric peptide that forms a four-α-helical bundle structure. The intra- and interchain disulfide bonds are labelled for clarity. (C) Homology model of *Dinoponera* ICK-like peptide modelled on μO-conotoxin MrVIB from the venom of the cone snail *Conus marmoreus* (PDB Accession 1RMK; UniProtKB Accession CO16B_CONMR). In all panels, the peptide backbone is shown as a grey tube; β-sheets are represented by cyan arrows, α-helices are depicted as green/yellow spirals and disulfide bonds are shown as yellow tubes. The N-terminus (N) and C-terminus (C) of each peptide are also labelled. (D) Schematic representation of an ICK-like peptide. The pseudo-knot is formed when one disulfide bridge (C_{III}–C_{VI}) passes through a ring formed by two other disulfides (C_I–C_{IV} and C_{II}–C_V) and the intervening backbone. For clarification of colours in this figure, refer to the web version of this article.

For the purposes of this review, these peptides have been classified based on their structure and classified into three main groups; (i) linear, (ii) dimeric and (iii) inhibitor cystine knot (ICK)-like peptides.

3.1. Linear peptides

Most of the proteomic studies on ant venoms have so far revealed that the majority of the proteinaceous component of ant venoms are small, polycationic linear peptides with masses below 5 kDa (Cologna et al., 2013; Johnson et al., 2010; Orivel et al., 2001; Rifflet et al., 2012). This is consistent with studies performed on other hymenopteran (wasp and bee) venoms (Argiolas and Pisano, 1985; Baptista-Saidemberg et al., 2011; de Souza et al., 2004; Dias et al., 2014; Favreau et al., 2006; Gomes et al., 2014; Mendes et al., 2004; Qiu et al., 2012). Many of these linear peptides have antimicrobial properties and some possess additional insecticidal activity. Examples include ponericins from the neotropical ant *Pachycondyla goeldii* (Orivel et al., 2001) now reclassified as *Neoponera goeldii* (Schmidt and Shattuck, 2014), certain dinoponeratoxins (from *Dinoponera australis*) (Cologna et al., 2013) and pilosulins from the Australian jack jumper ant *Myrmecia pilosula*, which have been shown to have antimicrobial activity (Inagaki et al., 2004; Zelezetsky et al., 2005). These antimicrobial peptides demonstrate broad spectrum antibacterial activity and include α-helix antimicrobial peptides, and peptides with homology to the antimicrobial mucroporins, cecropins, brevinins, gaegurins, temporins and demaseptins (Cologna et al., 2013; Davies et al., 2004; Inagaki et al., 2004; Johnson et al., 2010; Kuhn-Nentwig, 2003; Mackintosh et al., 1995, 1998; Orivel et al., 2001; Viljakainen and Pamilo, 2008; von Sicard et al., 1989; Zelezetsky et al., 2005).

3.1.1. Poneratoxin

In 1991, the first ant venom peptide toxin, poneratoxin, was isolated and sequenced (Piek et al., 1991b). Poneratoxin is a 25-residue peptide neurotoxin derived from the bullet ant *Paraponera clavata* (subfamily Paraponerinae) with no apparent homology to other known peptides (Fig. 4Aa) (Piek et al., 1991a). The 3D NMR structure of poneratoxin has also been determined (Szolajska et al., 2004) and revealed a 'V'-shaped peptide with two α-helices connected by a β-turn (Fig. 4Ab). It has been shown to modulate voltage-gated sodium (Na_V) channels of both vertebrates and invertebrates and blocks synaptic transmission in the insect CNS. Poneratoxin induces long-lasting plateau action potentials and repetitive firing due to the presence of a slowly developing inward sodium current that activates at hyperpolarising potentials. This results from a potential toxin-induced interconversion between a fast and a slow conducting state of the Na_V channel (Duval et al., 1992; Hendrich et al., 2001; Szolajska et al., 2004).

3.1.2. Ponericins

Ponericins are a group of 27 peptides characterised from the venom of the ponerine ants, *Pachycondyla goeldii* (Orivel et al., 2001), *Pachycondyla apicalis*, *Pachycondyla inversa* (Orivel, 2000) and *Pachycondyla commutata* (Touchard and Aili, unpublished data) (now all renamed *Neoponera* spp.; Schmidt and Shattuck, 2014). Ponericins possess amphipathic α-helical structures in polar environments, and have been shown to exhibit haemolysis,

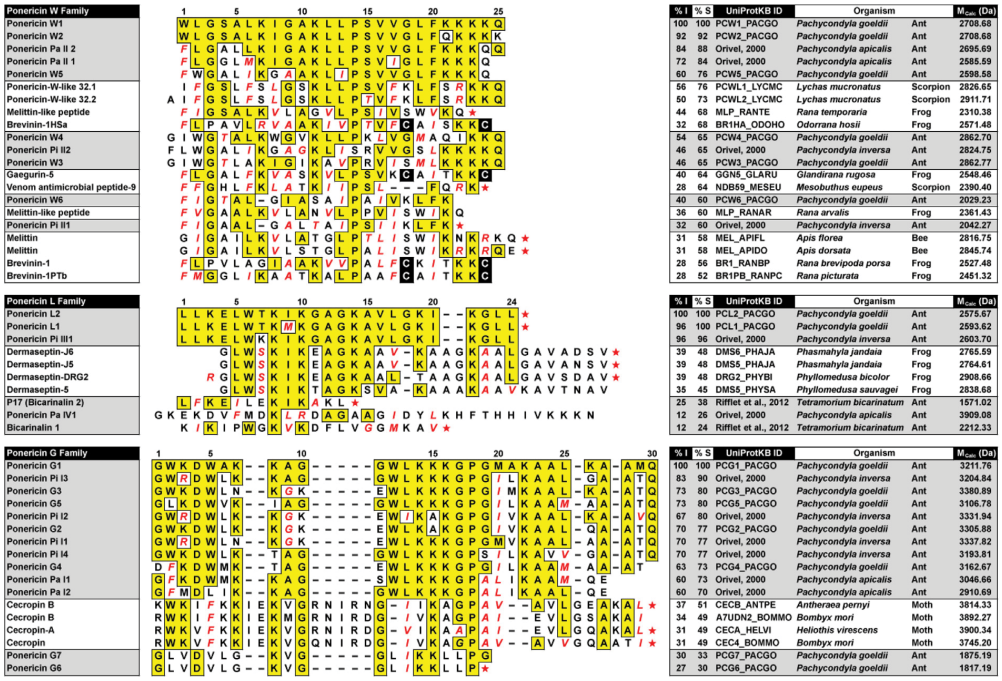


Fig. 5. Sequence alignment of bicarinalin and three ponericin families of linear peptides. Toxin names boxed in light grey are derived from ants. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are highlighted in black while red stars represent C-terminal amidation. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to the first peptide of each family, while percentage similarity (%S) includes conservatively substituted residues. M_{calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. Note added in proof: *Pachycondyla* spp. listed in this figure have all recently been renamed *Neoponera* spp. (Schmidt and Shattuck, 2014). For clarification of colours in this figure, refer to the web version of this article.

antibacterial activity against both Gram-positive and Gram-negative bacteria, as well as insecticidal activity (Orivel et al., 2001). Ponericins have been classified into three different families ('G', 'W' and 'L') based on sequence homology (Fig. 5). The ponericins show considerable sequence homology with other previously characterised peptides. For example, ponericin G peptides show homology to cecropin-like peptides from moths, flies, beetles and butterflies (Lee et al., 2013); ponericin W peptides have homology with the cytolytic peptide bee peptide melittin and gaegurins from frogs (Lee et al., 2011; Palma, 2013); and ponericin L peptides with dermaseptins isolated from the skin of *Phasmahyla* and *Phyllomedusa* frogs (Amiche and Galanth, 2011; Nicolas and Amiche, 2013). Given the known actions of these other peptides as cytolytic agents, ponericins may also form amphipathic α-helical structures in cell membranes, although only ponericin W peptides appear to have additional haemolytic actions. This function may be important in preventing the spread of microbial pathogens in ant colonies following ingestion of contaminated prey or their introduction into their colony following paralysis and subsequent transport of the prey into the colony (Lai et al., 2012).

3.1.3. Dinoponeratoxins

The giant Neotropical hunting ant *Dinoponera australis* (subfamily Ponerinae) is a solitary foraging, predatory ant whose venom paralyses invertebrates and causes a range of systemic effects in vertebrates (Haddad Junior et al., 2005). Envenomation in humans is rare, although stings have been reported to produce rapid and excruciating pain, diaphoresis, nausea, vomiting, tachycardia and lymphadenopathy (Haddad Junior et al., 2005). Liquid chromatography–mass spectrometry (LC-MS) analysis of the venom identified over 75 proteinaceous components with numerous small mass peptides (429–3214 Da) and a wide range of hydrophobicity and abundance. The six most abundant peptides were sequenced by tandem MS and Edman degradation and named dinoponeratoxins ('Da' toxins) (Johnson et al., 2010). Subsequently similar dinoponeratoxin peptides have been isolated and sequenced from the related ant *Dinoponera quadriceps* - henceforth known as 'Dq' toxins (Cologna et al., 2013). All 21 Dq and Da dinoponeratoxins show various degrees of homology with existing linear peptides, and can be separated into six groups (Fig. 6).

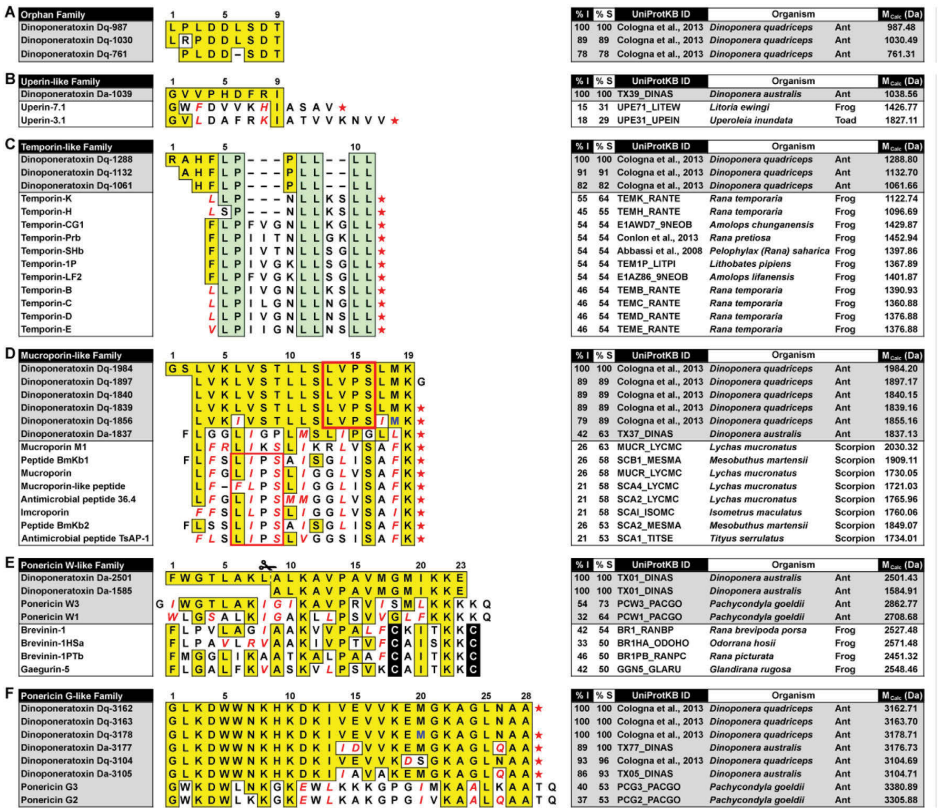


Fig. 6. Alignment of the dinoponeratoxin families of linear peptides. Toxin names boxed in light grey are derived from ants. Identical residues are boxed in yellow while conservative substitutions are shown in red italic text. Numbering is according to the first peptide in each family. Gaps were introduced to optimize the alignments. Red stars indicate an amidated C-terminus and cysteines are highlighted in black. Percentage identity (%I) is relative to the first peptide in each alignment while percentage similarity (%S) includes conservatively substituted residues. Apart from uperin peptides in panel B, only homologies greater than 50% are displayed. M_{Calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. (C) Residues Leu^{5}, Pro^{6}, and Leu$^{8-11}$ (numbering from Dq-1288), a common motif within temporin peptide families, are highlighted in green. (D) The residues K(V/L/I)IPS within the red boxes are thought to be critical for function in the scorpion antimicrobial peptide pandinin 2 and the scorpion peptides in panel D (Harrison et al., 2014). (E) Dinaponeratoxin Da-2501 is cleaved at the position marked to generate Da-1585. The blue Met^{18} in Dinoponeratoxin Dq-1837 (panel D) and Met^{20} in Dinoponeratoxin Dq-3178 (panel F) indicate residues that are probably oxidized. Note: *Pachycondyla goeldii* has recently been renamed *Neoponera goeldii* (Schmidt and Shattuck, 2014). For clarification of colours in this figure legend, refer to the web version of this article.

Group I are short 7–9 residue Dq peptides forming a three-member orphan peptide family with no homology to existing peptides (Fig. 6A), and no known biological activity (Cologna et al., 2013). Group II has only one member, Da-1039 (Fig. 6B), with only very limited homology to the uperin family of antibacterial frog skin secretions (Bradford et al., 1996; Steinborner et al., 1997). Group III comprise three 9–11 residue Dq toxins with moderate homology with the temporin family of antibacterial frog skin secretions (Abbassi et al., 2008; Rinaldi and Conlon, 2013; Simmaco et al., 1996). Temporins are one of the largest groups of antimicrobial peptides within the cationic host defence peptide family. They were originally isolated from skin secretions of the frog *Rana temporaria*, and are amphipathic α-helical peptides of 8–19 residues with a low net positive charge (0 to +3) and C-terminal amidation (Mangoni et al., 2007; Suzuki et al., 2007). The reasonably high homology of the temporin-like Dq toxins would suggest potentially similar biological activity, especially given the conservation of the common Pro and Leu residues found in temporin peptide families (Simmaco et al., 1996; residues in green boxes in Fig 6C). This leucine-rich tail has previously been shown to be important for membrane interaction (Avitabile et al., 2013). The antimicrobial activity of temporins is associated with an alteration of the cytoplasmic membrane permeability, without destruction of cell integrity (Mangoni et al., 2004). Temporins are particularly active against Gram-positive bacteria but most do not affect eukaryotic cells. However, they may act in a more complex way to inhibit various metabolic functions of the cell (Epand and Vogel, 1999; Park et al., 1998).

Group IV is the largest group of dinoponeratoxins and have masses between 1837 and 1984 Da with 17–19 residues. These have significant homology (53–63% similarity) with the antibacterial cationic host defence peptides BmKb (caerin-like) and mucroporin originally isolated from the venom of the scorpions *Mesobuthus martensii* and *Lychas mucronatus*, respectively (Dai et al., 2008; Zeng et al., 2004). These antimicrobial peptides are now found in a range of scorpion species and are being investigated as novel anti-infective drugs or lead compounds, for treating antibiotic-resistant microbial infections (Harrison et al., 2014).

Group V is a recently discovered collection of 15 dinoponeratoxins from the venom of *Dinoponera quadriceps* sequenced from a total of 354 peptides found in this venom (Cologna et al., 2013). These were found to share homology with the ponericin W family, dinoponeratoxins (from *D. australis*) and poneratoxin. These peptides also revealed both antimicrobial and antifungal activities (Cologna et al., 2013).

Group VI comprises of the ant venom peptides Da-3105 and Da-3177, from the giant Neotropical hunting ant *D. australis*, which show considerable homology to ponericin G2 and may possess similar bioactivity.

3.1.4. Bicarinalins

Two novel peptides, bicarinalin 1 and P17 (bicarinalin 2) have been isolated and characterised from the venom of the ant *Tetramorium bicarinatum* (Myrmicinae) from a total of 31 peptides identified in this venom (Rifflet et al., 2012). Interestingly, these peptides show very low homology with known peptide toxins (Fig. 5). Bicarinalin 1 exhibits all the characteristics of an amphipathic helical peptide and has broad and potent antibacterial activity similar to melittin, pilosulin and defensin but with weaker haemolytic activity (Rifflet et al., 2012; Téné et al., 2014). Accordingly, it is being investigated as an anti-infective agent for use against emerging antibiotic-resistant pathogens. Recently the venom gland transcriptome of *Tetramorium bicarinatum*, one of the world's most broadly distributed ant species, has also been published (Bouzid et al., 2013). Transcribed *T. bicarinatum* venom gland ESTs revealed allergenic/cytotoxic peptides, with homology to pilosulins 1, 3 and 5, and paralytic peptide toxins, one of which possesses homology with the insect cytokine precursor uENF2. These allergenic/cytotoxic and paralytic toxins contributed close to 70% of the total EST cDNAs.

3.2. Dimeric peptides

Dimeric peptides are peptides with two subunits that are linked covalently with a disulfide bond (Sarray et al., 2013) and peptide dimerization is currently being investigated as a potential way to increase the activity of certain peptide toxins (Vizzavona et al., 2009). Except for snake venoms (Osipov et al., 2008), a dimeric scaffold in peptides is quite rare in venomous animals, although it has occasionally been reported in the venoms of some scorpions (Zamudio et al., 1997), spiders (Santos et al., 1992) and marine cone snails (Loughnan et al., 2006). In the case of ant venoms, dimeric peptides seem to be common in the subfamilies Ectatomminae, Myrmeciinae and Pseudomyrmecinae (see below), but have not yet been described in other subfamilies. The amino acid sequences and

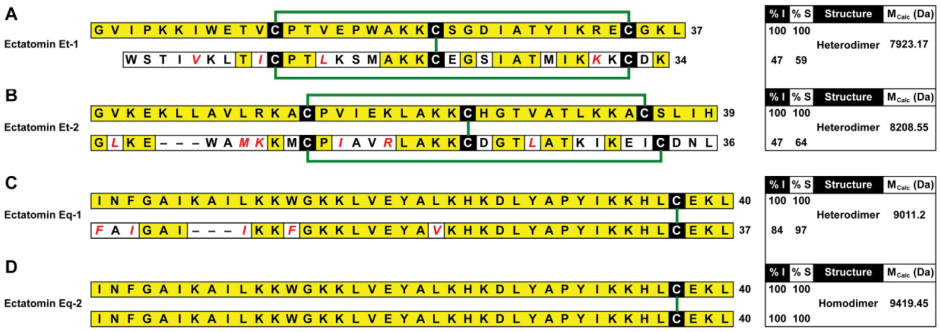


Fig. 7. Sequences and structures of the ectatomin family of dimeric ant peptides. Identical residues are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are highlighted in black and the predicted disulfide-bonding pattern is shown in green between the sequences. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to the longer (upper) chain for each peptide while percentage similarity (%S) includes conservatively substituted residues. M_{Calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. The heterodimeric ectatomin Et peptides (A–B) are from *Ectatomma tuberculatum* while ectatomin Eq peptides (C–D) are from *Ectatomma brunneum* (formerly *E. quadridens*). (A) Sequences for ectatomin Et-1 toxins are from UniProtKB Accessions ECAA_ECTTU and ECAB_ECTTU. (B–D) Remaining sequences are from Pluzhnikov et al. (2000). For clarification of colours in this figure, refer to the web version of this article.

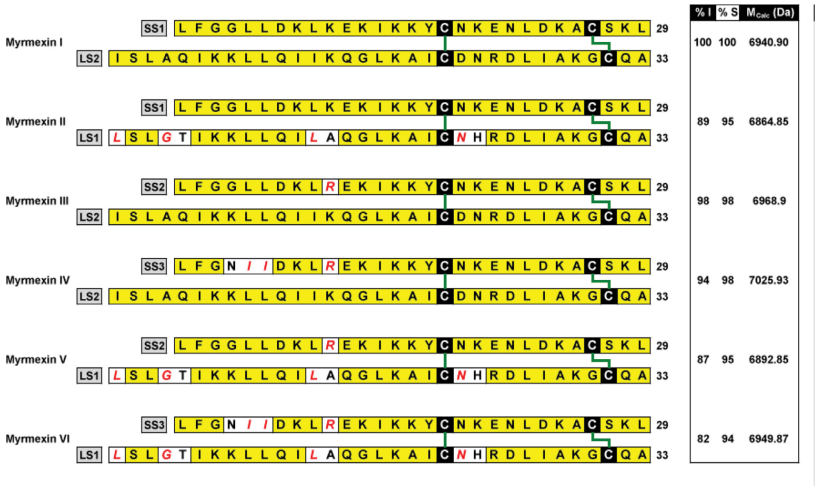


Fig. 8. Sequences and structures of the myrmexin family of heterodimeric peptides from the venom of the ant *Pseudomyrmex triplarinus*. Each myrmexin is composed of a short subunit (SS1, SS2 or SS3; grey boxes) and a long subunit (LS1 or LS2; grey boxes) linked by two disulfide bonds. Identical residues are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are highlighted in black and the predicted inter-chain disulfide-bonding pattern is shown in green between the sequences. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to myrmexin I while percentage similarity (%S) includes conservatively substituted residues. M_{Calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. Sequences are taken from Pan and Hink (2000). For clarification of colours in this figure, refer to the web version of this article.

disulfide connectivity of the known dimeric peptide-toxins are reported in Figs. 7–9.

3.2.1. *Ectatomins*

One of the most potent neurotoxic peptides isolated from ant venoms, is ectatomin (Et-1), from the venom of the ant *Ectatomma tuberculatum* (Ectatomminae) (Pluzhnikov et al., 1994) and its homologue, ectatomin Et-2 (Pluzhnikov et al., 2000). These peptides are highly basic heterodimeric complexes consisting of two highly homologous amphiphilic polypeptide chains linked together by one inter-chain disulfide bond (Arseniev et al., 1994). Each chain also possesses an intra-chain disulfide bond (Fig. 7A–B). Disulfide

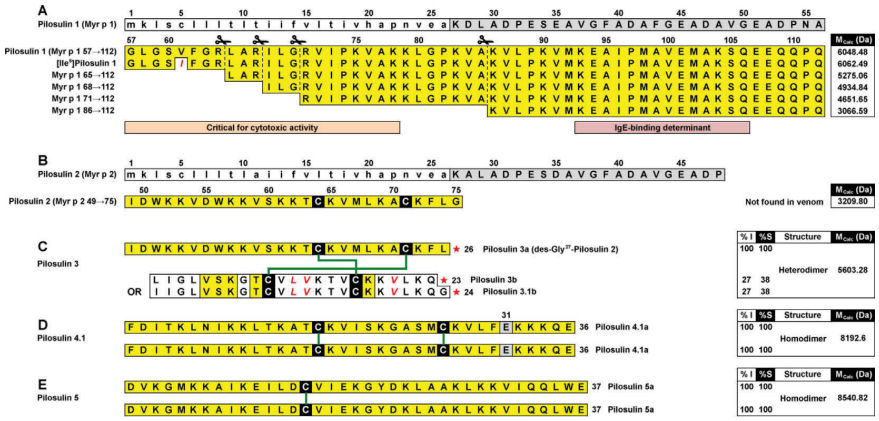


Fig. 9. Sequences and structures of the pilosulin family of linear and dimeric ant peptides. (A–B) Translated sequences of the linear peptides pilosulin 1 (A; from *Myrmecia pilosula*; UniProtKB accession MYR1_MYRPI) and pilosulin 2 (B; from *Myrmecia pilosula*; UniProtKB accession MYR3A_MYRPI). Sequences represent complete prepropeptides, where signal peptides are boxed in white and in lowercase, propeptide sequences are boxed in gray and the mature peptides are boxed in yellow. (A) In addition to the natural variant [Ile^5] pilosulin 1, pilosulin 1 undergoes cleavage at the sites marked above the mature peptide sequence to yield four additional peptides, while the residues important for cytotoxic activity and IgE binding are highlighted beneath the sequences. (B) Pilosulin 2 does not appear to be found in venom but undergoes post-translational modification to yield the monomer pilosulin 3a (des-Gly^{27}-Pilosulin 2) that forms the heterodimer pilosulin 3 (C) with the monomer pilosulin 3b (MYR3B_MYRPI). A natural variant, pilosulin 3.1b, can be found in the venom of *Myrmecia banksi* (MYR3_MYRBA). (D–E) Pilosulin 4.1 (Wiese et al., 2006) and pilosulin 5 (MYR5_MYRBA) are homodimers from *Myrmecia banksi*. In the case of pilosulin 4.1, cDNA cloning predicted a homodimer of pilosulin 4a (MYR4_MYRBA), but this was not detected in venom and the [Glu^{31}]pilosulin 4 variant (pilosulin 4.1) found in venom is shown. For the dimeric peptides (C–E), cysteines are highlighted in black and the predicted disulfide-bonding pattern is shown in green between the sequences. Red stars indicate an amidated C-terminus. M_{Calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. For clarification of colours in this figure, refer to the web version of this article.

bonds render venom peptides resistant to a number of different proteases and environmental extremes resulting in stable peptide toxins (King and Hardy, 2013). The three dimensional structure of Et-1 was determined by NMR and revealed that each ectatomin chain comprises two anti-parallel α-helices linked by a hinge region of four amino acid residues and a disulfide bridge (Fig. 4B) (Nolde et al., 1995). Two other ectatomins (Eq-1 and Eq-2) have also been isolated from the venom of *Ectatomma brunneum* (previously *Ectatomma quadridens*). These novel Eq ectatomins are also dimeric and linked by one inter-chain disulfide bond. However, they lack the intra-chain disulfide bond present in Et-1 and -2 (Pluzhnikov et al., 2000).

Et-1 appears to account for the major toxic effect of *Ectatomma tuberculatum* venom causing toxic effects in both mammals and insects (Pluzhnikov et al., 1999). At high concentrations (0.50–1 μM), Et-1 is a pore-forming peptide that inserts into cellular and artificial membranes but is not internalized. It produces haemolytic and cytolytic effects on rabbit erythrocytes, *Xenopus laevis* oocytes, rat cardiomyocytes and both insect and vertebrate cell lines. In *X. laevis* oocyte membranes, this arises due to the formation of nonselective cationic channels by two Et-1 molecules and appears to involve binding to lipids rather than a specific receptor. The increase in cell permeability, with resultant ion leakage, results in cell death (Pluzhnikov et al., 1994, 1999). At much lower concentrations (1–10 nM), Et-1 is capable of inhibiting whole-cell L-type calcium currents in isolated rat ventricular myocytes. Importantly, it prevents β-adrenoceptor- or adenylate cyclase-mediated activation of calcium currents suggesting that Et-1 interacts directly or allosterically with agonist-bound β-adrenoceptors preventing activation of calcium channels further down the signal transduction cascade. The modulation of calcium channels and possibly β-adrenoceptors by Et-1 may underlie its potent toxicity by interfering with the process of muscle contraction, neurotransmitter release and neuromodulation (Pluzhnikov et al., 1999).

3.2.2. Myrmexins

In vitro and clinical studies have shown that *Pseudomyrmex triplarinus* (Pseudomyrmecinae) ant venom decreases pain and inflammation in patients with rheumatoid arthritis and reduces swelling in animal models of inflammation (Altman et al., 1984; Hink and Butz, 1985; Schultz and Arnold, 1984). Myrmexins are a family of six related polypeptides (myrmexins I–VI) that have been purified from the venom of *Pseudomyrmex triplarinus*. These peptides are heterodimeric complexes comprising a combination of a short subunit of 29 residues (SS1, SS2 or SS3) and a long subunit of 33 residues (LS1 or LS2) stabilized by two inter-chain disulfide bonds (Pan and Hink, 2000) (Fig. 8). Unfortunately, it is not known at present which of the myrmexin peptides are associated with the anti-inflammatory activity observed with whole venom. Three additional myrmexin-like polypeptides from the venom of the related ant *Pseudomyrmex penetrator* (one heterodimeric and two homodimeric) have also been identified, however, they are yet to be sequenced (Touchard et al., 2014b). These myrmexins may represent a new class of toxins present in Pseudomyrmecine ants.

3.2.3. Pilosulins

Australian ants of the *Myrmecia pilosula* species complex (Myrmeciinae), also known as jack jumper ants, have a painful sting that is responsible for around 90% of life-threatening ant sting allergies in Australia (Brown et al., 2003; Douglas et al., 1998; Street et al., 1994). In South Eastern Australia around 2.7% of the population are allergic to *Myrmecia pilosula* venom, with approximately 50% of allergic people experiencing life-threatening reactions (Brown et al., 2003). The toxicity of the venom appears to result from the presence of a variety of histamine-like, haemolytic and eicosanoid-releasing factors, peptides such as pilosulins, and enzymes including phospholipases, hyaluronidase, and phosphatases (Matuszek et al., 1994a, 1992, 1994b; McGain and Winkel, 2002).

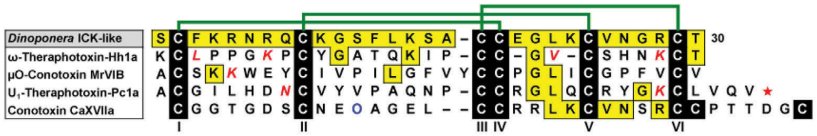
Using cDNA sequencing, two major protein allergens from *Myrmecia pilosula* sharing a common leader sequence have been identified (Donovan et al., 1993, 1995, 1994; Street et al., 1996). They encode the 112 and 75 amino acid prepropeptides Myr p 1 and Myr p 2, respectively (Fig. 9A–B). Pilosulin 1, the mature peptide product from residue 57 to 112 of Myr p 1 (Myr p 1 57 → 112), is a 6048 Da linear allergenic basic peptide that exhibits haemolytic and cytotoxic activity and is one of the major allergens that have been identified in this venom (Donovan et al., 1993, 1994; Wu et al., 1998). However, pilosulin 1 exists mainly, and sometimes exclusively, as a Val5Ile substituted isoform known as [Ile^5]pilosulin 1 (Davies et al., 2004) (Fig. 9A). Pilosulin 1 is also cleaved to form four additional N-terminally truncated isoforms with varying degrees of cytotoxic activity (Fig. 9A).

Pilosulin 2 (Myr p 2 49 → 75) has never been detected in whole venom in its monomeric form (Donovan and Baldo, 1997). However, a des-Gly^{27} pilosulin 2 peptide (renamed pilosulin 3a) has been found as part of the 5603 Da heterodimeric peptide pilosulin 3. The additional subunit of pilosulin 3 from *Myrmecia pilosula* is the 23 residue pilosulin 3b (Davies et al., 2004), or the variant pilosulin 3.1b from *Myrmecia banksi* (Inagaki et al., 2004), thought to be part of the *Myrmecia pilosula* species complex (Imai et al., 1994) (Fig. 9C). Pilosulin 3 displays antimicrobial activity, and is the major allergen in *M. pilosula* venom, along with [Ile5]pilosulin 1 accounting for 80% of the total venom peptide content. Pilosulin 4a peptide was originally identified via cDNA cloning (Inagaki et al., 2004) but was not detected in venom, while its Asp31Glu variant pilosulin 4.1a was found to be present only as a homodimeric peptide, pilosulin 4.1 (Wiese et al., 2006) (Fig. 9D). cDNA cloning also revealed the presence of a novel bioactive dimeric peptide pilosulin 5 connected by a single disulfide bond. Synthetic pilosulin 5 dimer causes significant histamine release that may be related to the weak homology of the peptide to the wasp peptide mastoparan (Inagaki et al., 2008a).

Although the monomeric pilosulin peptides (pilosulin 2, 3.2b, 4 and 5) all show antibacterial and histamine-releasing activities (Inagaki et al., 2004, 2008a) and some pilosulins, particularly 3a and to a lesser extent 4.1 and [Ile^5]pilosulin 1, are known to be highly allergenic (Wiese et al., 2007), the biological activities of these peptides have not been fully investigated.

3.3. ICK-like peptides

The inhibitor cystine knot (ICK) structural motif is an evolutionary conserved structure that has been found in plants, fungi, viruses, antimicrobial peptides from horseshoe crabs (tachystatins) and the venoms of many organisms such as spiders, scorpions, cone snails, insects (bees) and sea anemones (Barbault et al., 2003; Bloch and Cohen, 2014; Cammue et al., 1992; Gilly et al., 2011; Osaki et al., 1999; Pallaghy et al., 1994; Rodríguez et al., 2014; Zhu et al., 2003). The ICK motif is defined as an embedded ring formed by two disulfide bonds Cys(I–IV) and Cys(II–V) and their connecting backbone segments through which is threaded a third disulfide bond Cys(III–VI), forming a cystine knot. It is invariably associated with a nearby anti-parallel β-sheet and appears to be a highly effective motif for stabilizing peptide structures (Fig. 4D). Peptides with an ICK motif represent attractive scaffolds in drug design because of their inherent chemical stability and resistance to proteases provided by the fold and the wide range of amino acid sequences that can be accommodated in the structure (Craik et al., 2001; Norton and Pallaghy, 1998; Pallaghy et al., 1994; Zhu et al., 2003). While



	%I	%S	UniProtKB ID	Organism		M_{calc} (Da)
Dinoponera ICK-like	100	100	Torres et al. 2014	*Dinoponera quadriceps*	Ant	3319.51
ω-Theraphotoxin-Hh1a	33	47	TXH10_HAPSC	*Haplopelma huwenum*	Spider	2929.30
μO-Conotoxin MrVIB	32	35	CO16B_CONMR	*Conus marmoreus*	Cone snail	3403.58
U_1-Theraphotoxin-Pc1a	27	33	TXFK1_PSACA	*Psalmopoeus cambridgei*	Spider	3615.66
Conotoxin CaXVIIa	32	32	CXY17_CONCB	*Conus caracteristicus*	Cone snail	3444.30

Fig. 10. Structure and sequence alignment of the *Dinoponera* ICK-like peptide. The upper panel shows the disulfide bonding connectivity and alignment with homologous peptides. Cysteines are highlighted in black and the predicted disulfide-bonding pattern, similar to other ICK peptides, is shown in green above the sequences. Identical residues are boxed in yellow while conservative substitutions are shown in red italic text. Gaps were introduced to optimize the alignments. Red stars indicate an amidated C-terminus while the blue O[11] in conotoxin CaXVIIa indicates a hydroxyproline residue. The lower panel shows the percentage identity (%I) relative to *Dinoponera* ICK-like peptide while percentage similarity (%S) includes conservatively substituted residues. M_{calc}, Theoretical monoisotopic mass calculated using GPMAW 9.20 software. For clarification of colours in this figure, refer to the web version of this article.

large numbers of ICK peptide toxins have been reported from other arthropod venoms such as spiders and scorpions, only two types of ant venom peptides displaying this structural motif are currently known – *Dinoponera* ICK-like peptide and SKTXs.

3.3.1. *Dinoponera* ICK-like toxin

The recent transcriptome analysis of the venom glands of the ant *Dinoponera quadriceps* (Ponerinae) has confirmed the presence and sequence of the first ICK-like peptide in ant venoms (Torres et al., 2014). This *Dinoponera* ICK-like peptide is a minor component of the venom of *Dinoponera quadriceps* but has a VI/VII cysteine framework (–C–C–CC–C–C–) consistent with other ICK toxins (Fig. 4C). This peptide shows limited homology to the ICK toxins ω-theraphotoxin-Hh1a and μO-conotoxin MrVIB peptides found in tarantula and cone snail venoms, respectively, both of which exhibit neurotoxic activity via activity on voltage-gated ion channels (Liu et al., 2006; McIntosh et al., 1995) (Fig. 10).

3.3.2. SKTXs

The venom of the ant *Strumigenys kumadori* (Myrmicinae) also possesses ICK-like peptides which have been named SKTXs (Inagaki et al., 2008b). SKTXs are thought to modulate Na_V channels of *Drosophila*, however, this study remains unpublished and sequences of these peptides are still unknown.

4. Conclusion and perspectives

Until recently, the low yield of venom from ant species has severely restricted the biochemical and pharmacological characterisation of ant venom peptides. However, advances in the development of miniaturized bioassays and improvements in the sensitivity of mass spectrometry and NMR spectroscopy now allow broader investigations of the small quantities of venom peptides provided by small animals, especially ants. Indeed, mass spectrometry has been used as a method to improve the accuracy of taxonomic findings to reveal cryptic ant species within species complexes (Touchard et al., 2014a). This chemotaxonomic tool can therefore contribute to more rapid species identification and more accurate taxonomies.

The limited number of studies to date has revealed a number of unique structures across a broad range of ant subfamilies that differ from those described in other animal venoms. Given the diversity in ant species and distribution, ant venoms therefore represent vast sources of potentially novel bioactive toxins that could be exploited in drug and bioinsecticide discovery programs. For example, there is increasing awareness that peptides represent an under-utilized source of lead compounds for new therapeutics. Arguably, the largest source of chemical diversity comes from peptides derived from animal venoms. In animal venoms the evolutionary pressure for improved prey capture and/or defence has resulted in complex preoptimised combinatorial peptide libraries with extremely diverse pharmacologies that interact with a wide range of molecular targets. The discovery that these peptides bind to their cognate receptors and ion channels with high affinity and selectivity means that many are now being investigated as sources of lead compounds in therapeutic discovery pipelines (Bosmans et al., 2009; Escoubas and King, 2009; Lewis and Garcia, 2003; Vetter et al., 2011; Vetter and Lewis, 2012). Hence, there is a growing number of novel peptide or peptidomimetic therapeutics appearing on the drug market, or in clinical trials, which are derived from toxins from the venoms of cone snails, snakes, Gila monster, scorpions, spiders and sea anemones. Ants could also provide a unique source of potential therapeutic leads, especially antimicrobials and neuroactive compounds.

Since some venomous animals, particularly arachnids and ants, prey upon insects their venom contains large numbers of insecticidal peptide toxins that have evolved to kill or paralyse insect prey. These toxins often modulate the function of their targets with high insect selectivity, lacking any overt toxicity against their vertebrate counterparts (Bende et al., 2013; Gurevitz et al., 2007; Karbat et al., 2004; Wang et al., 2000, 2001; Windley et al., 2011), which can even extend to unique insect family selectivity (Bende et al., 2014). Hence, many of these toxins are being explored as novel insecticides in biopesticide discovery programs (King and Hardy, 2013; Smith et al., 2013; Windley et al., 2012). The limited number of studies on ant venoms would indicate that potential insect-selective peptide neurotoxins are present in their venoms and could be exploited as novel insecticides leads.

Acknowledgements

Financial support for this study was provided by an Australian Postgraduate Award to Samira R. Aili, a *Programme Convergence 2007-2013, Région Guyane* from the European Community (BI-Appli, 115/SGAR-DE/2011/052274) to Alain Dejean, and a BIOPEPMED grant from the *Programme Amazonie II* of the French *Centre National de la Recherche Scientifique* to Pierre Escoubas. This work has also benefited from an "*Investissement d'Avenir*" grant managed by the *Agence Nationale de la Recherche* (CEBA, ANR-10-LABX-25-01).

Conflict of interest

The authors declare that there are no conflicts of interest.

Transparency document

Transparency document related to this article can be found online at http://dx.doi.org/10.1016/j.toxicon.2014.10.021.

References

Abbassi, F., Oury, B., Blasco, T., Sereno, D., Bolbach, G., Nicolas, P., Hani, K., Amiche, M., Ladram, A., 2008. Isolation, characterization and molecular cloning of new temporins from the skin of the North African ranid *Pelophylax saharica*. Peptides 29, 1526–1533.

Agosti, D., Johnson, N.F., 2005. Antbase. World Wide Web Electronic Publication antbase.org, version (05/2005). Available online: http://osuc.biosci.ohio-state.edu/hymenoptera/tsa.sppcount?the_taxon=Formicidae (accessed 24.08.14.).

Altman, R.D., Schultz, D.R., Collins-Yudiskas, B., Aldrich, J., Arnold, P.I., Brown, H.E., 1984. The effects of a partially purified fraction of an ant venom in rheumatoid arthritis. Arthritis Rheum. 27, 277–284.

Amiche, M., Galanth, C., 2011. Dermaseptins as models for the elucidation of membrane-acting helical amphipathic antimicrobial peptides. Curr. Pharm. Biotechnol. 12, 1184–1193.

AntWeb, 2014. Available online: http://www.antweb.org (accessed 27.08.14.).

Argiolas, A., Pisano, J.J., 1985. Bombolitins, a new class of mast cell degranulating peptides from the venom of the bumblebee *Megabombus pennsylvanicus*. J. Biol. Chem. 260, 1437–1444.

Arseniev, A., Pluzhnikov, K., Nolde, D., Sobol, A., Torgov, M.Y., Sukhanov, S., Grishin, E., 1994. Toxic principle of selva ant venom is a pore-forming protein transformer. FEBS Lett. 347, 112–116.

Avitabile, C., Netti, F., Orefice, G., Palmieri, M., Nocerino, N., Malgieri, G., D'Andrea, L.D., Capparelli, R., Fattorusso, R., Romanelli, A., 2013. Design, structural and functional characterization of a temporin-1b analog active against gram-negative bacteria. Biochim. Biophys. Acta 1830, 3767–3775.

Baptista-Saidemberg, N.B., Saidemberg, D.M., Palma, M.S., 2011. Profiling the peptidome of the venom from the social wasp *Agelaia pallipes pallipes*. J. Proteomics 74, 2123–2137.

Barbault, F., Landon, C., Guenneugues, M., Meyer, J.-P., Schott, V., Dimarcq, J.-L., Vovelle, F., 2003. Solution structure of alo-3: a new knottin-type antifungal peptide from the insect *Acrocinus longimanus*. Biochemistry 42, 14434–14442.

Bende, N.S., Dziemborowicz, S., Mobli, M., Herzig, V., Gilchrist, J., Wagner, J., Nicholson, G.M., King, G.F., Bosmans, F., 2014. A distinct sodium channel voltage-sensor locus determines insect selectivity of the spider toxin Dc1a. Nat. Commun. 5, 4350.

Bende, N.S., Kang, E., Herzig, V., Bosmans, F., Nicholson, G.M., Mobli, M., King, G.F., 2013. The insecticidal neurotoxin Aps III is an atypical knottin peptide that potently blocks insect voltage-gated sodium channels. Biochem. Pharmacol. 85, 1542–1554.

Bloch, G., Cohen, M., 2014. The expression and phylogenetics of the inhibitor cysteine knot peptide OCLP1 in the honey bee *Apis mellifera*. J. Insect Physiol. 65, 1–8.

Bosmans, F., Escoubas, P., Nicholson, G.M., 2009. Spider venom peptides as leads for drug and insecticide design. In: de Lima, M.E., Pimenta, A.M.C., Martin-Eauclaire, M.-F., Zingali, R.B., Rochat, H. (Eds.), Animal Toxins: State of the Art. Perspectives in Health and Biotechnology. Federal University of Minas Gerais Press, Belo Horizonte, pp. 269–290.

Bouzid, W., Klopp, C., Verdenaud, M., Ducancel, F., Vetillard, A., 2013. Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (hymenoptera: formicidae): the first transcriptome analysis of an ant species. Toxicon 70, 70–81.

Bradford, A.M., Raftery, M.J., Bowie, J.H., Tyler, M.J., Wallace, J.C., Adams, G.W., Severini, C., 1996. Novel uperin peptides from the dorsal glands of the Australian floodplain toadlet *Uperoleia inundata*. Aust. J. Chem. 49, 475–484.

Brady, S.G., Fisher, B.L., Schultz, T.R., Ward, P.S., 2014. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol. Biol. 14, 93–106.

Brady, S.G., Schultz, T.R., Fisher, B.L., Ward, P.S., 2006. Evaluating alternative hypotheses for the early evolution and diversification of ants. Proc. Natl. Acad. Sci. U. S. A. 103, 18172–18177.

Brand, J.M., 1978. Fire ant venom alkaloids: their contribution to chemosystematics and biochemical evolution. Biochem. Syst. Ecol. 6, 337–340.

Brown, S.G.A., Franks, R.W., Baldo, B.A., Heddle, R.J., 2003. Prevalence, severity, and natural history of jack jumper ant venom allergy in Tasmania. J. Allergy Clin. Immunol. 111, 187–192.

Cammue, B.P., De Bolle, M.F., Terras, F.R., Proost, P., Van Damme, J., Rees, S.B., Vanderleyden, J., Broekaert, W.F., 1992. Isolation and characterization of a novel class of plant antimicrobial peptides form *Mirabilis jalapa* L. seeds. J. Biol. Chem. 267, 2228–2233.

Casewell, N.R., Wuster, W., Vonk, F.J., Harrison, R.A., Fry, B.G., 2013. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol. Evol. 28, 219–229.

Cerdá, X., Dejean, A., 2011. Predation by ants on arthropods and other animals. In: Polidori, C. (Ed.), Predation in the Hymenoptera: an Evolutionary Perspective. Transworld Research Network, Kerala - India, pp. 39–78.

Cologna, C.T., Cardoso J. dos, S., Jourdan, E., Degueldre, M., Upert, G., Gilles, N., Uetanabaro, A.P., Costa Neto, E.M., Thonart, P., de Pauw, E., Quinton, L., 2013. Peptidomic comparison and characterization of the major components of the venom of the giant ant *Dinoponera quadriceps* collected in four different areas of Brazil. J. Proteomics 94, 413–422.

Craik, D.J., Daly, N.L., Waine, C., 2001. The cystine knot motif in toxins and implications for drug design. Toxicon 39, 43–60.

Dai, C., Ma, Y., Zhao, Z., Zhao, R., Wang, Q., Wu, Y., Cao, Z., Li, W., 2008. Mucroporin, the first cationic host defense peptide from the venom of *Lychas mucronatus*. Antimicrob. Agents Chemother. 52, 3967–3972.

Davies, N.W., Wiese, M.D., Brown, S.G., 2004. Characterisation of major peptides in 'jack jumper' ant venom by mass spectrometry. Toxicon 43, 173–183.

de Souza, B.M., Marques, M.R., Tomazela, D.M., Eberlin, M.N., Mendes, M.A., Palma, M.S., 2004. Mass spectrometric characterization of two novel inflammatory peptides from the venom of the social wasp *Polybia paulista*. Rapid Commun. Mass Spectrom. 18, 1095–1102.

Dejean, A., Labrière, N., Touchard, A., Petitclerc, F., Roux, O., 2014. Nesting habits shape feeding preferences and predatory behavior in an ant genus. Naturwissenschaften 101, 323–330.

Dias, N.B., de Souza, B.M., Gomes, P.C., Palma, M.S., 2014. Peptide diversity in the venom of the social wasp *Polybia paulista* (Hymenoptera): a comparison of the intra- and inter-colony compositions. Peptides 51, 122–130.

Donovan, G.R., Baldo, B.A., 1997. In: Pilosulin 2 from Ant Venom, Cloning and Expression of a CDNA Encoding it and its Antihypertensive Properties. PCT International Application, p. 27. WO 9713854.

Donovan, G.R., Baldo, B.A., Sutherland, S., 1993. Molecular cloning and characterization of a major allergen (Myr p 1) from the venom of the Australian jumper ant, *Myrmecia pilosula*. Biochim. Biophys. Acta 1171, 272–280.

Donovan, G.R., Street, M.D., Baldo, B.A., 1995. Separation of jumper ant (*Myrmecia pilosula*) venom allergens: a novel group of highly basic proteins. Electrophoresis 16, 804–810.

Donovan, G.R., Street, M.D., Baldo, B.A., Alewood, D., Alewood, P., Sutherland, S., 1994. Identification of an IgE-binding determinant of the major allergen *Myr p 1* from the venom of the Australian jumper ant *Myrmecia pilosula*. Biochim. Biophys. Acta 1204, 48–52.

Douglas, R.G., Weiner, J.M., Abramson, M.J., O'Hehir, R.E., 1998. Prevalence of severe ant-venom allergy in southeastern Australia. J. Allergy Clin. Immunol. 101, 129–131.

Duval, A., Malécot, C.O., Pelhate, M., Piek, T., 1992. Poneratoxin, a new toxin from an ant venom, reveals an interconversion between two gating modes of the Na channels in frog skeletal muscle fibres. Pflügers Arch. 420, 239–247.

Epand, R.M., Vogel, H.J., 1999. Diversity of antimicrobial peptides and their mechanisms of action. Biochim. Biophys. Acta 1462, 11–28.

Escoubas, P., King, G.F., 2009. Venomics as a drug discovery platform. Expert Rev. Proteomics 6, 221–224.

Fautin, D.G., 2014. Hexacorallians of the World. Available online: http://geoportal.kgs.ku.edu/hexacoral/anemone2/index.cfm (accessed 24.08.14.).

Favreau, P., Menin, L., Michalet, S., Perret, F., Cheneval, O., Stocklin, M., Bulet, P., Stocklin, R., 2006. Mass spectrometry strategies for venom mapping and peptide sequencing from crude venoms: case applications with single arthropod specimen. Toxicon 47, 676–687.

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.

Gomes, P.C., de Souza, B.M., Dias, N.B., Brigatte, P., Mourelle, D., Arcuri, H.A., dos Santos Cabrera, M.P., Stabeli, R.G., Neto, J.R., Palma, M.S., 2014. Structure-function relationships of the peptide paulistine: a novel toxin from the venom of the social wasp *Polybia paulista*. Biochim. Biophys. Acta 1840, 170–183.

Gurevitz, M., Karbat, I., Cohen, L., Ilan, N., Kahn, R., Turkov, M., Stankiewicz, M., Stühmer, W., Dong, K., Gordon, D., 2007. The insecticidal potential of scorpion β-toxins. Toxicon 49, 473–489.

Haddad Junior, V., Cardoso, J.L., Moraes, R.H., 2005. Description of an injury in a human caused by a false tocandira (*Dinoponera gigantea*, Perty, 1833) with a revision on folkloric, pharmacological and clinical aspects of the giant ants of the genera *Paraponera* and *Dinoponera* (sub-family Ponerinae). Rev. Inst. Med. Trop. Sao Paulo 47, 235–238.

Hallan, J., 2005. Synopsis of the Described Scorpiones of the World. Available online: http://insects.tamu.edu/research/collection/hallan/acari/Scorpiones1.htm (accessed 24.08.14.).

Harrison, P.L., Abdel-Rahman, M.A., Miller, K., Strong, P.N., 2014. Antimicrobial peptides from scorpion venoms. Toxicon 88C, 115–137.

Hendrich, A.B., Mozrzymas, J.W., Konopińska, D., Scuka, M., 2001. The effect of poneratoxin on neuromuscular transmission in the rat diaphragm. Cell. Mol. Biol. Lett. 7, 195–202.

Herzig, V., Wood, D.L., Newell, F., Chaumeil, P.A., Kaas, Q., Binford, G.J., Nicholson, G.M., Gorse, D., King, G.F., 2011. ArachnoServer 2.0, an updated online resource for spider toxin sequences and structures. Nucleic Acids Res. 39, D653–D657.

Hink, W.F., Butz, J.J., 1985. Primary culture of ant venom gland cells. In Vitro Cell Dev Biol 21, 333–339.

Hölldobler, B., Wilson, E.O., 1990. The Ants. Belknap Press (Harvard University Press), Cambridge, MA.

Imai, H.T., Taylor, R.W., Crozier, R.H., 1994. Experimental bases for the minimum interaction theory. I. Chromosome evolution in ants of the *Myrmecia pilosula* species complex (hymenoptera: formicidae: myrmeciinae). Jpn. J. Genet. 69, 137–182.
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., 2008a. Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (jack jumper ant). Arch. Biochem. Biophys. 477, 411–416.
Inagaki, H., Masuko, K., Kudo, T., 2008b. SKTXs: peptides identified from the ant *Strumigenys kumadori* that block sodium channels. In: 8th-Asia-Pacific Congress on Animal, Plant and Microbial Toxins, Hanoi, Vietnam, p. 75.
Johnson, B.R., Borowiec, M.L., Chiu, J.C., Lee, E.K., Atallah, J., Ward, P.S., 2013. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. Curr. Biol. 23, 2058–2062.
Johnson, S.R., Copello, J.A., Evans, M.S., Suarez, A.V., 2010. A biochemical characterization of the major peptides from the venom of the giant neotropical hunting ant *Dinoponera australis*. Toxicon 55, 702–710.
Jones, T., Blum, M., Howard, R., McDaniel, C., Fales, H., DuBois, M., Torres, J., 1982. Venom chemistry of ants in the genus *Monomorium*. J. Chem. Ecol. 8, 285–300.
Jones, T., Torres, J., Spande, T., Garraffo, H., Blum, M., Snelling, R., 1996. Chemistry of venom alkaloids in some *Solenopsis* (diplorhoptrum) species from Puerto Rico. J. Chem. Ecol. 22, 1221–1236.
Jones, T., Zottig, V., Robertson, H., Snelling, R., 2003. The venom alkaloids from some African *Monomorium* species. J. Chem. Ecol. 29, 2721–2727.
Jones, T.H., Andersen, A.N., Kenny, J.C., 2009. Venom alkaloid chemistry of Australian species of the *Monomorium rothsteini* complex, with particular reference to taxonomic implications. Chem. Biodivers. 6, 1034–1041.
Jones, T.H., Stahly, S.M., Don, A.W., Blum, M.S., 1988. Chemotaxonomic implications of the venom chemistry of some *Monomorium* "antarcticum" populations. J. Chem. Ecol. 14, 2197–2212.
Karbat, I., Frolow, F., Froy, O., Gilles, N., Cohen, L., Turkov, M., Gordon, D., Gurevitz, M., 2004. Molecular basis of the high insecticidal potency of scorpion α-toxins. J. Biol. Chem. 279, 31679–31686.
Kem, W.R., Wildeboer, K., LeFrancois, S., Raja, M., Marszalec, W., Braekman, J.C., 2004. Nicotinic receptor inhibition by *Tetraponera* ant alkaloids. Cell. Mol. Neurobiol. 24, 535–551.
King, G.F., Hardy, M.C., 2013. Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. Annu. Rev. Entomol. 58, 475–496.
Kohn, A.J., Anderson, T., 2009. The *Conus* Biodiversity Website. Available online: http://biology.burke.washington.edu/conus/index.php (accessed 24.08.14.).
Kuhn-Nentwig, L., 2003. Antimicrobial and cytolytic peptides of venomous arthropods. Cell. Mol. Life Sci. 60, 2651–2668.
Lai, L.C., Kuo, T.C., Huang, R.N., Wu, W.J., 2012. The insecticidal activities of fire ant (hymenoptera: formicidae) venoms against *Plutella xylostella* (lepidoptera: plutellidae) larvae. J. Econ. Entomol. 105, 1591–1596.
Lee B.J., Seo M.D., Kang S.J., Kim H.J., 2011. Analogues of antimicrobial and anticancer peptide synthesized and produced from Gaegurin 5. Promeditech, Inc. (Seoul, KR), Patent 12/301,028.
Lee, E.J., Lee, J.H., Kim, J.K., Lee, D.G., 2013. Structure-activity relationships of cecropin-like peptides and their interactions with phospholipid membrane. Biochem. Mol. Biol. Rep. 46, 284–289.
Lewis, J.C., Day, A.J., De la Lande, I.S., 1968. Phospholipase A in the venom of the Australian bulldog ant *Myrmecia pyriformis*. Toxicon 6, 109–112.
Lewis, R.J., Garcia, M.L., 2003. Therapeutic potential of venom peptides. Nat. Rev. Drug Discov. 2, 790–802.
Liu, Z., Dai, J., Dai, L., Deng, M., Hu, Z., Hu, W., Liang, S., 2006. Function and solution structure of huwentoxin-X, a specific blocker of N-type calcium channels, from the Chinese bird spider *Ornithoctonus huwena*. J. Biol. Chem. 281, 8628–8635.
Loughnan, M., Nicke, A., Jones, A., Schroeder, C.I., Nevin, S.T., Adams, D.J., Alewood, P.F., Lewis, R.J., 2006. Identification of a novel class of nicotinic receptor antagonists: dimeric conotoxins VxXIIA, VxXIIB, and VxXIIC from *Conus vexillum*. J. Biol. Chem. 281, 24745–24755.
Mackintosh, J., Flood, J., Veal, D., Beattie, A., 1999. Increase in levels of microbiota recoverable from male and larval *Myrmecia gulosa* (fabricius) (hymenoptera: formicidae) following segregation from worker ants. Aust. J. Entomol. 38, 124–126.
Mackintosh, J., Trimble, J., Beattie, A., Veal, D., Jones, M., Karuso, P., 1995. Antimicrobial mode of action of secretions from the metapleural gland of *Myrmecia gulosa* (Australian bull ant). Can. J. Microbiol. 41, 136–144.
Mackintosh, J.A., Veal, D.A., Beattie, A.J., Gooley, A.A., 1998. Isolation from an ant *Myrmecia gulosa* of two inducible O-glycosylated proline-rich antibacterial peptides. J. Biol. Chem. 273, 6139–6143.
Mangoni, M.L., Marcellini, H.G., Simmaco, M., 2007. Biological characterization and modes of action of temporins and bombinins H, multiple forms of short and mildly cationic anti-microbial peptides from amphibian skin. J. Pept. Sci. 13, 603–613.
Mangoni, M.L., Papo, N., Barra, D., Simmaco, M., Bozzi, A., Di Giulio, A., Rinaldi, A.C., 2004. Effects of the antimicrobial peptide temporin L on cell morphology, membrane permeability and viability of *Escherichia coli*. Biochem. J. 380, 859–865.
Maschwitz, U., Hahn, M., Schönegge, P., 1979. Paralysis of prey in ponerine ants. Naturwissenschaften 66, 213–214.
Maschwitz, U., Jessen, K., Maschwitz, E., 1981. Foaming in *Pachycondyla*: a new defense mechanism in ants. Behav. Ecol. Sociobiol. 9, 79–81.
Matuszek, M.A., Hodgson, W.C., King, R.G., Sutherland, S.K., 1994a. Some enzymic activities of two Australian ant venoms: a jumper ant *Myrmecia pilosula* and a bulldog ant *Myrmecia pyriformis*. Toxicon 32, 1543–1549.
Matuszek, M.A., Hodgson, W.C., Sutherland, S.K., King, R.G., 1992. Pharmacological studies of jumper ant (*Myrmecia pilosula*) venom: evidence for the presence of histamine, and haemolytic and eicosanoid-releasing factors. Toxicon 30, 1081–1091.
Matuszek, M.A., Hodgson, W.C., Sutherland, S.K., King, R.G., 1994b. Pharmacological studies of the venom of an Australian bulldog ant (*Myrmecia pyriformis*). Nat. Toxins 2, 36–43.
McGain, F., Winkel, K.D., 2002. Ant sting mortality in Australia. Toxicon 40, 1095–1100.
McIntosh, J.M., Hasson, A., Spira, M.E., Gray, W.R., Li, W., Marsh, M., Hillyard, D.R., Olivera, B.M., 1995. A new family of conotoxins that blocks voltage-gated sodium channels. J. Biol. Chem. 270, 16796–16802.
Mendes, M.A., de Souza, B.M., Marques, M.R., Palma, M.S., 2004. Structural and biological characterization of two novel peptides from the venom of the neotropical social wasp *Agelaia pallipes pallipes*. Toxicon 44, 67–74.
Moreau, C.S., Bell, C.D., Vila, R., Archibald, S.B., Pierce, N.E., 2006. Phylogeny of the ants: diversification in the age of angiosperms. Science 312, 101–104.
Nicolas, P., Amiche, M., 2013. Dermaseptins. In: Kastin, A.J. (Ed.), Handbook of Biologically Active Peptides. Elsevier, Amsterdam, pp. 350–363.
Nolde, D.E., Sobol, A.G., Pluzhnikov, K.A., Grishin, E.V., Arseniev, A.S., 1995. Three-dimensional structure of ectatomin from *Ectatomma tuberculatum* ant venom. J. Biomol. NMR 5, 1–13.
Norton, R.S., Pallaghy, P.K., 1998. The cystine knot structure of ion channel toxins and related polypeptides. Toxicon 36, 1573–1583.
Orivel, J., 2000. L'adaptation à la vie arboricole de la fourmi *Pachycondyla goeldii* (Hymenoptera: Ponerinae) (PhD Thesis). Université Paris XIII.
Orivel, J., Dejean, A., 2001. Comparative effect of the venoms of ants of the genus *Pachycondyla* (hymenoptera: ponerinae). Toxicon 39, 195–201.
Orivel, J., Redeker, V., Le Caer, J.P., Krier, F., Revol-Junelles, A.M., Longeon, A., Chaffotte, A., Dejean, A., Rossier, J., 2001. Ponericins, new antibacterial and insecticidal peptides from the venom of the ant *Pachycondyla goeldii*. J. Biol. Chem. 276, 17823–17829.
Ortiz, G., Mathias, M.I., 2006. Venom gland of *Pachycondyla striata* worker ants (hymenoptera: ponerinae). Ultrastructural characterization. Micron 37, 243–248.
Osaki, T., Omotezako, M., Nagayama, R., Hirata, M., Iwanaga, S., Kasahara, J., Hattori, J., Ito, I., Sugiyama, H., Kawabata, S., 1999. Horseshoe crab hemocyte-derived antimicrobial polypeptides, tachystatins, with sequence similarity to spider neurotoxins. J. Biol. Chem. 274, 26172–26178.
Osipov, A.V., Kasheverov, I.E., Makarova, Y.V., Starkov, V.G., Vorontsova, O.V., Ziganshin, R., Andreeva, T.V., Serebryakova, M.V., Benoit, A., Hogg, R.C., Bertrand, D., Tsetlin, V.I., Utkin, Y.N., 2008. Naturally occurring disulfide-bound dimers of three-fingered toxins: a paradigm for biological activity diversification. J. Biol. Chem. 283, 14571–14580.
Pallaghy, P.K., Norton, R.S., Nielsen, K.J., Craik, D.J., 1994. A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides. Protein Sci. 3, 1833–1839.
Palma, M.S., 2013. Hymenoptera insect peptides. In: Kastin, A.J. (Ed.), Handbook of Biologically Active Peptides. Elsevier, Amsterdam, pp. 416–422.
Pan, J., Hink, W.F., 2000. Isolation and characterization of myrmexins, six isoforms of venom proteins with anti-inflammatory activity from the tropical ant, *Pseudomyrmex triplarinus*. Toxicon 38, 1403–1413.
Park, C.B., Kim, H.S., Kim, S.C., 1998. Mechanism of action of the antimicrobial peptide buforin II: buforin II kills microorganisms by penetrating the cell membrane and inhibiting cellular functions. Biochem. Biophys. Res. Commun. 244, 253–257.
Piek, T., Duval, A., Hue, B., Karst, H., Lapied, B., Mantel, P., Nakajima, T., Pelhate, M., Schmidt, J.O., 1991a. Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata*. Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol. 99, 487–495.
Piek, T., Hue, B., Mantel, P., Nakajima, T., Schmidt, J.O., 1991b. Pharmacological characterization and chemical fractionation of the venom of the ponerine ant, *Paraponera clavata* (F.). Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol. 99, 481–486.
Platnick, N.I., 2014. The World Spider Catalog, Version 14.5. American Museum of Natural History. Available online: http://research.amnh.org/iz/spiders/catalog (accessed 24.08.14.).
Pluzhnikov, K., Nolde, D., Tertyshnikova, S., Sukhanov, S., Sobol, A., Torgov, M., Filippov, A., Arseniev, A., Grishin, E., 1994. Structural and functional studies of toxic principle of *Ectatomma tuberculatum* ant venom. Bioorg. Khimiia 20, 857–871 (in Russian).
Pluzhnikov, K., Nosyreva, E., Shevchenko, L., Kokoz, Y., Schmalz, D., Hucho, F., Grishin, E., 1999. Analysis of ectatomin action on cell membranes. Eur. J. Biochem. 262, 501–506.
Pluzhnikov, K.A., Shevchenko, L.V., Grishin, E.V., 2000. Ant polypeptide toxins. In: Rochat, H., Martin-Eauclaire, M.-F. (Eds.), Methods and Tools in Biosciences and Medicine: Animal Toxins. Birkhäuser Verlag, Basel, Switzerland, pp. 90–98.
Qiu, Y., Choo, Y.M., Yoon, H.J., Jin, B.R., 2012. Molecular cloning and antibacterial activity of bombolitin isolated from the venom of a bumblebee, *Bombus terrestris*. J. Asia Pacific Entomol. 15, 21–25.
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.

Rifflet, A., Gavalda, S., Tene, N., Orivel, J., Leprince, J., Guilhaudis, L., Genin, E., Vetillard, A., Treilhou, M., 2012. Identification and characterization of a novel antimicrobial peptide from the venom of the ant *Tetramorium bicarinatum*. Peptides 38, 363–370.
Rinaldi, A.C., Conlon, J.M., 2013. Temporins. In: Kastin, A.J. (Ed.), Handbook of Biologically Active Peptides. Elsevier, Amsterdam, pp. 400–406.
Rodríguez, A.A., Salceda, E., Garateix, A.G., Zaharenko, A.J., Peigneur, S., López, O., Pons, T., Richardson, M., Díaz, M., Hernández, Y., Ständker, L., Tytgat, J., Soto, E., 2014. A novel sea anemone peptide that inhibits acid-sensing ion channels. Peptides 53, 3–12.
Rykaczewska-Czerwinska, M., Radosz, A., Konopinska, D., Wrobel, M., Plech, A., 2008. Antinociceptive effect of poneratoxin (PoTX) in rats. Pestycydy 1-2, 135–141.
Santos, A.D., Imperial, J.S., Chaudhary, T., Beavis, R.C., Chait, B.T., Hunsperger, J.P., Olivera, B.M., Adams, M.E., Hillyard, D.R., 1992. Heterodimeric structure of the spider toxin ω-agatoxin IA revealed by precursor analysis and mass spectrometry. J. Biol. Chem. 267, 20701–20705.
Sarray, S., Luis, J., Ayeb, M.E., Marrakchi, N., 2013. Snake venom peptides: promising molecules with anti-tumor effects. In: Hernández-Ledesma, B., Hsieh, C.-C. (Eds.), Bioactive Food Peptides in Health and Disease. InTech, Croatia, pp. 219–238.
Schmidt, C.A., Shattuck, S.O., 2014. The higher classification of the ant subfamily Ponerinae (Hymenoptera: Formicidae), with a review of ponerine ecology and behavior. Zootaxa 3817, 1–242.
Schmidt, J.O., 1982. Biochemistry of insect venoms. Annu. Rev. Entomol. 27, 339–368.
Schmidt, J.O., 1990. Hymenopteran venoms: striving toward the ultimate defense against vertebrates. In: Evans, D.L., Schmidt, J.O. (Eds.), Insect Defenses: Adaptive Mechanisms and Strategies of Prey and Predators. State University of New York Press, Albany, pp. 387–419.
Schmidt, J.O., 2014. Evolutionary responses of solitary and social hymenoptera to predation by primates and overwhelmingly powerful vertebrate predators. J. Hum. Evol. 71, 12–19.
Schmidt, J.O., Blum, M.S., 1978a. A harvester ant venom: chemistry and pharmacology. Science 200, 1064–1066.
Schmidt, J.O., Blum, M.S., 1978b. Pharmacological and toxicological properties of harvester ant, *Pogonomyrmex badius*, venom. Toxicon 16, 645–651.
Schmidt, J.O., Blum, M.S., Overal, W.L., 1983. Hemolytic activities of stinging insect venoms. Arch. Insect Biochem. Physiol. 1, 155–160.
Schmidt, J.O., Snelling, G.C., 2009. *Pogonomyrmex anzensis* cole: does an unusual harvester ant species have an unusual venom? J. Hym Res. 18, 322–325.
Schultz, D.R., Arnold, P.I., 1984. Immunochemical and clinical studies of venom from the ant *Pseudomyrmex sp*. In: Tu, A.T. (Ed.), Handbook of Natural Toxicon, Insect Poisons, Allergens and Other Invertebrate Venoms. Marcel Dekker, New York, pp. 243–290.
Simmaco, M., Mignogna, G., Canofeni, S., Miele, R., Mangoni, M.L., Barra, D., 1996. Temporins, antimicrobial peptides from the European red frog *Rana temporaria*. Eur. J. Biochem. 242, 788–792.
Smith, J.J., Herzig, V., King, G.F., Alewood, P.F., 2013. The insecticidal potential of venom peptides. Cell. Mol. Life Sci. 70, 3665–3693.
Starr, C.K., 1985. A simple pain scale for field comparison of hymenopteran stings. J. Entomol. S. C. 20, 225–231.
Steinborner, S.T., Bowie, J.H., Tyler, M.J., Wallace, J.C., 1997. An unusual combination of peptides from the skin glands of Ewing's tree frog, *Litoria ewingi*. Sequence determination and antimicrobial activity. Aust. J. Chem. 50, 889–894.
Street, M.D., Donovan, G.R., Baldo, B.A., 1996. Molecular cloning and characterization of the major allergen *Myr p* II from the venom of the jumper ant *Myrmecia pilosula: Myr p* I and *Myr p* II share a common protein leader sequence. Biochim. Biophys. Acta 1305, 87–97.
Street, M.D., Donovan, G.R., Baldo, B.A., Sutherland, S., 1994. Immediate allergic reactions to *Myrmecia* ant stings: immunochemical analysis of *Myrmecia* venoms. Clin. Exp. Allergy 24, 590–597.
Suzuki, H., Iwamuro, S., Ohnuma, A., Coquet, L., Leprince, J., Jouenne, T., Vaudry, H., Taylor, C.K., Abel, P.W., Conlon, J.M., 2007. Expression of genes encoding antimicrobial and bradykinin-related peptides in skin of the stream brown frog *Rana sakuraii*. Peptides 28, 505–514.
Szolajska, E., Poznanski, J., Ferber, M.L., Michalik, J., Gout, E., Fender, P., Bailly, I., Dublet, B., Chroboczek, J., 2004. Poneratoxin, a neurotoxin from ant venom. Structure and expression in insect cells and construction of a bio-insecticide. Eur. J. Biochem. 271, 2127–2136.
Téné, N., Roche-Chatain, V., Rifflet, A., Bonnafé, E., Lefranc, B., Leprince, J., Treilhou, M., 2014. Potent bactericidal effects of bicarinalin against strains of the *Enterobacter* and *Cronobacter* genera. Food Control 42, 202–206.
Torres, A.F.C., Huang, C., Chong, C.-M., Leung, S.W., Prieto-da-Silva, A.R.B., Havt, A., Quinet, Y.P., Martins, A.M.C., Lee, S.M.Y., Rádis-Baptista, G., 2014. Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: insights into the polypeptide toxin arsenal of Hymenopterans. PLoS One 9, e87556.
Touchard, A., Dauvois, M., Arguel, M.J., Petitclerc, F., Leblanc, M., Dejean, A., Orivel, J., Nicholson, G.M., Escoubas, P., 2014a. Elucidation of the unexplored biodiversity of ant venom peptidomes via MALDI-TOF mass spectrometry and its application for chemotaxonomy. J. Proteomics 105, 217–231.
Touchard, A., Labrière, N., Roux, O., Petitclerc, F., Orivel, J., Escoubas, P., Koh, J.M., Nicholson, G.M., Dejean, A., 2014b. Venom toxicity and composition in three *Pseudomyrmex* ant species having different nesting modes. Toxicon 88, 67–76.
Uetz, P., Hošek, J., 2014. The Reptile Database. Available online: http://www.reptile-database.org (accessed 24.08.14.).
van Emden, H., 2013. Subclass pterygota, division endopterygota, order hymenoptera (sawflies, ants, bees and wasps) – c. 120,000 described species. In: Handbook of Agricultural Entomology. John Wiley & Sons, pp. 193–220.
Vetter, I., Davis, J.L., Rash, L.D., Anangi, R., Mobli, M., Alewood, P.F., Lewis, R.J., King, G.F., 2011. Venomics: a new paradigm for natural products-based drug discovery. Amino Acids 40, 15–28.
Vetter, I., Lewis, R.J., 2012. Therapeutic potential of cone snail venom peptides (conopeptides). Curr. Top. Med. Chem..
Viljakainen, L., Pamilo, P., 2008. Selection on an antimicrobial peptide defensin in ants. J. Mol. Evol. 67, 643–652.
Vizzavona, J., Zufferey, A., Rose, K., 2009. Synthesis and characterization of dimeric venom peptides. Adv. Exp. Med. Biol. 611, 493–494.
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.
Wang, X., Connor, M., Smith, R., Maciejewski, M.W., Howden, M.E., Nicholson, G.M., Christie, M.J., King, G.F., 2000. Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. Nat. Struct. Biol. 7, 505–513.
Wang, X.H., Connor, M., Wilson, D., Wilson, H.I., Nicholson, G.M., Smith, R., Shaw, D., Mackay, J.P., Alewood, P.F., Christie, M.J., King, G.F., 2001. Discovery and structure of a potent and highly specific blocker of insect calcium channels. J. Biol. Chem. 276, 40306–40312.
Wiese, M.D., Brown, S.G., 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.
Wiese, M.D., Chataway, T.K., Davies, N.W., Milne, R.W., Brown, S.G., Gai, W.P., Heddle, R.J., 2006. Proteomic analysis of *Myrmecia pilosula* (jack jumper) ant venom. Toxicon 47, 208–217.
Wilson, E.O., 1971. The Insect Societies. Harvard University Press, Cambridge, MA.
Wilson, E.O., 1990. Success and Dominance in Ecosystems: the Case of the Social Insects. Ecology Institute, Oldendorf/Luhe, Germany.
Windley, M.J., Escoubas, P., Valenzuela, S.M., Nicholson, G.M., 2011. A novel family of insect-selective peptide neurotoxins targeting insect large-conductance calcium-activated K^+ channels isolated from the venom of the theraphosid spider *Eucratoscelus constrictus*. Mol. Pharmacol. 80, 1–13.
Windley, M.J., Herzig, V., Dziemborowicz, S.A., Hardy, M.C., King, G.F., Nicholson, G.M., 2012. Spider-venom peptides as bioinsecticides. Toxins 4, 191–227.
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.
Yek, S.H., Mueller, U.G., 2011. The metapleural gland of ants. Biol. Rev. Camb. Philos. Soc. 86, 774–791.
Yi, G.B., McClendon, D., Desaiah, D., Goddard, J., Lister, A., Moffitt, J., Meer, R.K., deShazo, R., Lee, K.S., Rockhold, R.W., 2003. Fire ant venom alkaloid, isosolenopsin A, a potent and selective inhibitor of neuronal nitric oxide synthase. Int. J. Toxicol. 22, 81–86.
Zamudio, F.Z., Conde, R., Arevalo, C., Becerril, B., Martin, B.M., Valdivia, H.H., Possani, L.D., 1997. The mechanism of inhibition of ryanodine receptor channels by imperatoxin I, a heterodimeric protein from the scorpion *Pandinus imperator*. J. Biol. Chem. 272, 11886–11894.
Zelezetsky, I., Pag, U., Antcheva, N., Sahl, H.G., Tossi, A., 2005. Identification and optimization of an antimicrobial peptide from the ant venom toxin pilosulin. Arch. Biochem. Biophys. 434, 358–364.
Zeng, X.C., Wang, S.X., Zhu, Y., Zhu, S.Y., Li, W.X., 2004. Identification and functional characterization of novel scorpion venom peptides with no disulfide bridge from *Buthus martensii* Karsch. Peptides 25, 143–150.
Zhang, Z., Zhu, S., 2012. Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species. Dev. Comp. Immunol. 38, 262–274.
Zhu, S., Darbon, H., Dyason, K., Verdonck, F., Tytgat, J., 2003. Evolutionary origin of inhibitor cystine knot peptides. FASEB J. 17, 1765–1767.

Chapter Three

The Biochemical Toxin Arsenal from Ant Venoms

Chapter Three: The Biochemical Toxin Arsenal from Ant Venoms

Compound Abstract

The previous chapter illustrated the huge diversity of ant venom peptide toxins that could be useful in drug and insecticide discovery. However, another major component of ant venoms that plays a role in their defensive and predatory functions is the protein component as well as the other compounds found in this venom such as alkaloids. There was no review summarising this aspect of ant venom, therefore this chapter aims to summarise what is currently known about the other non-peptide components of ant venoms. This chapter re-examined the ant venom peptides in more detail and proposes a rational nomenclature system similar to that of other venom peptides. In addition, it describes the types and functions of proteins found in ant venoms as well as alkaloids which, together with the peptides make up a large proportion of the venom.

Certificate of authorship and originality

I certify that the work present in chapter three of this thesis has not previously been submitted as part of the requirements for a degree. I also certify that I carried out a significant amount of the work presented in this paper.

The authors listed on this manuscript contributed in the following way:

- Samira Aili: Section 1 and section 4, parts of section 2. Also proofed all sections and edited all subsequent drafts.
- Axel Touchard: Parts of sections 2 and 3 and proofed other sections
- Eduardo Gonçalves Paterson Fox: section 5
- Pierre Escoubas, Jerome Orivel, Graham Nicholson and Alain Dejean provided intellectual input and proof-read the manuscript

Principal supervisor

Professor Graham Nicholson	Samira Ryma Aili
_____	_____
Signature	Signature
_____	_____
Date	Date

toxins

MDPI

Review

The Biochemical Toxin Arsenal from Ant Venoms

Axel Touchard [1,2,*,†], Samira R. Aili [3,†], Eduardo Gonçalves Paterson Fox [4], Pierre Escoubas [5], Jérôme Orivel [1], Graham M. Nicholson [3] and Alain Dejean [1,6]

Received: 22 December 2015; Accepted: 8 January 2016; Published: 20 January 2016
Academic Editor: Glenn F. King

1 CNRS, UMR Écologie des Forêts de Guyane (AgroParisTech, CIRAD, CNRS, INRA, Université de Guyane, Université des Antilles), Campus Agronomique, BP 316, Kourou Cedex 97379, France; Jerome.Orivel@ecofog.gf (J.O.); alain.dejean@wanadoo.fr (A.D.)
2 BTSB (Biochimie et Toxicologie des Substances Bioactives) Université de Champollion, Place de Verdun, Albi 81012, France
3 Neurotoxin Research Group, School of Medical & Molecular Biosciences, University of Technology Sydney, Broadway, Sydney, NSW 2007, Australia; samira.aili@uts.edu.au (S.R.A.); graham.nicholson@uts.edu.au (G.M.N.)
4 Red Imported Fire Ant Research Center, South China Agricultural University, Guangzhou 510642, China; ofoxofox@gmail.com
5 VenomeTech, 473 Route des Dolines—Villa 3, Valbonne 06560, France; escoubas@venometech.com
6 Laboratoire Écologie Fonctionnelle et Environnement, 118 Route de Narbonne, Toulouse 31062, France
* Correspondence: t.axel@hotmail.fr; Tel.: +33-5-6348-1997; Fax: +33-5-6348-6432
† These authors contributed equally to this work.

Abstract: Ants (Formicidae) represent a taxonomically diverse group of hymenopterans with over 13,000 extant species, the majority of which inject or spray secretions from a venom gland. The evolutionary success of ants is mostly due to their unique eusociality that has permitted them to develop complex collaborative strategies, partly involving their venom secretions, to defend their nest against predators, microbial pathogens, ant competitors, and to hunt prey. Activities of ant venom include paralytic, cytolytic, haemolytic, allergenic, pro-inflammatory, insecticidal, antimicrobial, and pain-producing pharmacologic activities, while non-toxic functions include roles in chemical communication involving trail and sex pheromones, deterrents, and aggregators. While these diverse activities in ant venoms have until now been largely understudied due to the small venom yield from ants, modern analytical and venomic techniques are beginning to reveal the diversity of toxin structure and function. As such, ant venoms are distinct from other venomous animals, not only rich in linear, dimeric and disulfide-bonded peptides and bioactive proteins, but also other volatile and non-volatile compounds such as alkaloids and hydrocarbons. The present review details the unique structures and pharmacologies of known ant venom proteinaceous and alkaloidal toxins and their potential as a source of novel bioinsecticides and therapeutic agents.

Keywords: ant venom; toxins; venom biochemistry; alkaloids; formic acid; peptides; enzymes

1. Introduction

Nature contains a vast diversity of bioactive molecules and is hence a source of inspiration for chemists, biochemists and the pharmaceutical industry searching for molecules of potential therapeutic benefit or insecticidal activity. Amongst natural products, venoms are a promising source for the discovery of unique molecules as they offer a formidable array of biological properties. Venoms are complex cocktails of toxins that have been fine-tuned and pre-optimized during the course of evolution for greater efficacy and target selectivity towards prey capture as well as defence against predators [1].

Amongst venomous animals, insects represent a diverse group of organisms, with 120,000 extant hymenopteran species [2].

Similar to other hymenopterans, ants (Vespoidea: Formicidae) have evolved a venom apparatus that is derived from the ancestral reproductive system [3]. Ants are one of the most abundant groups of venomous organisms and dominate most terrestrial environments [4,5], with around 13,165 extant species described thus far [6] and an estimated total of ~25,000 species belonging to 16 different subfamilies [7–9] (Figure 1A).

Their stunning ecological diversity has also contributed to broad species diversity and, presumably, to the evolution of multiple venom types. For example, ground-dwelling ants of the subfamilies Ponerinae and Myrmicinae contain venoms that are either used for generalist predation or are specialized to prey on a locally abundant arthropod taxon (*i.e.*, isopods, myriapods, collembolans or termites). Ants that are restricted to one prey taxon and therefore may possess a specialist venom include *Psalidomyrmex procerus* (Ponerinae), which preys only on earthworms, *Strumigenys* spp. (Myrmicinae) which specializes in collembolan predation and *Megaponera analis* (Ponerinae), which preys upon a limited number of termite species [10]. While many ants overcome their prey by attacking in large numbers, some ant species are solitary hunters, suggesting that their venom is potent enough to rapidly subdue their prey similarly to solitary wasps that specifically prey on caterpillars, crickets or spiders [11].

This great taxonomical and ecological diversity has therefore allowed ants to employ their venom for several purposes such as predation and defence against predators and competitors. It can also be used for defence against microbial pathogens, communication, and as a herbicide [12–14]. Ant venoms have been found to contain an extraordinary diversity of toxins and other types of molecules including salts, sugars, formic acid, biogenic amines, alkaloids, free amino acids, hydrocarbons, peptides and proteins [13,15–21]. However, due to the small size of these organisms, the amount of venom produced by each ant is scarce in that some ants only produce around 10 μg or less of dry venom whilst other ants are capable of producing up to 300 μg, compared to spiders, scorpions and snakes that produce 0.1–300 mg of dry venom per individual [22]. This has certainly contributed to the low number of studies of ant venoms. Nevertheless, the present review aims to describe the current knowledge of the wide range of toxins present in ant venoms and their functional roles.

2. Toxins from Non-Stinging Ants

Of all ant species, only 71% are considered to be stinging species due to the fact that a few ant subfamilies have lost their ability to sting over the course of evolution. Instead of injecting their venoms, Formicinae spray their venom (which is secreted by the venom gland), whereas Dolichoderinae and Aneuretinae spray their targets with substances (*i.e.*, ketones and iridoïds) secreted by their pygidial glands [23]. Among the army ants (Dorylinae), ants from the genus *Dorylus* have a non-functional stinger and their venom glands are a source of trail pheromones. *Dorylus* species do not employ venom for prey immobilization but overcome their prey as a result of their overwhelming numbers (group hunting behaviour) and their disabling bites (Figure 1A) [24].

While stinging species inject their secretions with a stinger, stingless ants from the subfamily Formicinae spray their venoms through a special opening called the acidopore, a round orifice surrounded by a fringe of hairs constituting a unique feature of formicine ants. The peculiarity of this venom apparatus has considerably affected the chemical nature of the components secreted by the venom glands, promoting the natural selection of volatile compounds. Formic acid (methanoic acid) is the predominant compound and is presumably present in the venoms of all Formicinae, although acetic acid can also be present [25]. Formic acid, present in concentrations of up to 70% (v/v), is an alarm pheromone that, along with acetic acid, is an efficient defensive compound against competitors and predators, including vertebrates [23,26]. The major precursors for its biosynthesis are the amino acids serine and glycine [27]. By self-grooming their acidopore, *Lasius neglectus* (Formicinae) workers uptake venom into their mouth and apply the acid on brood in their colony in order to inhibit the

growth of fungal pathogens [25]. Also, upon occasional aggressive encounters, *Nylanderia fulva* (Formicinae) workers can apply their formic acid-rich venom onto their cuticle to detoxify venom alkaloids of the fire ant *Solenopsis invicta* (Myrmicinae) [28].

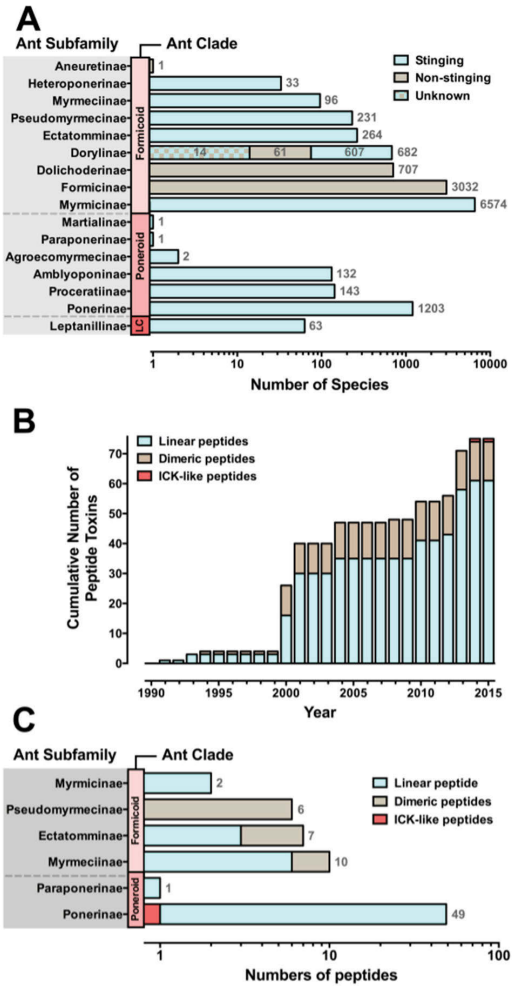


Figure 1. Species diversity of ant venom peptides. Panels A and B show data updated from Aili *et al.* 2014 [29]. (**A**) Ants have been grouped according to three clades, where LC represents the single-genus Leptanilloid clade. Stinging ants are represented by cyan bars and comprise around 71% of all ant species. Non-stinging ant subfamilies are depicted by brown bars. The total number of species in each subfamily is noted at the right of each bar. The newly described subfamily Dorylinae [7] is composed of several junior synonyms including stinging ants that belong to the subfamilies Aenictinae, Cerapachyinae, Ecitoninae, Leptanilloidinae, and ants belonging to the original subfamily Dorylinae that lost their ability to sting during evolution. Also, it is unclear whether the ants that once belonged to the junior synonym subfamily Aenictogitoninae, now subsumed in Dorylinae, are venomous or not, as only males have been observed and females are yet to be described [30]; (**B**) Cumulative total number of peptide-toxin sequences, showing the three main structural classes: cyan, linear peptides; brown, dimeric peptides; red, ICK-like peptides. Ant venom peptides remain barely investigated, with only 75 peptides sequenced to date; (**C**) Distribution of venom peptide structural classes by ant subfamily.

Several arboreal ant species use their venom gland secretion as a herbicide to destroy plants, mainly encroaching vines that compete with their host plant. An example of this is the Amazonian species *Myrmelachista schumanni* (Formicinae) that creates the "devil's gardens"—large stands of trees almost exclusively comprised of *Duroia hirsuta*, a myrmecophyte sheltering *M. schumanni* colonies in its hollow stems (domatia). Workers of this plant-ant species kill all trees situated around their host plants with their venom, which is mostly composed of formic acid. This facilitates the growth and establishment of their host plant [14,31].

3. Peptides

3.1. Ant Venom Peptides

In most animal venoms, peptides (<100 amino acids) are the predominant class of toxins and have been investigated extensively in organisms such as scorpions, cone snails, and spiders [32–35]. Although ant venoms remain very much unexplored, recent studies have revealed that the venoms of stinging ants (those belonging to the subfamilies Paraponerinae, Ponerinae, Amblyoponerinae, Dorylinae, Myrmeciinae, Pseudomyrmecinae, Myrmicinae, and Ectatomminae) are also rich in peptides, similar to other venomous animals [36,37]. To date, 75 venom peptides from only six ant subfamilies (11 ant species) have been fully sequenced (Figure 1B) [29]. These peptides have previously been classified based on their structure into three main groups: linear, dimeric, and inhibitor cystine knot (ICK)-like peptides (for a complete review of ant venom peptides, see Aili *et al.* [29]). An alternate approach to classifying ant venom peptides is to base the nomenclature on biological activity: cytolytic and neurotoxic peptides. This review will address these ant venom peptides based on their biological functions.

3.1.1. Cytolytic Peptides

Most proteomic studies on ant venoms have confirmed the prevalence of small, linear peptides (devoid of disulfide bonds) with fewer than 35 residues [19,36,37]. Most of these small peptides are cytolytic and often possess insecticidal, haemolytic and/or antimicrobial properties. Examples include ponericins from the ant *Neoponera goeldii* (Ponerinae; formerly *Pachycondyla goeldii*) that exhibit haemolytic activity, antibacterial activity against both Gram-positive and Gram-negative bacteria, as well as insecticidal activity [12]. Ponericins have been classified into three different families ("G", "W", and "L") based on sequence homology [29]. Numerous additional ponericin-like peptides were also sequenced from the venom of *Neoponera apicalis* and *N. inversa* (Ponerinae; formerly *Pachycondyla apicalis* and *P. inversa*) [38]. Thus, although the biological function of these peptides has not been characterized, their strong homology to G-, W- and L-family ponericins from *N. goeldii* suggests that they should have both antimicrobial and insecticidal activities, however, this remains to be proven. Additional homologous toxins include dinoponeratoxins from both *Dinoponera australis* and *D. quadriceps* (Ponerinae) [37,39], and bicarinalins from *Tetramorium bicarinatum* (Myrmicinae) [40]. Multiple alignment analyses have shown that these linear venom peptides display various degrees of sequence homology to each other and that they can be separated into several families [41–44]. More recently, three novel antimicrobial linear peptides with little homology to ponericin peptides have been isolated from the venom of the ant *Ectatomma brunneum* (Ectatomminae; formerly known as *E. quadridens*) [45].

Another group of peptides from ant venoms are pilosulins that constitute the major allergens of the venom of *Myrmecia pilosula* (Myrmeciinae). Pilosulin 1 is a long linear peptide (57 amino acids) and displays haemolytic and cytolytic activities [46,47]. Pilosulins 3, 4, and 5 are a group of homo- and heterodimeric peptides. Although these peptides possess some antimicrobial activity, and are classified as allergens, their biological function remains unknown [41,42].

The biological function of such membrane perturbing peptides in ant venoms is likely to be varied. Among spider and scorpion venoms, cytolytic peptides are believed to act as

membrane-disrupting agents, facilitating the passage of other disulfide-rich neurotoxins through cellular barriers to their molecular targets [48]. These linear cytolytic peptides also have direct toxic effects on prey, although this insecticidal activity is often moderate in comparison to disulfide-rich neurotoxins [32]. However, in some cases, spiders and scorpions use cytolytic-based venoms rather than neurotoxic-based venoms [49–51]. For example, the cyto-insectotoxins present in the venom of the spider *Lachesana tarabaevi* have a potent insecticidal effect and are the major insecticidal toxins in this venom [52]. Most ants seem to have evolved cytolytic-based venoms [12,36,37] and have probably developed a strategy similar to that of the spider *L. tarabaevi* for subduing their prey. The cytolytic peptides seen in most ant venoms act synergistically against different kinds of cells, disrupting membranes and rapidly killing prey.

Due to their non-selective activity, membrane-perturbing venom peptides are able to target the membranes of bacterial cells and, therefore, often exhibit some antimicrobial activity. This antimicrobial activity may be a bonus function for ants as it helps with the social immunity of the colony [53,54]. In eusocial insects, promiscuity and the relative genetic homogeneity of individuals creates ideal circumstances for the spread of infectious diseases in their nests. Therefore, the presence of multiple membrane-perturbing peptides with antimicrobial activities in ant venoms is also believed to be a way to disinfect captured prey before to bringing them back to the nest [12]. Another hypothetical function would be to assist in pre-digestion of prey. This is important since adult ants only feed on liquids due to their inability to digest solid food as a result of their narrow and constricted waists. Cytolytic activity combined with an enzymatic activity would help the degradation of cellular membranes of prey and, therefore, liquefy prey as do spider venoms [55].

Membrane-perturbing peptides are promising candidates for the future development of novel antimicrobial, insecticidal, and anticancer drugs, and have been well investigated for this purpose for many years. However, pharmaceutical research has struggled to find valid lead drug candidates as, to date, these peptides cannot sufficiently discriminate between the membranes of pathogenic cells or erythrocytes and other human host cells [48].

3.1.2. Neurotoxic Peptides

Neurotoxic peptides are widely expressed in animal venoms to assist in the rapid immobilization of prey. These neurotoxins act on a broad diversity of molecular targets, mostly ion channels, with varied selectivity, specificity, and efficacy. Many peptidic toxins modulating ion channels have been discovered in arthropod venoms. As several ant venoms have paralytic effects on arthropods, it is clear that they also contain neurotoxins that induce paralysis [11,56]. However, studies investigating the neurotoxic properties of ant venom peptides are rare and only two neurotoxic peptides have been characterized so far, as discussed below.

The first neurotoxic peptide that was isolated and characterized was poneratoxin, a small 25 residue linear peptide derived from the bullet ant *Paraponera clavata* (Paraponerinae). It has been shown to be capable of modulating voltage-gated sodium (Na_v) channels of both vertebrates and invertebrates, blocking synaptic transmission in the insect CNS [57,58]. Poneratoxin causes repetitive firing and prolongation of action potentials due to the presence of a slowly developing inward sodium current that activates at hyperpolarizing potentials. This results from a potential toxin-induced interconversion between a fast and a slow conducting state of the Na_V channel [59–61]. Due to its high efficacy, this peptide has been used in the construction of a novel bioinsecticide employing a recombinant poneratoxin-producing baculovirus [61].

The other neurotoxic peptide isolated from ant venom is a dimeric peptide, Ectatomin Et-1, from the ant *Ectatomma tuberculatum* which has been found to be the most potent neurotoxic peptide isolated from ant venoms [62]. This peptide, which is one of a family of four related peptides from *Ectatomma* spp., is a voltage-gated calcium (Ca_v) channel blocker, and also a pore-forming peptide cytotoxic to vertebrate and invertebrate cells [63,64].

3.1.3. Uncharacterized Peptides

Dimeric peptides are highly stable as they comprise two subunits that are linked by one or several disulfide bonds. Among ant venoms, dimeric peptides seem to be common in the formicoid subfamilies Ectatomminae, Pseudomyrmecinae, and Myrmeciinae (Figure 1C). In addition to pilosulins and ectatomins, the myrmexins are a group of six heterodimeric peptides isolated from the venom of *Pseudomyrmex triplarinus* (Pseudomyrmecinae) whose biological functions remain unknown [65]. Homo- and heterodimeric peptides have also been shown to be present in the venoms of *P. penetrator* [66] and *Tetraponera* sp. (Pseudomyrmecinae) [36] although their amino acid sequences and biological activity also remain uncharacterized.

The recent transcriptome analysis of the venom glands of the giant ant *Dinoponera quadriceps* has confirmed the sequence of a third structural class of ant venom peptides: the ICK-like peptides. ICK peptides contain three disulfide bonds, forming a pseudo knot, and are very stable. These are present in the venoms of cone snails and spiders, and typically have neurotoxic properties [67,68]. These peptides are a minor component of the venom of the giant ant, and their role and biological activity is still unknown [69].

Until recently, the limited amount of venom has restricted the biochemical characterization of ant venom peptides. However, recent investigations using high resolution technologies to probe ant venom peptidomes have revealed the vast and unexplored structural diversity of peptidic toxins with many small, linear peptides as well as several peptides structured by disulfide bonds that constitute novel structural classes of toxins with a likely novel, associated pharmacology [19,36,66]. Unfortunately, the number of characterized ant venom peptides is vanishingly small compared with the enormous peptide diversity revealed among ant venoms. This diversity, combined with the great ecological and taxonomical diversity of ants, suggests that ant venom peptides constitute a promising new source in the search for both novel drugs and insecticides.

3.2. Proposed Rational Nomenclature System for Ant Venom Peptides

Cutting-edge technologies such as integrated venomics represent a new gateway to exploring the venom peptides of small organisms such as ants [70–72], and have led to increases in the rate of description of novel peptidic toxins. Stinging ant venoms represent an untapped source of toxins, and the total number of peptides has been estimated to be in excess of 1 million [36]. Thus far, there is no consistent nomenclature for naming newly identified peptidic toxins in ant venoms. This may cause considerable confusion, as presumably at some point there will be a surge of ant venom-derived toxins being identified with the advent of new, more rapid and sensitive analytical strategies. It will also be difficult to quickly compare toxins and establish evolutionary relationships with no consistent nomenclature. The use of the same toxin name for different peptidic toxins with similar functions regardless of the relatedness of the source ant species is advantageous, as it allows for the quick identification of the peptide's properties; however, it does not reveal any evolutionary relationships. This is illustrated by the ponericins, a family of antimicrobial peptides, which were originally isolated from three different species in the genus *Neoponera* (Ponerinae) [12,38] and later from the unrelated ant *Ectatomma brunneum* (Ectatomminae) [45]. Another issue is the use of multiple names for the same toxin, such as with Myr p 1, pilosulin 1, and Myr p 1.0101, where all names refer to the same peptide, a linear allergenic peptide from the venom of *Myrmecia pilosula* [47,73]. Table 1 also highlights the confusing similarity in the names of ant venom alkaloids *vs.* peptides (e.g., solenopsins *vs.* ponericins) where toxins have completely different biochemical structure and function yet they are simply named after the organism from which they were obtained.

Table 1. Generic names for peptidic toxins from stinging ant subfamilies and their corresponding abbreviations.

Subfamily	Generic Toxin Name	Toxin Abbreviation
Agroecomyrmecinae	Agroecomyrmecitoxin	AGRTX
Amblyoponerinae	Amblyotoxin	ABTX
Dorylinae	Dorylitoxin	DRTX
Ectatomminae	Ectatotoxin	ECTX
Heteroponerinae	Heteroponeritoxin	HETX
Leptanilinae	Leptanilitoxin	LETX
Martialinae	Martialitoxin	MATX
Myrmeciinae	Myrmeciitoxin	MIITX
Myrmicinae	Myrmicitoxin	MYRTX
Paraponerinae	Paraponeritoxin	PPOTX
Ponerinae	Poneritoxin	PONTX
Proceratiinae	Proceratoxin	PROTX
Pseudomyrmecinae	Pseudomyrmecitoxin	PSDTX

This highlights the unmet need for a standardized nomenclature system for ant venom peptides in order to avoid confusion. Wiese *et al.* (2004) proposed a standardized nomenclature system for the *Myrmecia pilosula* venom allergens according to the International Union of Immunological Societies (IUIS) [74]. Although this system has been very useful for the pilosulins, it does not seem very practical for the naming of all ant venom peptides. This is because it provides no details on the biological activity or the molecular target of the toxin and it provides only minimal taxonomic information with no reference to subfamilies. It is therefore of great importance to adopt a nomenclature with sufficient detail and which follows the patterns of nomenclature used for other venomous organisms.

Accordingly, we propose adopting the standard nomenclature system used in naming spider [75], centipede [76] and sea anemone venom peptides [77] for naming ant venom peptides. The nomenclature is as follows:

i The toxin name should begin with a Greek letter prefix denoting the biological activity or molecular target (if known) of the peptide; see King *et al.* for a summary [75]. Where the target is not known the toxin should have a prefix of "U". As only a few pharmacological activities have been determined to date this will be an ongoing process. Haemolytic, cytolytic or antibacterial peptides that have activity against bacteria, fungus, insect or vertebrate cells are denoted by the Greek letter "M" to denote a general action to cause membrane perturbation. Neurotoxic peptides (*i.e.*, poneratoxin and ectatomin) which target voltage-gated sodium or calcium ion channels have been identified by the prefixes "δ" and "ω", respectively.

ii The Greek letter prefix will be followed by a generic toxin name. As all ants are grouped into a single family (Formicidae), we propose to slightly modify King's nomenclature which uses family names and use the 13 extant stinging subfamily names instead (Figure 1A). This will allow the toxins to be compared and will highlight the evolutionary relationship between different toxins. A list of the proposed generic toxin names and their corresponding abbreviations is proposed in Table 1 for all extant subfamilies of stinging ants. These names and their abbreviations have been carefully chosen so that they do not overlap with current toxins from other venomous animals nor other chemical groups. NB: non-stinging ants are thought to contain mostly non-peptidic venom components, and are therefore not included.

iii The toxin name is then followed by an uppercase letter that indicates the genus of the ant and a lowercase letter which identifies the species of the ant from which it was isolated. An additional one or two lowercase letters may be required to distinguish species with the same first letters. Due to several taxonomic revisions concerning ants, their species names are often subject to modifications; therefore, all ant venom studies should follow the world's largest online ant database AntCat [6] when defining the most current species name.

iv Finally, an alpha-numerical code will be used to separate different structural classes of peptides based on their molecular scaffold and amino acid sequences. An Arabic numeral will be used to distinguish different toxins from the same species with little amino acid homology or different three-dimensional structures. A lowercase letter will also be added in order to distinguish isotoxins. The isotoxins are named based on the sequence alignment analyses presented in the review of Aili *et al.* [29]. The definition of isotoxin groups by Olivera *et al.* [77] will be used to distinguish isotoxins. Toxins from the same ant species will be classified in the same isotoxin group when there is $\geqslant$ 50% similarity in molecular size, biological function as well as amino acid sequence.

We have applied this proposed nomenclature to all the known peptidic toxins isolated and sequenced from the venoms of both poneroid (Table 2) and formicoid (Table 3) ants. Using this rational nomenclature method, we have uniformly renamed the 75 currently sequenced peptidic ant toxins. This new nomenclature will provide a clearer means of identifying and classifying former toxins as well as new peptides, which will facilitate future exploration of ant venom peptides.

Table 2. Venom peptide toxins from poneroid ant species renamed according to the proposed rational nomenclature system.

Species (Subfamily)	Original Toxin Name	Proposed Toxin Name	Abbreviation	Reference
Paraponera clavata (Paraponerinae)	Poneratoxin	δ-Paraponeritoxin-Pc1a	δ-PPOTX-Pc1a	[57]
Neoponera goeldii (Ponerinae)	Ponericin G1	M-poneritoxin-Ng3a	M-PONTX-Ng3a	[12]
	Ponericin G2	U_1-poneritoxin-Ng3b	U_1-PONTX-Ng3b	[12]
	Ponericin G3	M-poneritoxin-Ng3c	M-PONTX-Ng3c	[12]
	Ponericin G4	M-poneritoxin-Ng3d	M-PONTX-Ng3d	[12]
	Ponericin G5	U_1-poneritoxin-Ng3e	U_1-PONTX-Ng3e	[12]
	Ponericin G6	M-poneritoxin-Ng3f	M-PONTX-Ng3f	[12]
	Ponericin G7	U_1-poneritoxin-Ng3g	U_1-PONTX-Ng3g	[12]
	Ponericin L1	U_1-poneritoxin-Ng2a	U_1-PONTX-Ng2a	[12]
	Ponericin L2	M-poneritoxin-Ng2b	M-PONTX-Ng2b	[12]
	Ponericin W1	M-poneritoxin-Ng1a	M-PONTX-Ng1a	[12]
	Ponericin W2	U_1-poneritoxin-Ng1b	U_1-PONTX-Ng1b	[12]
	Ponericin W3	M-poneritoxin-Ng1c	M-PONTX-Ng1c	[12]
	Ponericin W4	M-poneritoxin-Ng1d	M-PONTX-Ng1d	[12]
	Ponericin W5	M-poneritoxin-Ng1e	M-PONTX-Ng1e	[12]
	Ponericin W6	M-poneritoxin-Ng1f	M-PONTX-Ng1f	[12]
Neoponera inversa (Ponerinae)	Ponericin Pi I1	U_1-poneritoxin-Ni3a	U_1-PONTX-Ni3a	[38]
	Ponericin Pi I2	U_1-poneritoxin-Ni3b	U_1-PONTX-Ni3b	[38]
	Ponericin Pi I3	U_1-poneritoxin-Ni3c	U_1-PONTX-Ni3c	[38]
	Ponericin Pi I4	U_1-poneritoxin-Ni3d	U_1-PONTX-Ni3d	[38]
	Ponericin Pi II1	U_1-poneritoxin-Ni1a	U_1-PONTX-Ni1a	[38]
	Ponericin Pi II2	U_1-poneritoxin-Ni1b	U_1-PONTX-Ni1b	[38]
	Ponericin Pi III1	U_1-poneritoxin-Ni2a	U_1-PONTX-Ni2a	[38]
Neoponera apicalis (Ponerinae)	Ponericin Pa I1	U_1-poneritoxin-Na3a	U_1-PONTX-Na3a	[38]
	Ponericin Pa I2	U_1-poneritoxin-Na3b	U_1-PONTX-Na3b	[38]
	Ponericin Pa II 1	U_1-poneritoxin-Na1a	U_1-PONTX-Na1a	[38]
	Ponericin Pa II 2	U_1-poneritoxin-Na1b	U_1-PONTX-Na1b	[38]
	Ponericin Pa IV1	U_1-poneritoxin-Na2a	U_1-PONTX-Na2a	[38]
Dinoponera australis (Ponerinae)	Dinoponeratoxin Da-1039	U_1-poneritoxin-Da1a	U_1-PONTX-Da1a	[39]
	Dinoponeratoxin Da-1585	U_1-poneritoxin-Da3a	U_1-PONTX-Da3a	[39]
	Dinoponeratoxin Da-1837	U_1-poneritoxin-Da2a	U_1-PONTX-Da2a	[39]
	Dinoponeratoxin Da-2501	U_1-poneritoxin-Da3b	U_1-PONTX-Da3b	[39]
	Dinoponeratoxin Da-3105	U_1-poneritoxin-Da4a	U_1-PONTX-Da4a	[39]
	Dinoponeratoxin Da-3177	M-poneritoxin-Da4b	M-PONTX-Da4b	[39]
Dinoponera quadriceps (Ponerinae)	Dinoponeratoxin Dq-762	U_1-poneritoxin-Dq1a	U_1-PONTX-Dq1a	[37]
	Dinoponeratoxin Dq-987	U_1-poneritoxin-Dq1b	U_1-PONTX-Dq1b	[37]
	Dinoponeratoxin Dq-1031	U_1-poneritoxin-Dq1c	U_1-PONTX-Dq1c	[37]
	Dinoponeratoxin Dq-1062	U_1-poneritoxin-Dq2a	U_1-PONTX-Dq2a	[37]
	Dinoponeratoxin Dq-1133	U_1-poneritoxin-Dq2b	U_1-PONTX-Dq2b	[37]
	Dinoponeratoxin Dq-1289	U_1-poneritoxin-Dq2c	U_1-PONTX-Dq2c	[37]
	Dinoponeratoxin Dq-1839	U_1-poneritoxin-Dq3a	U_1-PONTX-Dq3a	[37]
	Dinoponeratoxin Dq-1840	U_1-poneritoxin-Dq3b	U_1-PONTX-Dq3b	[37]
	Dinoponeratoxin Dq-1856	U_1-poneritoxin-Dq3c	U_1-PONTX-Dq3c	[37]
	Dinoponeratoxin Dq-1897	U_1-poneritoxin-Dq3d	U_1-PONTX-Dq3d	[37]
	Dinoponeratoxin Dq-1984	U_1-poneritoxin-Dq3e	U_1-PONTX-Dq3e	[37]
	Dinoponeratoxin Dq-3104	M-poneritoxin-Dq4a	M-PONTX-Dq4a	[37]
	Dinoponeratoxin Dq-3162	M-poneritoxin-Dq4b	M-PONTX-Dq4b	[37]
	Dinoponeratoxin Dq-3163	U_1-poneritoxin-Dq4c	U_1-PONTX-Dq4c	[37]
	Dinoponeratoxin Dq-3178	U_1-poneritoxin-Dq4d	U_1-PONTX-Dq4d	[37]
	Dinoponeratoxin ICK-like	U_1-poneritoxin-Dq5a	U_1-PONTX-Dq5a	[69]

Table 3. Venom peptide toxins from formicoid ant species renamed according to the proposed rational nomenclature system.

Species (Subfamily)	Original Toxin Name	Proposed Toxin Name	Abbreviation	Reference
Tetramorium bicarinatum (Myrmicinae)	Bicarinalin 1 P 17	M-myrmicitoxin-Tb1a U_1-myrmicitoxin-Tb2a	M-MYRTX-Tb1a U_1-MYRTX-Tb2a	[40] [40]
Ectatomma tuberculatum (Ectatomminae)	Ectatomin-Et1 Ectatomin-Et2	ω/M-ectatotoxin-Et1a U_1-ectatotoxin-Et1b	ω/M-ECTX-Et1a U_1-ECTX-Et1b	[62] [78]
Ectatomma brunneum (Ectatomminae)	Ectatomin-Eq1 Ectatomin-Eq2 Ponericin-Q42 Ponericin-Q49 Ponericin-Q50	U_1-ectatotoxin-Eb1a U_1-ectatotoxin-Eb1b M-ectatotoxin-Eb2a M-ectatotoxin-Eb2b M-ectatotoxin-Eb2c	U_1-ECTX-Eb1a U_1-ECTX-Eb1b M-ECTX-Eb2a M-ECTX-Eb2b M-ECTX-Eb2c	[78] [78] [45] [45] [45]
Pseudomyrmex triplarinus (Pseudomyrmecinae)	Myrmexin I Myrmexin II Myrmexin III Myrmexin IV Myrmexin V Myrmexin VI	U_1-pseudomyrmecitoxin-Pt1a U_1-pseudomyrmecitoxin-Pt1b U_1-pseudomyrmecitoxin-Pt1c U_1-pseudomyrmecitoxin-Pt1d U_1-pseudomyrmecitoxin-Pt1e U_1-pseudomyrmecitoxin-Pt1f	U_1-PSDTX-Pt1a U_1-PSDTX-Pt1b U_1-PSDTX-Pt1c U_1-PSDTX-Pt1d U_1-PSDTX-Pt1e U_1-PSDTX-Pt1f	[65] [65] [65] [65] [65] [65]
Myrmecia pilosula (Myrmeciinae)	Myr p 157–112 Myr p 1 57–112 (Ile5) Myr p 1 65–112 Myr p 1 68–112 Myr p 1 71–112 Myr p 1 86–112 Pilosulin 3a Pilosulin 3b Pilosulin 4 Pilosulin 5	M-myrmeciitoxin-Mp1a M-myrmeciitoxin-Mp1b M-myrmeciitoxin-Mp1c M-myrmeciitoxin-Mp1d M-myrmeciitoxin-Mp1e U_1-myrmeciitoxin-Mp1f M-myrmeciitoxin-Mp2a M-myrmeciitoxin-Mp2b M-myrmeciitoxin-Mp3a M-myrmeciitoxin-Mp4a	M-MIITX-Mp1a M-MIITX-Mp1b M-MIITX-Mp1c M-MIITX-Mp1d M-MIITX-Mp1e U_1-MIITX-Mp1f M-MIITX-Mp2a M-MIITX-Mp2b M-MIITX-Mp3a M-MIITX-Mp4a	[79] [79] [79] [79] [79] [79] [41] [41] [41] [42]

4. Ant Venom Proteins

Although ants are amongst the most abundant and diverse of all social insects [5], there remains limited information in the literature regarding their venom protein characteristics. Most published studies have investigated the allergenic properties of ant venoms [43,74,80]. This is especially true for proteomic data, even though such information can give insights into the functions of venom components [81]. This paucity of data is mainly due to the limited amount of venom that can be obtained from stinging ants [82] and the laborious nature of venom dissections and extractions [83,84]. Nevertheless, current data shows that ant venom proteins are highly diverse, as is the case with the peptide component. This diversity is further exemplified with the vastly different patterns of venom protein expression across ant subfamilies which has been attributed to both phylogenetic and behavioural differences between ants [82,85].

One of the first studies to report the presence of proteins in ant venoms was that of Leluk *et al.* [81] which found proteins ranging from 24 to 75 kDa in all six ants investigated *(Dinoponera grandis*, *Diacamma* sp., *Paraponera clavata*, *Odontoponera transversa*, *Pogonomyrmex rugosus*, and *Po. maricopa*). The two most investigated ants using proteomics are *Myrmecia pilosula* (Australian jack jumper ant) and *Solenopsis invicta* (red imported fire ant) due to frequent allergic reactions to their sting which can lead to anaphylaxis and, in some extreme cases, death [83,85]. In fact, the first ever published study proving the presence of proteins in ant venoms was performed on the red imported fire ant in 1979, where the authors managed to extract and enzymatically assay venom proteins by employing chromatographic separation on a massive amount of manually-milked venom (*ca.* 120 mg) [86].

Myrmecia pilosula venom is mainly composed of peptides, however, it does contain six proteins between 26 and 90 kDa [43,74]. Most of its activity was originally attributed to the pilosulin peptides (see Section 3.1.1), however, it was later found that the proteins also play a role in the allergic reactions [43,74,80,87]. An interesting feature of this venom, which had hindered investigations of its composition in the past, is its highly basic nature which makes venom proteins more difficult to separate based on isoelectric point (p*I*) when using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) [80]. The basic nature of these venom proteins is common with defensive venoms, such as that of bees, and the proteins responsible for this effect usually cause painful or cytolytic effects [81]. However, this is not common amongst ants, as the original study performed by Leluk *et al.* (1989)

revealed that the majority of the six ants investigated contained acidic venoms, particularly that of *Paraponera clavata* [81].

The proteome of *Solenopsis invicta* has only recently been investigated using 2D-PAGE based on a commercial protein extract, due to the small amount of protein present in the venom in comparison to its high alkaloid content (>95% alkaloids; see Section 5.2) [74,84,88]. It has been postulated that one reason for *S. invicta* venom being less proteinaceous than those of other ant venoms is that this ant evolved more recently compared to the more ancestral ants with higher venom protein content [89]. Other venoms that have been shown to be proteinaceous in nature are those from Ponerinae, Dorylinae, Pseudomyrmecinae (e.g., *Pseudomyrmex triplarinus* with proteins making up 42% of the venom's dry weight [90]), and even some Myrmicinae such as *Myrmica* spp. and *Pogonomyrmex* spp.

The proteins identified in ant venoms have been assigned to one of the following three major functional protein groups: housekeeping proteins, body muscle proteins, or true venom proteins (classification modified from [83]). Previous transcriptomic studies have revealed that the majority of transcripts identified from the venom glands (~40%–65%) are housekeeping proteins such as ribosomal proteins which come from the venom gland tissues [69,85,87]. These predicted housekeeping and body muscle proteins have also been predicted by other approaches such as proteomics [83] and are therefore not covered in this review as they are not true venom components. Transcriptomics has revealed that true venom proteins make up a small fraction of the transcripts being expressed in venom gland tissues (<1%–5%).While there are indications of several new venom gland components using transcriptomics, one must be cautious in considering that not all potential transcripts identified are necessarily translated into proteins, and must be confirmed using proteomic techniques. Therefore, true venom components which have been confirmed by proteomic studies have been categorized into (1) toxic venom proteins; (2) non-toxic proteins involved in protecting other venom components and the gland tissue; and (3) non-toxic proteins involved in chemical communication.

4.1. Toxic Venom Proteins

The present review will discuss those proteins which have been associated with venom diffusion and toxicity to prey or predators as well as major allergenic proteins revealed by venom gland proteomic and transcriptomic studies. These proteins can be further classified into five subgroups: (1) neurotoxins; (2) proteins that promote venom diffusion or modulate prey defense mechanisms; (3) proteins that promote tissue damage or cause inflammation; (4) allergens; and (5) antimicrobial proteins involved in colony/food asepsis. Ant venom toxic proteins commonly interfere with tissue signalling, lipid homeostasis, protein–protein interactions or trafficking of vesicles [69]. While only two ant venom gland transcriptomes have been published to date, these have revealed an enormous amount of novel information regarding potential proteins in the venom gland [69,85,91]. It is clear that further transcriptomic studies are necessary, as it would make the current difficult task of novel protein annotation a lot clearer [83]. Moreover, a significant number of predicted proteins are apparently unique to ant venoms, as they are not homologous to previously deposited sequences in databanks from other tissues or organisms for the first time [69,85,91].

4.1.1. Neurotoxic Proteins

An increasing number of proteins that cause neurotoxicity are being revealed in ant venoms. For example, the proteomic investigation of *Solenopsis invicta* venom [83] revealed the presence of three 18–43.1 kDa neurotoxins similar to proteins from other arthropods. One of these proteins is homologous to U_5-ctenitoxin-Pk1a-like protein which has been implicated in causing spastic paralysis in mice [92]. The other protein found is homologous to the neurotoxic alpha-toxin Tc48a-like protein, which is also lethal to mice through its action on Na_V channels [83,93]. The third neurotoxic protein was homologous to *Scolopendra* (Chilopoda) toxin-like proteins which are not lethal to vertebrates, but are neurotoxic to insects and crustaceans [94,95].

Phospholipases have been described as one of the major proteins in several hymenopteran venoms and are considered potent neurotoxic, cytotoxic and allergenic proteins [82,96]. Phospholipases (PL) hydrolyze the different ester linkages in phospholipids. There are five major types: PLA_1, PLA_2, and PLC (which cleave ester bonds at positions sn-1, sn-2, and sn-3, respectively), PLD (which is mainly found in plants that attacks the nitrogenous base of the phospholipids) and PLB (which cleaves lysophospholipids) [96]. The most commonly reported phospholipase in ant venoms is PLA_2 [83,97–99], however, there have been isolated reports of PLA_1, PLB, and PLD as well [69,99,100]. Venom phospholipases from various animals have been demonstrated to induce neurotoxicity, platelet activation, allergic reactions, haemolysis, and tissue damage. Unlike snake venom phospholipases which are lethal to their prey [69,101], ant venom phospholipases have not been implicated in causing lethality of prey, however, it is likely that they have synergistic activity with other toxic proteins which cause lethality [83,102].

4.1.2. Proteins that Promote Venom Diffusion or Modulate Victim Defense Mechanisms

Examples of proteins involved with tissue damage would include phospholipases, hyaluronidases, proteases, and venom acid phosphatases. Hyaluronidase is implicated in aiding the spread of venom through the host tissues. This results from the hydrolysis of hyaluronic acid and chondroitin sulphate which are essential components of connective tissues [82], thereby increasing membrane permeability, reducing viscosity, and making tissues more permeable to venom neurotoxins [82,103]. This function has been used clinically to assist in the absorption of fluids administrated by subcutaneous or intramuscular injection, and to improve the diffusion of local anaesthetics [82]. Hyaluronidase activity was observed in all nine ant venoms tested by Schmidt *et al.* [102], however, the activity was lower in comparison to that observed in wasp venoms. The ants with the highest hyaluronidase activity were the tropical ants *Paraponera clavata* and *Ectatomma tuberculatum* [102]. Other ants which have been found to contain hyaluronidase activity in the venom are *Myrmecia pyriformis* [104], *Pseudomyrmex triplarinus* [90] and *Solenopsis invicta* [86].

Proteases are responsible for moderate necrosis in some tissues of patients following envenomation by various venomous animals. Little information is available on venom proteases in insects, especially in ant venoms [82], and clinical reports of necrosis from ant stings are likely a result of secondary bacterial infections [60]. However, proteases have been reported in *Eciton burchellii* in very high levels [102]. Transcriptomic analysis of the venom gland of the ant *Tetramorium bicarinatum* suggested that the main toxin-like proteins are metalloproteinases that degrade proteins and hydrolyse specific peptide bonds [85]. The presence of a metalloproteinase in ant venom is significant as they are thought to be involved in disruption of the host's coagulation cascade as well as in generating a more digestible prey [85]. A metalloproteinase has also been found in the venom of the fire ant *Solenopsis invicta* using proteomics techniques [83]. In wasps, metalloproteinases have been associated with inflammation, necrosis, oedema, and skin damage after massive attacks on humans [83].

Other proteases which have been identified in several hymenopterans and some ants are carboxylesterases [69,105]. These enzymes hydrolyse carboxylic acid esters into acids and alcohols; this enzyme has been considered to have a protective function for the organism as it promotes cellular detoxification by inactivating carcinogens and toxicants. Pesticides and drugs usually contain ester moieties that are susceptible to these enzymes and are therefore degraded by this enzyme. This enzyme has been found in the genomes of the ants *Harpegnathos saltator*, *Camponotus floridanus*, *Acromyrmex echinatior* [106], and, more recently, in *Dinoponera quadriceps* through transcriptomic analysis [69].

Another interesting finding of the investigation by Schmidt *et al.* [102] was the presence of phosphodiesterase activity in the venom of the ants *Ectatomma tuberculatum* and *Paraponera clavata*. Phosphodiesterases are more common in snake venoms and have not yet been reported in insect venoms [102,107]. They have been associated with catalysing the action of other active or toxic venom component functions [102] which can cause cell lysis or DNA/RNA degradation in the prey [107].

Several enzymes potentially involved with targeting major host defence cascades were also revealed through transcriptomics analysis of *Tetramorium bicarinatum* venom [85,91]. An example of such a protein is phenoloxidase which is a multicopperoxidase which generates highly reactive and toxic quinine intermediates that clear bacterial infections from the insect. This is because they cross-link bacteria to a protein on the cytoplasmic membrane of haemocytes [108,109]. This protein has been identified as an important venom compound among parasitic wasps to disrupt their host's immune systems.

4.1.3. Proteins that Promote Tissue Damage or Cause Inflammation

Phospholipases are a common protein found in ant and other hymenopteran venoms that, in addition to the activities described in Section 4.1.1, cause disruption of the phospholipid membrane leading to pore formation, inflammation, and cell lysis [83,110,111]. One of the earliest reports of phospholipase activity in ant venoms was in the bulldog ant *Myrmecia pyriformis* [112]. However, since then many other ants have been reported to express phospholipases in their venom either through enzymatic, proteomic or transcriptomic studies. For example, the ants *Paraponera clavata, Pogonomyrmex occidentalis, Pogonomyrmex badius, Eciton burchellii* [102], *Pseudomyrmex triplarinus* [90], *Dinoponera grandis, Solenopsis invicta*, and *Ectatomma tuberculatum* [96] have all been reported to have phospholipase activity using enzymatic studies. Through transcriptomic techniques, the additional ant species *Dinoponera quadriceps* [69] and *Tetramorium bicarinatum* [91] have also been reported to have phospholipases. According to current data, the phospholipase activity of ant venoms seems to be lower than that of wasps. While the activity of *Pogonomyrmex badius* is comparable to that of the yellow jacket wasp [102], *Tetramorium caespitum* has with no reported phospholipase activity [89]. This indicates that venom phospholipase activity is not broadly distributed in all ant species.

As previously mentioned, PLDs have not been reported in hymenopteran venoms, however they have been recently predicted in two ant venoms, *Dinoponera quadriceps* and *Solenopsis invicta* [69,83]. The presence of PLD in ant venoms is a significant finding as it has only been previously reported in spider venoms, highlighting the need for confirmation by enzymatic assays. The enzyme has been often referred to as sphingomyelinase D where it can hydrolyze sphingomyelin containing membranes, or phospholipase D by virtue of its wider spectrum of lipid substrates. In the brown spider *Loxoceles gaucho,* sphingomyelinase activity results in characteristic dermonecrotic lesions, which typically follow a massive inflammatory response [113,114].

The 2D-PAGE analysis of *Solenopsis invicta* [83] revealed the presence of several other proteins that could promote tissue damage. These included myotoxin 2-like proteins previously found in snake venom proteins that have been reported to cause necrosis of tissue by increasing cytolysis and microvascular permeability [83,115] and PSTx 60-like protein previously identified from sea anemones which also promotes tissue damage due to a haemolytic action [116].

Venom acid phosphatases are common toxic inflammatory enzymes in venoms and have known digestive functions and toxic actions to cause histamine release and cell lysis [117,118]. Venom acid phosphatases also seem to be common in ant venoms in different levels of abundance, with *Eciton burchellii* and *Ectatomma tuberculatum* having the highest reported activity [98,102]. Interestingly, ants have higher activities of this enzyme overall compared to wasp venoms [102]. This enzyme is also a typical tissue enzyme, so it has been suggested that it might be a contamination from the venom gland tissues. However, venom collected by electrical stimulation was found to contain these proteins suggesting this is a true venom protein [102].

4.1.4. Allergens

An allergen is any substance capable of eliciting an allergic reaction. Often, this can culminate in anaphylactic shock, which is a serious reaction involving oedema and systemic smooth muscle stimulation. The autacoid histamine is one of the key molecules involved in the mediation of hypersensitivity. Whilst all known allergens are relatively large molecules, with most allergens

consisting of proteins or protein conjugates, the exact chemical properties leading to allergenicity are not well understood. There are also other substances which can induce the release of histamine, potentially leading to a hypersensitive state. In summary, (i) different families of proteins may work as potent allergens, although this cannot be reliably predicted from their sequence, nor does it depend on enzymatic activity; and (ii) some enzymes and non-proteinaceous substances can induce histamine release, either directly or by activating the immune system via local reaction products, thus also acting as allergens.

Allergic reactions, as well as anaphylaxis, are a common manifestation of stings by species of the order Hymenoptera, and ants are no exception. Indeed, it has been reported that perhaps over 50% of venom secretion proteins are allergenic proteins [85]. In Australia, most allergic reactions to ants are attributed to ants of the genus *Myrmecia*, particularly the jack jumper ant *Myrmecia pilosula* [43,74,80]. This ant has been alleged to rival the fire ant *Solenopsis invicta* in terms of venom allergenicity. Within hymenopterans, proteins ranging from 20 to 50 kDa are usually the source of these allergenic effects [80]. However, up until fairly recently, most allergenic activity was attributed to the peptide components. For example, *M. pilosula* allergenicity was mainly attributed to the pilosulins, however, seven proteins (20–90 kDa) have now been discovered and are believed to contribute to the allergic effects manifested after a sting [74,81].

An example of an ant whose allergenicity was attributed to proteins is that of *Solenopsis invicta* whose four main allergens (Sol i 1–4) are between 14 and 37 kDa [98,99,119–121]. Initially, it was believed that they all possessed phospholipase activities which was causing the allergenic effects [98], due to previous work suggesting that phospholipases cause the release of histamine [104]. A pioneering study on these allergens using enzymatic assays ruled out hyaluronidases and venom acid phosphatases as the cause of the allergenic effects of fire ant venoms [98]. To date, little is known about the biological activities of most of these proteins. Sol i 1 contains both phospholipase A_1 and B activity and was shown to be more related to vespid phospholipases than bee phospholipases [83,99]. Sol i 2 has had its crystal structure recently determined (see Figure 2) [119,122], and was suggested to play a role in binding temporarily to hydrophobic factors such as trail pheromones [122]. Sol i 3 is a dimeric protein that is a member of the antigen 5 protein family with no known enzymatic activity (see Figure 3) [83,123–125]. Sol i 4 is homologous to Sol i 2 in sequence but occurs as a monomer [122] its biological function is also still unknown. Like Sol i 2, it is unique to ant venoms and does not seem to be homologous to any bee or vespid proteins [83,124]. Further potential allergens have been recently identified among fire ant venom proteins, which proved to be more diverse than previously thought, however specific immunological tests are necessary to confirm which ones are the most allergenic proteins [83]. Homologs of Sol i 2 and Sol i 3 have been predicted from the transcriptome of *Dinoponera quadriceps*, however, they have several unique amino acids which show species-specific diversification of these proteins [69].

Brachyponera chinensis (formerly *Pachycondyla chinensis*) is another ant whose proteins account for the majority of the allergic manifestations. Thus far, nine proteins (all >10 kDa) have been identified as allergenic using 2D-PAGE and western blots. The major allergenic protein seems to be a 23 kDa protein with a p*I* 8.7 given that the majority of IgE proteins from hypersensitive patients reacted against this protein. This protein was found to be homologous to *Solenopsis invicta* Sol i 3 and the antigen 5 family of proteins [124].

4.1.5. Antimicrobial Proteins

Antimicrobial proteins have bactericidal activity and include proteins with a colony asepsis role, preventing contamination of stored food as well as colony individuals, including the brood. An interesting group of proteins involved with colony asepsis from *Solenopsis invicta* venom are the bactericidal transferrins [83]. These proteins chelate free Fe^{3+} in biological fluids, making it unavailable for use by bacteria that need it for their survival [126].

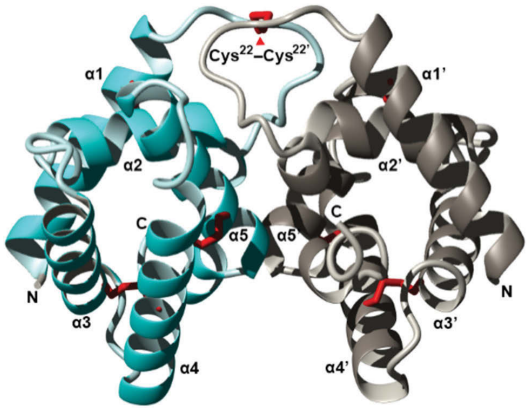


Figure 2. Crystal structure of the fire ant (*Solenopsis invicta*) venom allergen Sol i 2 dimer (PDB accession 2YGU) shown as a ribbon diagram. The two monomers (cyan and silver) dimerized by a disulfide-bond on symmetrical residues Cys^{22}. Each monomer is composed of five α-helices, with helices α2–α5 surrounding a central hydrophobic cavity. Figure modified from Borer *et al.* Redrawn from [122].

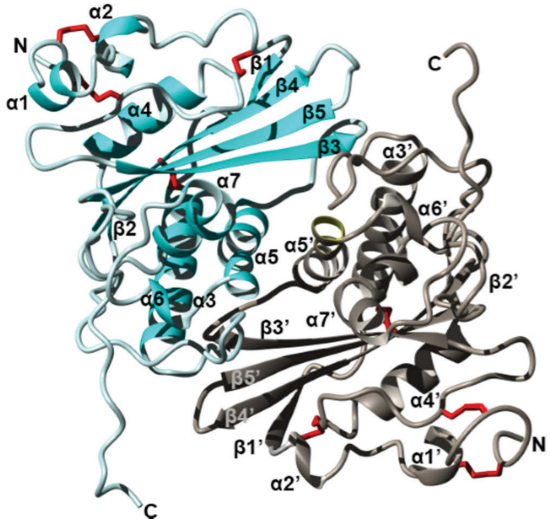


Figure 3. Crystal structure of the fire ant (*Solenopsis invicta*) venom allergen Sol i 3 dimer (PDB accession 2VZN). Ribbon diagram revealing the overall structure of each monomer which contains seven helices (α1–α7) and five beta strands (β1–β5), arranged as three stacked layers, giving rise to an α–β–α sandwich. Two units (cyan and silver) form a dimer by non-disulfide bonds involving symmetrical residues in helix α5 and α5′. Disulfide bridges are shown in red, and N and C termini are labelled. Figure modified from Padavattan *et al.* Redrawn from [125].

4.2. Identified Proteins with Unknown Functions

Due to the lack of proteomic data, a recurring issue with ant venom proteomic studies is the abundance of unassigned and unannotated predicted proteins in database searches. According to published venom gland transcriptomes [69,85,91], there are thousands of unique hypothetical proteins which could not be assigned to any biological function or previously described protein through

sequence searches. Such proteins could include unique venom toxins that could be investigated as potential insecticidal or antimicrobial drugs.

5. Ant Alkaloids

Alkaloids are defined as a heterogeneous assembly of secondary metabolite cyclic compounds containing nitrogen atoms in a negative oxidation state [127]. Nowadays, around 14,000 different alkaloids are known [127] with the inevitable result that the chemistry of alkaloids is very complex with numerous chemical subdivisions. They are primarily found in plants, particularly in the Angiosperma, where alkaloid production pathways seem to have diversified mainly as a protection against defoliation by herbivores [128]. However, it is not only plants that contain alkaloids, with a number of alkaloids having been isolated from fungi, and different classes of vertebrates (e.g., numerous toads, the musk deer, and beavers) and invertebrates (mainly marine sponges, myriapods, and insects). Alkaloid-rich insects are particularly prevalent among lepidopterans, beetles, and ants [129].

Alkaloids in ants were first reported in the early 1970s [130] and they have been reported in an increasing number of different ant groups, particularly as venom secretions (for a summary see Table 4). Although wasp venoms may contain amines and several other low molecular weight compounds, venom alkaloids seem to be a particularity of ants within Hymenopterans [127,131].

Table 4. Ant genera containing venom alkaloids.

Subfamily	Ant Genus	Structural Family	Trivial Name	Reference
Myrmicinae	*Atta Acromyrmex*	Pyrroles	Trail pheromone	[132]
	Messor	Pyridines	Anabaseine Anabasine	[133,134]
	Aphaenogaster	Pyridines	Anabaseine	[135]
	Megalomyrmex	Pyrrolidines Pyrrolines Pyrrolizidines	-	[136]
	Monomorium	Farnesylamine Pyrrolidines Indolizidines	Monomorines (trail pheromones)	[137]
	Myrmicaria	Polycyclic indolizidines Pyrrolo-indolizidines	Myrmicarins	[138]
	Solenopsis	Piperidine and piperideine Dialkylpyrrolidines and Pirrolines Indolizidines	Solenopsins Histrionicotoxins Gephyrotoxin	[139–141]
	Carebarella [1]	Pyrrolidines	Histrionicotoxins Gephyrotoxin	[142]
	Leptothorax *Harpagoxenus*	Alkylpyrrolidines	–	[143]
Formicinae	*Nylanderia* *Brachymyrmex*	Alkyl-hydroxyl-indolizidines	Pumiliotoxins [2]	[144]
Pseudomyrmecinae	*Tetraponera*	Pyrimidines	Tetraponerines	[145]

[1] This genus has been very recently incorporated within *Solenopsis* [146]; [2] Pumiliotoxins are important alkaloids isolated from mixed whole ant extracts of other groups, but as yet it is unknown whether these alkaloids come from the venom apparatus.

In the following sections, only alkaloids with toxic activities found in ant venoms are discussed; for an overview of general alkaloid structures and classification see Anisewiski [127]. Ant alkaloids can be either monocyclic, bicyclic, tricyclic or polycyclic (the latter being derived from tricyclic alkaloids), thus displaying considerable diversity (Figure 4). The venom from ants of the same species group may contain several different alkaloids and isomers, however they all tend to share the same basic structure (exemplified by the alkaloids from the venom of fire ant workers in Figure 5). Venom alkaloids in ants, particularly in those groups where they are predominant compounds, play a central role in their biology.

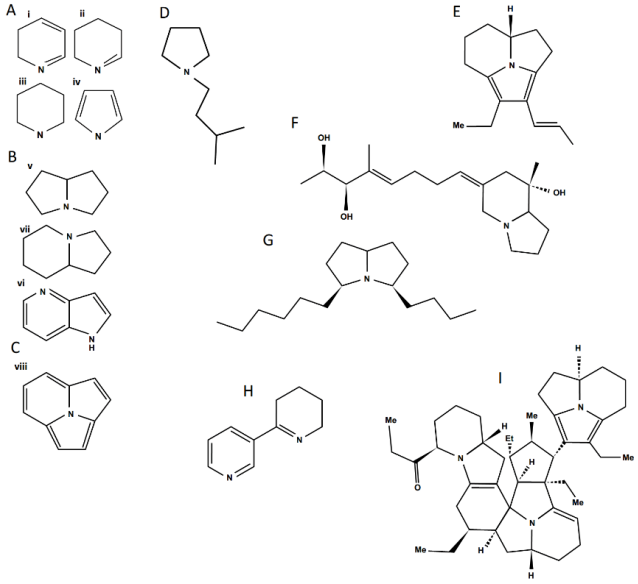


Figure 4. Structural diversity among ant venom alkaloids. A—monocyclic alkaloid families, i. Pyridine, ii. Piperideine, iii. Piperidine, iv. Pyrrole; B—bicyclic alkaloid families, v. Pyrrolizidine, vi. Indolizidine, vii. Pyrrolopyridine; C—trycliclic alkaloid family, viii. Pyrroloindolizidine; Examples of ant alkaloids: D—anabaseine; E—pumiliotoxin; F—monomorine; G—another *Monomorium* venom alkaloid; H—anabaseine; I—a complex myrmicarin.

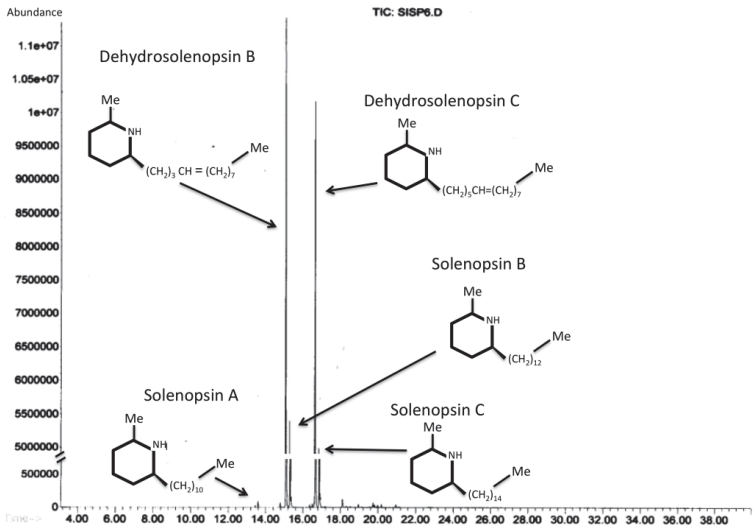


Figure 5. Examples of the most abundant venom alkaloids found in the venom of *Solenopsis invicta*. GC-MS chromatogram of hexane solvent in which fire ant workers were immersed.

5.1. Production of Alkaloids in Ant Venoms

As previously mentioned, most alkaloids are primarily described from plant extracts, illustrated by the well-studied compounds coniine and nicotine. Plants take advantage of secondary metabolites,

such as alkaloids, as a deterrent against herbivores. Therefore, it is often presumed that ants can sequester alkaloids into their venoms by feeding on plants [147], however, laboratory ant colonies can produce alkaloids in the absence of such foods, thus demonstrating they synthesize the compounds directly (personal observation by EGPF). This observation singles out ants, from other hymenopterans, for their capacity to produce copious amounts of bioactive alkaloids, with different groups of ants producing particular groups of alkaloids (Table 4). Venom alkaloids are believed to be produced inside the convoluted gland of the venom apparatus, as mentioned in anatomical studies of the venom apparatus of *Solenopsis* fire ants [148,149]. These authors even demonstrated local tissue damage possibly caused by synthesis and storage of such toxic compounds. Exactly how these ants produce the alkaloids is presently unknown, however, a few biochemical pathways have been proposed [150]. Interestingly it has been mentioned that among the transcripts of a fire ant venom gland, there are several enzymes related to the mevalonate pathway of synthesis of polyketides, which is the biochemical pathway attributed to this class of alkaloids [20]. It is also possible that a microbial symbiont may have been involved in the production of intermediary compounds, as was suggested for some alkaloids found in sponges [151] and briefly hinted at in ants by Saporito *et al.* [144].

To date, alkaloids are known to be prominent within the venom secretions of the subfamily Myrmicinae, particularly within the tribe Solenopsidini, which includes such genera as *Solenopsis*, *Monomorium*, *Allomerus*, and *Megalomyrmex*. These ants usually either infest the nests of other species (e.g., thief ants within *Solenopsis* and *Megalomyrmex*) or they are slow-moving hardy foragers such as the flower ants from the genus *Monomorium*. Alkaloids are usually bitter and frequently poisonous when ingested, thus these ants are granted protection against most potential predators. Moreover, venom alkaloids aid these ants to manipulate and avoid their host species and competitors.

The literature concerning the biological activities of alkaloidal compounds detected in ant venoms is fragmented. Ant alkaloids have only been investigated in a minority of ant species where they are associated with a range of biological activities, with different activities often reported for the same compounds across different ant groups. Nevertheless, several alkaloids detected in ants are shared with other distant organisms (e.g., anabaseine alkaloids are also found in nemertine worms, and tobacco plants), and a few have been synthesized for specific investigations (e.g., synthetic solenopsins for biomedical studies). In such instances, there are studies of their chemical and biological properties published in sources unrelated to myrmecology (see Anisewiski [127]). Non-toxic activities include those relating to communication and behavioral modulation including trail, alarm and sex pheromones and attractants [131,152] but these are outside of the scope of this review and will not be covered here. For a more complete overview on the biological activities of ant venom alkaloids, please refer to Jones & Blum 1983, Brossi 1987, Escoubas & Blum 1990, Anisewisky 2015 [127,153–155].

Alkaloids with toxic adverse effects on other organisms include: (i) herbicidal effects recorded from the alkaloid-rich venom of *Solenopsis* [153]; (ii) arthropod toxins that are used against competitors, predators, and prey either by spraying, injection or topical application, mainly recorded from *Monomorium*, *Solenopsis*, and *Tetraponera*; (iii) antimicrobials which remain to date poorly studied in ant venoms excepting studies with *Solenopsis* and some tests with *Monomorium*; and (iv) mammalian toxins, as demonstrated by tests on mammals and mammalian cells, reported mainly from *Solenopsis*, but also from *Monomorium* and *Tetraponera* [156]. Considering the disproportionate number of studies regarding the toxic effects of fire ant venom alkaloids, they are discussed in further detail below.

5.2. Solenopsins: A Case Study of Ant Venom Alkaloids

The chemistry and physiological effects of venom alkaloids have been best studied among the fire ants (*Solenopsis* spp.; Myrmicinae), a notable group of about 20 species [157]. Compared to other *Solenopsis* ants which typically behave as thief-ants, fire ants are much larger and faster, foraging in the open and using venom to subdue larger prey, defend their ground and food, and discourage predators. They are widely reputed for their aggressiveness combined with the burning sensation caused by their stings [157]. When injected into the skin their poorly-soluble venom alkaloids cause a

local inflammatory reaction, and can lead to pustule formation within hours. Due to some fire ants being regarded as one of the top-rated global invasive pests, there is a growing body of literature about their biology and venom alkaloids.

5.2.1. Solenopsin Chemistry

Alkaloids in fire ant venoms are mainly hydrophobic piperidines called solenopsins (generally similar in structure to coniine and nicotine [127]) and piperideines in much lesser amounts. These are oxygen-free polyketide alkaloids as their nitrogen atom is inserted into a polyketide carbon skeleton, and they are not metabolically derived from amino acids [127]. Structurally, these are compounds with a piperidinic ring, often unsaturated, attached to the side by a hydrocarbon chain of variable length [127] (Figure 2). The solenopsins come in many isomeric forms with slightly different chemical and biological properties for each configuration [158]. There are several piperideines found in trace amounts in the venom, which are currently thought to be unstable intermediates for the synthesis of solenopsins, but these remain largely unstudied [158].

The solenopsins can be easily extracted from *Solenopsis* ants either by directly dissecting the venom glands or by dipping the ants in organic solvents [21,159]. The extract can then be partially purified through traditional thin-layer or silica column chromatography based on the relative affinity of solenopsins for silicates and different solvents. Given this facile extraction procedure, copious amounts of venom can be obtained from whole nests (further details are given in Fox *et al*. [84]). Female individuals of any given caste will carry a unique mixture of solenopsins [130], as shown in Figure 3. Unfortunately, because of the shared chemical properties between the different isomers, complete purification of each of these compounds is currently not feasible [160]. Thus, until preparative purification methods to separate isomers are devised, the study of the biological and physiological effects of solenopsins depends either on the synthesis of each compound, or testing with natural mixed extracts (see Fox [20]).

5.2.2. Solenopsin Pharmacology

In general, solenopsins are regarded as the main toxic component of fire ant venoms. Apart from a burning sensation, oedema and pustule formation, they have been found to possess necrotic, haemolytic, antibiotic, and insecticidal activities [23,153]. Many alkaloids have antimicrobial properties to prevent infections from materials and prey brought into the colony. Solenopsins, in general, are no exception with potent antimicrobial activity against fungi and gram-positive bacteria [161,162], while solenopsin A was effective against gram-negative bacteria. The ants appear to employ this activity to disinfect their surroundings and brood by vigorously shaking their gaster and spreading the venom throughout the nest [163,164]. Also, solenopsins are effective as insect repellents and as insecticides, mainly against lepidopterans [165,166]. Such a property is invaluable to these ants since they are aggressive predators and competitors of other ants compared to thief ants which invade the nests of other ants to pillage resources and the brood.

In mammals, solenopsins were demonstrated to cause a number of complex physiological alterations, such as blockade of the neuromuscular junction [167], triggering histamine production in mastocytes [168], inhibiting ATP-dependent sodium-potassium pumps, and respiratory chains [169,170], activating platelets and neutrophils [171], and inhibiting neuronal nitric oxide synthase [172]. Following intravenous injection, synthetic isosolenopsin A and solenopsin A were capable of severely impairing both the central nervous system and cardiovascular systems of mice [173], which shows their capacity to cross the blood-brain barrier. Doses of 3–30 mg/kg were particularly toxic, causing a range of effects from dizziness, cardiorespiratory complications, seizures, and death [173]. These toxic effects are beneficial to these ants as active predators of both vertebrate and invertebrate prey, and also in defending their nests. Synthetic solenopsin A has been shown to possess a potent inhibitory activity against class-1 phosphatidylinositol-3-kinase signalling and

angiogenesis in mice embryos and zebra fish, making this alkaloid a potential lead therapeutic for the treatment of cancer [174].

6. Other Toxins

A set of additional small molecule compounds have been found in ant venoms. This includes alkylated pyrazines that are usually not considered alkaloids [127], being more common as mandibular gland secretions, identified as venom components in some ants (e.g., *Atta bisphaerica*) [152,175]. The venom glands of some species have also been reported to contain monoterpene hydrocarbons. For example, the primary venom compound of *Myrmicaria natalensis* is the cyclic terpene limonene [176], however, the venom also contains α-pinene, β-pinene, sabinene, terpinolene, β-myrcene, α-phallendrene, α-terpinene, and caphene [177]. All of these compounds can be highly toxic to insects minding the fact that this ant species preys on termites. The presence of immunologically active heterogeneous polyanonic polysaccharides have also been reported in the venoms of *Pseudomyrmex* spp. [178]. There are also non-alkaloidal amines such as pteridines that have been identified from some ants including *Formica* and *Lasius*, actidines found in *Megaponera* and *Dorymyrmex*, and histamine which is abundant in venoms of *Myrmecia* spp. [131,179].

Finally, it should be mentioned that the secreted venom of ants is even more complex due to interactions with secretions from their Dufour's gland. The Dufour's gland is an accessory organ attached to the venom gland and considered part of the venom apparatus [180]. Its secretions may act synergistically with toxins originating from the venom gland and contribute to the toxicity of the secreted venom. For example, the Dufour's gland of *Crematogaster scutellaris* produces long-chain primary acetates (non-toxic) which are converted to highly electrophilic aldehydes (toxic) by enzymes from the venom gland during the venom secretion [181].

7. Conclusions

A host of recent studies have revealed that ant venoms are more complex and heterogeneous than initially thought, owing in particular to the newly uncovered complexity of their peptidome and proteome contents. Although the extant biodiversity of ant venoms remains largely unexplored, their high plasticity is suggested by recent studies of a variety of subfamilies or genera, showing highly different venom compositions. Formicinae ants and some Myrmicinae genera essentially produce non-proteinaceous venoms primarily composed of formic acid and alkaloids, respectively. In contrast, most ant species from other subfamilies have retained the ability to sting and, in turn, produce peptide- and protein-rich venoms. Evolution has therefore led to some ants abandoning their ability to sting, unlike wasps and most Apidae, whilst evolving a different set of chemical defenses along with modified predatory and defensive behaviours. To date, the molecular and structural complexity of ant venoms has barely been explored but the broad ecological diversity of ants strongly suggests that further structural peptide and protein diversity might be uncovered by extensive biochemical studies, leading to discovery of a much broader range of toxins than currently observed. Technological progress, particularly in deep-sequencing approaches, coupled with high-end transcriptomic, peptidomic, and enzymatic methods based on mass spectrometry and peptide *de novo* sequencing, will quickly allow for the detection and characterization of numerous novel peptides and enzymes, at sensitivity levels and a depth not previously attained. Despite the biochemical diversity potentially present in ant venoms, this review highlights the paucity of knowledge on the molecular pharmacology of most ant toxins. The biological function and mechanism of action of the majority of ant venom toxins described to date remain poorly characterized or simply unstudied. Thorough exploration of a broader taxonomic diversity will likely result in the discovery of novel bioactive toxins which may become useful tools for biopesticide or drug development and will shed insights on the molecular evolution of venom in Hymenoptera, the roles of venom in ant biology, the genetic makeup leading to ant venom diversification and its role in the successful conquest of almost all ecological niches by ants. The characterization of functional roles and pharmacological properties of this vast array of novel toxins

(particularly peptides) will certainly become one of the most functional and significant endeavors in future ant venom research, with a high application potential. Thus, the whole world of ant venom is still a vast uncharted scientific territory awaiting our attention.

Author Contributions: Axel Touchard, Samira R. Aili and Eduardo Gonçalves Paterson Fox conceived and wrote the review; Graham M. Nicholson, Pierre Escoubas, Jérôme Orivel and Alain Dejean reviewed the manuscript and provided valuable comments to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol. Evol.* **2013**, *28*, 219–229. [CrossRef] [PubMed]
2. Van Emden, H.F. Subclass pterygota, division endopterygota, order Hymenoptera (Sawflies, Ants, Bees and Wasps) 120,000 described species. In *Handbook of Agricultural Entomology*; John Wiley & Sons: Oxford, UK, 2013; pp. 193–220.
3. Robertson, P.L. A morphological and functional study of the venom apparatus in representatives of some major groups of Hymenoptera. *Aust J. Zool.* **1968**, *16*, 133–166. [CrossRef]
4. Hölldobler, B.; Wilson, E.-O. *The Ants*; Harvard University Press: Cambridge, MA, USA, 1990.
5. Wilson, E.O. *Success and Dominance in Ecosystems: The Case of the Social Insects*; Ecology Institute: Oldendorf/Luhe, Germany, 1990.
6. Bolton, B. An online catalog of the ants of the world. Available online: http://antcat.org (accessed on 1 October 2015).
7. Brady, S.; Fisher, B.; Schultz, T.; Ward, P. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. *BMC Evol. Biol.* **2014**, *14*. [CrossRef] [PubMed]
8. Ward, P.S. Taxonomy, phylogenetics and evolution. In *Ant ecology*; Lach, L., Parr, C.L., Abott, K.L., Eds.; Oxford University Press: Oxford, UK, 2010.
9. Ward, P.S. Phylogeny, classification, and species-level taxonomy of ants (Hymenoptera: formicidae). *Zootaxa* **2007**, *1668*, 549–563.
10. Cerdá, X.; Dejean, A. Predation by ants on arthropods and other animals. In *Predation in the Hymenoptera: an evolutionary perspective*; Polidori, C., Ed.; TransWorld Research Network: Kerala, India, 2011; pp. 39–78.
11. Orivel, J.; Dejean, A. Comparative effect of the venoms of ants of the genus *Pachycondyla* (Hymenoptera: Ponerinae). *Toxicon* **2001**, *39*, 195–201. [CrossRef]
12. Orivel, J.; Redeker, V.; Le Caer, J.P.; Krier, F.; Revol-Junelles, A.M.; Longeon, A.; Chaffotte, A.; Dejean, A.; Rossier, J. Ponericins, new antibacterial and insecticidal peptides from the venom of the ant *Pachycondyla goeldii*. *J. Biol. Chem.* **2001**, *276*, 17823–17829. [CrossRef] [PubMed]
13. Schmidt, J.O. Biochemistry of insect venoms. *Annu. Rev. Entomol.* **1982**, *27*, 339–368. [CrossRef] [PubMed]
14. Frederickson, M.E.; Gordon, D.M. The devil to pay: A cost of mutualism with *Myrmelachista schumanni* ants in 'devil's gardens' is increased herbivory on *Duroia hirsuta* trees. *Proc. Roy. Soc. Lond. B* **2007**, *274*, 1117–1123. [CrossRef] [PubMed]
15. Morgan, E.D.; Jungnickel, H.; Keegans, S.J.; do Nascimento, R.R.; Billen, J.; Gobin, B.; Ito, F. Comparative survey of abdominal gland secretions of the ant subfamily Ponerinae. *J. Chem. Ecol.* **2003**, *29*, 95–114. [CrossRef] [PubMed]
16. Schmidt, J.O. Chemistry, pharmacology and chemical ecology of ant venoms. In *Venoms of the Hymenoptera: Biochemical, Pharmacological and Behavioural Aspects*; Piek, T., Ed.; Academic Press: London, UK, 1986; pp. 425–508.
17. Brand, J.M. Fire ant venom alkaloids: Their contribution to chemosystematics and biochemical evolution. *Biochem. Syst. Ecol.* **1978**, *6*, 337–340. [CrossRef]
18. Schmidt, J.O.; Blum, M.S. The biochemical constituents of the venom of the harvester ant, *Pogonomyrmex badius*. *Comp. Biochem. Physiol. C* **1978**, *61C*, 239–247. [CrossRef]
19. Touchard, A.; Dauvois, M.; Arguel, M.-J.; Petitclerc, F.; Leblanc, M.; Dejean, A.; Orivel, J.; Nicholson, G.M.; Escoubas, P. Elucidation of the unexplored biodiversity of ant venom peptidomes via MALDI-TOF mass spectrometry and its application for chemotaxonomy. *J. Proteomics* **2014**, *105*, 217–231. [CrossRef] [PubMed]

20. Fox, E.G.P. Venom toxins of fire ants. In *Venom genomics and proteomics*; Gopalakrishnakone, P., Calvete, J.J., Eds.; Springer: Dordrecht, The Netherlands, 2014; pp. 1–16.
21. Fox, E.G.P.; Pianaro, A.; Solis, D.R.; Delabie, J.H.C.; Vairo, B.C.; Machado, E.d.A.; Bueno, O.C. Intraspecific and intracolonial variation in the profile of venom alkaloids and cuticular hydrocarbons of the fire ant *Solenopsis saevissima* Smith (Hymenoptera: Formicidae). *Psyche: Psyche: J. Entomol.* **2012**, *2012*, 1–10.
22. Schmidt, J.O. Ant venoms: A study of venom diversity. In *Pesticide and Venom Neurotoxicity*; Shankland, D.L., Hollingworth, R.M., Smyth, T., Jr., Eds.; Springer: New York, NY, USA, 1978; pp. 247–263.
23. Attygalle, A.B.; Morgan, E.D. Chemicals from the glands of ants. *Chem. Soc. Rev.* **1984**, *13*, 245–278. [CrossRef]
24. Billen, J.; Gobin, B. Trail following in army ants (Hymenoptera, Formicidae). *Neth. J. Zool.* **1996**, *46*, 272–280. [CrossRef]
25. Tragust, S.; Mitteregger, B.; Barone, V.; Konrad, M.; Ugelvig, L.V.; Cremer, S. Ants disinfect fungus-exposed brood by oral uptake and spread of their poison. *Curr. Biol.* **2013**, *23*, 76–82. [CrossRef] [PubMed]
26. Beard, R.L. Insect toxins and venoms. *Annu. Rev. Entomol.* **1963**, *8*, 1–18. [CrossRef] [PubMed]
27. Hefetz, A.; Blum, M.S. Biosynthesis of formic acid by the poison glands of formicine ants. *Biochim. Biophys. Acta* **1978**, *543*, 484–496. [CrossRef]
28. LeBrun, E.G.; Jones, N.T.; Gilbert, L.E. Chemical warfare among invaders: A detoxification interaction facilitates an ant invasion. *Science* **2014**, *343*, 1014–1017. [CrossRef] [PubMed]
29. Aili, S.R.; Touchard, A.; Escoubas, P.; Padula, M.P.; Orivel, J.; Dejean, A.; Nicholson, G.M. Diversity of peptide toxins from stinging ant venoms. *Toxicon* **2014**, *92*, 166–178. [CrossRef] [PubMed]
30. Brady, S.G.; Schultz, T.R.; Fisher, B.L.; Ward, P.S. Evaluating alternative hypotheses for the early evolution and diversification of ants. *Proc. Natl. Acad. Sci.* **2006**, *103*, 18172–18177. [CrossRef] [PubMed]
31. Frederickson, M.E.; Greene, M.J.; Gordon, D.M. Ecology: 'Devil's gardens' bedevilled by ants. *Nature* **2005**, *437*, 495–496. [CrossRef] [PubMed]
32. King, G.F.; Hardy, M.C. Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. *Annu. Rev. Entomol.* **2013**, *58*, 475–496. [CrossRef] [PubMed]
33. Possani, L.V.D.; Rodríguez de la Vega, R. Scorpion venom peptides. In *Handbook of Biologically Active Peptides*; Kastin, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2006; pp. 339–354.
34. Rodríguez de la Vega, R.C.; Schwartz, E.F.; Possani, L.D. Mining on scorpion venom biodiversity. *Toxicon* **2010**, *56*, 1155–1161. [CrossRef] [PubMed]
35. Terlau, H.; Olivera, B.M. Conus venoms: A rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **2004**, *84*, 41–68. [CrossRef] [PubMed]
36. Touchard, A.; Koh, J.M.S.; Aili, S.R.; Dejean, A.; Nicholson, G.M.; Orivel, J.; Escoubas, P. The complexity and structural diversity of ant venom peptidomes is revealed by mass spectrometry profiling. *Rapid Commun. Mass Spectrom.* **2015**, *29*, 385–396. [CrossRef] [PubMed]
37. Cologna, C.T.; Cardoso, J.D.S.; Jourdan, E.; Degueldre, M.; Upert, G.; Gilles, N.; Uetanabaro, A.P.T.; Costa Neto, E.M.; Thonart, P.; de Pauw, E. Peptidomic comparison and characterization of the major components of the venom of the giant ant *Dinoponera quadriceps* collected in four different areas of Brazil. *J. Proteomics* **2013**, *94*, 413–422. [CrossRef] [PubMed]
38. Orivel, J. L'adaptation à la vie arboricole de la fourmi *Pachycondyla goeldii* (Hymenoptera: ponerinae). Ph.D. Thesis, Université de Paris XIII, Paris, France, 2000. Available online: http://cat.inist.fr/?aModele=afficheN&cpsidt=200149 (accessed online: 19 January 2016) . (In French).
39. Johnson, S.R.; Copello, J.A.; Evans, M.S.; Suarez, A.V. A biochemical characterization of the major peptides from the venom of the giant Neotropical hunting ant *Dinoponera australis*. *Toxicon* **2010**, *55*, 702–710. [CrossRef] [PubMed]
40. Rifflet, A.; Gavalda, S.; Téné, N.; Orivel, J.; Leprince, J.; Guilhaudis, L.; Génin, E.; Vétillard, A.; Treilhou, M. Identification and characterization of a novel antimicrobial peptide from the venom of the ant *Tetramorium bicarinatum*. *Peptides* **2012**, *38*, 363–370. [CrossRef] [PubMed]
41. Inagaki, H.; Akagi, M.; Imai, H.T.; Taylor, R.W.; Kubo, T. 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.* **2004**, *428*, 170–178. [CrossRef] [PubMed]

42. Inagaki, H.; Akagi, M.; Imai, H.T.; Taylor, R.W.; Wiese, M.D.; Davies, N.W.; Kubo, T. Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (Jack Jumper Ant). *Arch. Biochem. Biophys.* **2008**, *477*, 411–416. [CrossRef] [PubMed]
43. Wiese, M.D.; Chataway, T.K.; Davies, N.W.; Milne, R.W.; Brown, S.G.; Gai, W.P.; Heddle, R.J. Proteomic analysis of *Myrmecia pilosula* (jack jumper) ant venom. *Toxicon* **2006**, *47*, 208–217. [CrossRef] [PubMed]
44. Wanandy, T.; Gueven, N.; Davies, N.W.; Brown, S.G.A.; Wiese, M.D. Pilosulins: A review of the structure and mode of action of venom peptides from an Australian ant *Myrmecia pilosula*. *Toxicon* **2015**, *98*, 54–61. [CrossRef] [PubMed]
45. Pluzhnikov, K.A.; Kozlov, S.A.; Vassilevski, A.A.; Vorontsova, O.V.; Feofanov, A.V.; Grishin, E.V. Linear antimicrobial peptides from *Ectatomma quadridens* ant venom. *Biochimie* **2014**, *107*, 211–215. [CrossRef] [PubMed]
46. King, M.A.; Wu, Q.X.; Donovan, G.R.; Baldo, B.A. Flow cytometric analysis of cell killing by the jumper ant venom peptide pilosulin 1. *Cytometry* **1998**, *32*, 268–273. [CrossRef]
47. Wu, Q.X.; King, M.A.; Donovan, G.R.; Alewood, D.; Alewood, P.; Sawyer, W.H.; Baldo, B.A. Cytotoxicity of pilosulin 1, a peptide from the venom of the jumper ant *Myrmecia pilosula*. *Biochim. Biophys. Acta* **1998**, *1425*, 74–80. [CrossRef]
48. Kuhn-Nentwig, L. Antimicrobial and cytolytic peptides of venomous arthropods. *Cell. Mol. Life Sci.* **2003**, *60*, 2651–2668. [CrossRef] [PubMed]
49. Zeng, X.-C.; Wang, S.-X.; Zhu, Y.; Zhu, S.-Y.; Li, W.-X. Identification and functional characterization of novel scorpion venom peptides with no disulfide bridge from *Buthus martensii* Karsch. *Peptides* **2004**, *25*, 143–150. [CrossRef]
50. Kozlov, S.A.; Vassilevski, A.A.; Feofanov, A.V.; Surovoy, A.Y.; Karpunin, D.V.; Grishin, E.V. Latarcins, antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaevi* (Zodariidae) thatexemplify biomolecular diversity. *J. Biol. Chem.* **2006**, *281*, 20983–20992. [CrossRef] [PubMed]
51. Kuzmenkov, A.I.; Fedorova, I.M.; Vassilevski, A.A.; Grishin, E.V. Cysteine-rich toxins from *Lachesana tarabaevi* spider venom with amphiphilic *C*-terminal segments. *Biochim. Biophys. Acta* **2013**, *1828*, 724–731. [CrossRef] [PubMed]
52. Vassilevski, A.; Kozlov, S.; Samsonova, O.; Egorova, N.; Karpunin, D.; Pluzhnikov, K.; Feofanov, A.; Grishin, E. Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. *Biochem. J.* **2008**, *411*, 687–696. [CrossRef] [PubMed]
53. Cremer, S.; Armitage, S.A.O.; Schmid-Hempel, P. Social immunity. *Curr. Biol.* **2007**, *17*, R693–R702. [CrossRef] [PubMed]
54. Turillazzi, S.; Mastrobuoni, G.; Dani, F.R.; Moneti, G.; Pieraccini, G.; la Marca, G.; Bartolucci, G.; Perito, B.; Lambardi, D.; Cavallini, V.; *et al.* Dominulin A and B: Two new antibacterial peptides identified on the cuticle and in the venom of the social paper wasp *Polistes dominulus* using MALDI-TOF, MALDI-TOF/TOF, and ESI-ion trap. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 376–383. [CrossRef] [PubMed]
55. Nicholson, G.M. Spider venom peptides. In *Handbook of biologically active peptides*; Kastin, A., Ed.; Elsevier: San Diego, CA, USA, 2006; pp. 461–472.
56. Maschwitz, U.; Hahn, M.; Schönegge, P. Paralysis of prey in ponerine ants. *Naturwissenschaften* **1979**, *66*, 213–214. [CrossRef]
57. Piek, T.; Duval, A.; Hue, B.; Karst, H.; Lapied, B.; Mantel, P.; Nakajima, T.; Pelhate, M.; Schmidt, J.O. Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata*. *Comp. Biochem. Physiol. B.-Biochem. Mol. Biol.* **1991**, *99*, 487–495. [CrossRef]
58. Piek, T.; Hue, B.; Mantel, P.; Nakajima, T.; Schmidt, J.O. Pharmacological characterization and chemical fractionation of the venom of the ponerine ant, *Paraponera clavata*. *Comp. Biochem. Physiol. B.-Biochem. Mol. Biol.* **1991**, *99*, 481–486. [CrossRef]
59. Duval, A.; Malecot, C.O.; Pelhate, M.; Piek, T. Poneratoxin, a new toxin from an ant venom, reveals an interconversion between two gating modes of the Na channels in frog skeletal muscle fibres. *Pflugers Arch.* **1992**, *420*, 239–247. [CrossRef] [PubMed]
60. Hendrich, A.B.; Mozrzymas, J.W.; Konopinska, D.; Scuka, M. The effect of poneratoxin on neuromuscular transmission in the rat diaphragm. *Cell. Mol. Biol. Lett.* **2002**, *7*, 195–202. [PubMed]

61. Szolajska, E.; Poznanski, J.; Ferber, M.L.; Michalik, J.; Gout, E.; Fender, P.; Bailly, I.; Dublet, B.; Chroboczek, J. Poneratoxin, a neurotoxin from ant venom. Structure and expression in insect cells and construction of a bio-insecticide. *Eur. J. Biochem.* **2004**, *271*, 2127–2136. [CrossRef] [PubMed]
62. Pluzhnikov, K.A.; Nol'de, D.E.; Tertyshnikova, S.M.; Sukhanov, S.V.; Sobol, A.G.; Torgov, M.; Filippov, A.K.; Arsen'ev, A.S.; Grishin, E.V. Structure activity study of the basic toxic component of venom from the ant *Ectatomma tuberculatum*. *Bioorg. Khim.* **1994**, *20*, 857–871.
63. Arseniev, A.S.; Pluzhnikov, K.A.; Nolde, D.E.; Sobol, A.G.; Torgov, M.; Sukhanov, S.V.; Grishin, E.V. Toxic principle of selva ant venom is a pore-forming protein transformer. *FEBS Lett.* **1994**, *347*, 112–116. [CrossRef]
64. Pluzhnikov, K.; Nosyreva, E.; Shevchenko, L.; Kokoz, Y.; Schmalz, D.; Hucho, F.; Grishin, E. Analysis of ectatomin action on cell membranes. *Eur. J. Biochem.* **1999**, *262*, 501–506. [CrossRef] [PubMed]
65. Pan, J.; Hink, W.F. Isolation and characterization of myrmexins, six isoforms of venom proteins with anti-inflammatory activity from the tropical ant, *Pseudomyrmex triplarinus*. *Toxicon* **2000**, *38*, 1403–1413. [CrossRef]
66. Touchard, A.; Labrière, N.; Roux, O.; Petitclerc, F.; Orivel, J.; Escoubas, P.; Koh, J.; Nicholson, G.M.; Dejean, A. Venom toxicity and composition in three *Pseudomyrmex* ant species having different nesting modes. *Toxicon* **2014**, *88*, 67–76. [CrossRef] [PubMed]
67. Craik, D.J.; Daly, N.L.; Waine, C. The cystine knot motif in toxins and implications for drug design. *Toxicon* **2001**, *39*, 43–60. [CrossRef]
68. Herzig, V.; King, G.F. The cystine knot is responsible for the exceptional stability of the insecticidal spider toxin ω-Hexatoxin-Hv1a. *Toxins* **2015**, *7*, 4366–4380. [CrossRef] [PubMed]
69. Torres, A.F.C.; Huang, C.; Chong, C.M.; Leung, S.W.; Prieto-da-Silva, Á.R.B.; Havt, A.; Quinet, Y.P.; Martins, A.M.C.; Lee, S.M.Y.; Rádis-Baptista, G. Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: Insights into the polypeptide toxin arsenal of hymenopterans. *PLoS One* **2014**, *9*. [CrossRef]
70. Lavergne, V.; Harliwong, I.; Jones, A.; Miller, D.; Taft, R.J.; Alewood, P.F. Optimized deep-targeted proteotranscriptomic profiling reveals unexplored *Conus* toxin diversity and novel cysteine frameworks. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3782–E3791. [CrossRef] [PubMed]
71. Escoubas, P.; Quinton, L.; Nicholson, G.M. Venomics: Unravelling the complexity of animal venoms with mass spectrometry. *J. Mass Spectrom.* **2008**, *43*, 279–295. [CrossRef] [PubMed]
72. Von Reumont, B.; Campbell, L.; Jenner, R. Quo Vadis Venomics? A roadmap to neglected venomous invertebrates. *Toxins* **2014**, *6*, 3488–3551. [CrossRef] [PubMed]
73. Donovan, G.R.; Street, M.D.; Baldo, B.A.; Alewood, D.; Alewood, P.; Sutherland, S. Identification of an IgE-binding determinant of the major allergen *Myr pI* from the venom of the Australian jumper ant Myrmecia pilosula. *Biochim. Biophys. Acta* **1994**, *1204*, 48–52. [CrossRef]
74. Wiese, M.D.; Brown, S.G.A.; Chataway, T.K.; Davies, N.W.; Milne, R.W.; Aulfrey, S.J.; Heddle, R.J. *Myrmecia pilosula* (Jack Jumper) ant venom: Identification of allergens and revised nomenclature. *Allergy* **2007**, *62*, 437–443. [CrossRef] [PubMed]
75. King, G.F.; Gentz, M.C.; Escoubas, P.; Nicholson, G.M. A rational nomenclature for naming peptide toxins from spiders and other venomous animals. *Toxicon* **2008**, *52*, 264–276. [CrossRef] [PubMed]
76. Undheim, E.A.; Jones, A.; Clauser, K.R.; Holland, J.W.; Pineda, S.S.; King, G.F.; Fry, B.G. Clawing through evolution: Toxin diversification and convergence in the ancient lineage Chilopoda (Centipedes). *Mol. Biol. Evol.* **2014**. [CrossRef] [PubMed]
77. Oliveira, J.S.; Fuentes-Silva, D.; King, G.F. Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. *Toxicon* **2012**, *60*, 539–550. [CrossRef] [PubMed]
78. Pluzhnikov, K.; Shevchenko, L.; Grishin, E. Ant polypeptide toxins. In *Methods and tools in biosciences and medicine: Animal toxins*; Rochat, H., Martin-Eauclaire, M.-F., Eds.; Birkhäuser: Basel, Switzerland, 2000; pp. 90–98.
79. Davies, N.W.; Wiese, M.D.; Brown, S.G. Characterisation of major peptides in 'jack jumper' ant venom by mass spectrometry. *Toxicon* **2004**, *43*, 173–183. [CrossRef] [PubMed]
80. Donovan, G.R.; Street, M.D.; Baldo, B.A. Separation of jumper ant (*Myrmecia pilosula*) venom allergens: A novel group of highly basic proteins. *Electrophoresis* **1995**, *16*, 804–810. [CrossRef] [PubMed]
81. Leluk, J.; Schmidt, J.; Jones, D. Comparative studies on the protein composition of hymenopteran venom reservoirs. *Toxicon* **1989**, *27*, 105–114. [CrossRef]

82. De Lima, P.R.; Brochetto-Braga, M.R. Hymenoptera venom review focusing on *Apis mellifera*. *J. Venom. Anim. Toxins* **2003**, *9*, 149–162.
83. dos Santos Pinto, J.R.A.; Fox, E.G.P.; Saidemberg, D.M.; Santos, L.D.; da Silva Menegasso, A.R.; Costa-Manso, E.; Machado, E.A.; Bueno, O.C.; Palma, M.S. Proteomic view of the venom from the fire ant *Solenopsis invicta* Buren. *J. Proteome Res.* **2012**, *11*, 4643–4653. [CrossRef] [PubMed]
84. Fox, E.G.P.; Solis, D.R.; dos Santos, L.D.; dos Santos Pinto, J.R.A.; da Silva Menegasso, A.R.; Silva, R.C.M.C.; Palma, M.S.; Bueno, O.C.; de Alcântara Machado, E. A simple, rapid method for the extraction of whole fire ant venom (Insecta: Formicidae: *Solenopsis*). *Toxicon* **2013**, *65*, 5–8.
85. Bouzid, W.; Klopp, C.; Verdenaud, M.; Ducancel, F.; Vétillard, A. Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (Hymenoptera: Formicidae): The first transcriptome analysis of an ant species. *Toxicon* **2013**, *70*, 70–81. [CrossRef] [PubMed]
86. Baer, H.; Liu, T.Y.; Anderson, M.C.; Blum, M.; Schmid, W.H.; James, F.J. Protein components of fire ant venom (*Solenopsis invicta*). *Toxicon* **1979**, *17*, 397–405. [CrossRef]
87. Torres, A.; Quinet, Y.; Havt, A.; Rádis-Baptista, G.; Martins, A. Molecular pharmacology and toxynology of venom from ants. In *An Integrated View of the Molecular Recognition and Toxinology—From Analytical Procedures to Biomedical Applications*; InTech: Rijeka, Croatia, 2013; pp. 207–222.
88. dos Santos, L.D.; da Silva Menegasso, A.R.; dos Santos Pinto, J.R.; Santos, K.S.; Castro, F.M.; Kalil, J.E.; Palma, M.S. Proteomic characterization of the multiple forms of the PLAs from the venom of the social wasp *Polybia paulista*. *Proteomics* **2011**, *11*, 1403–1412. [CrossRef] [PubMed]
89. von Sicard, N.A.; Candy, D.J.; Anderson, M. The biochemical composition of venom from the pavement ant (*Tetramorium caespitum* L.). *Toxicon* **1989**, *27*, 1127–1133. [CrossRef]
90. Hink, W.F.; Pappas, P.W.; Jaworski, D.C. Partial biochemical characterization of venom from the ant, *Pseudomyrmex triplarinus*. *Toxicon* **1994**, *32*, 763–772. [CrossRef]
91. Bouzid, W.; Verdenaud, M.; Klopp, C.; Ducancel, F.; Noirot, C.; Vétillard, A. De Novo sequencing and transcriptome analysis for *Tetramorium bicarinatum*: A comprehensive venom gland transcriptome analysis from an ant species. *BMC Genomics* **2014**, *15*, 987. [CrossRef] [PubMed]
92. Richardson, M.; Pimenta, A.M.C.; Bemquerer, M.P.; Santoro, M.M.; Beirao, P.S.L.; Lima, M.E.; Figueiredo, S.G.; Bloch, C., Jr.; Vasconcelos, E.A.R.; Campos, F.A.P.; *et al*. Comparison of the partial proteomes of the venoms of Brazilian spiders of the genus *Phoneutria*. *Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol.* **2006**, *142*, 173–187. [CrossRef] [PubMed]
93. Batista, C.V.F.; Zamudio, F.Z.; Lucas, S.; Fox, J.W.; Frau, A.; Prestipino, G.; Possani, L.D. Scorpion toxins from *Tityus cambridgei* that affect Na^+-channels. *Toxicon* **2002**, *40*, 557–562. [CrossRef]
94. Gutierrez Mdel, C.; Abarca, C.; Possani, L.D. A toxic fraction from scolopendra venom increases the basal release of neurotransmitters in the ventral ganglia of crustaceans. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2003**, *135*, 205–214. [CrossRef]
95. Rates, B.; Bemquerer, M.P.; Richardson, M.; Borges, M.H.; Morales, R.A.V.; De Lima, M.E.; Pimenta, A.M.C. Venomic analyses of *Scolopendra viridicornis nigra* and *Scolopendra angulata* (Centipede, Scolopendromorpha): Shedding light on venoms from a neglected group. *Toxicon* **2007**, *49*, 810–826. [CrossRef] [PubMed]
96. Zalat, S.; Schmidt, J.; Moawad, T.I. Lipase and phospholipase activities of Hymenoptera venoms (wasps and ants). *Egyptian J. Biol.* **2003**, *5*, 138–147.
97. Ford, S.A.; Baldo, B.A.; Weiner, J.; Sutherland, S. Identification of jack-jumper ant (*Myrmecia pilosula*) venom allergens. *Clin. Exp. Allergy* **1991**, *21*, 167–171. [CrossRef] [PubMed]
98. Hoffman, D.R.; Dove, D.E.; Jacobson, R.S. Allergens in Hymenoptera venom: XX. Isolation of four allergens from imported fire ant (*Solenopsis invicta*) venom. *J. Allergy Clin. Immunol.* **1988**, *82*, 818–827. [CrossRef]
99. Hoffman, D.R.; Sakell, R.H.; Schmidt, M. Sol i 1, the phospholipase allergen of imported fire ant venom. *J. Allergy Clin. Immunol.* **2005**, *115*, 611–616. [CrossRef] [PubMed]
100. Hoffman, D.R. Ant venoms. *Curr. Opin. Allergy Clin. Immunol.* **2010**, *10*, 342–346. [CrossRef] [PubMed]
101. Marcon, F.; Purtell, L.; Santos, J.; Hains, P.G.; Escoubas, P.; Graudins, A.; Nicholson, G.M. Characterization of monomeric and multimeric snake neurotoxins and other bioactive proteins from the venom of the lethal Australian common copperhead (*Austrelaps superbus*). *Biochem. Pharmacol.* **2013**, *85*, 1555–1573. [CrossRef] [PubMed]
102. Schmidt, J.O.; Blum, M.S.; Overal, W.L. Comparative enzymology of venoms from stinging Hymenoptera. *Toxicon* **1986**, *24*, 907–921. [CrossRef]

103. Kuhn-Nentwig, L.; Schaller, J.; Nentwig, W. Biochemistry, toxicology and ecology of the venom of the spider *Cupiennius salei* (Ctenidae). *Toxicon* **2004**, *43*, 543–553. [CrossRef] [PubMed]
104. Wanstall, J.C.; De la Lande, I. Fractionation of bulldog ant venom. *Toxicon* **1974**, *12*. [CrossRef]
105. Cui, F.; Lin, Z.; Wang, H.; Liu, S.; Chang, H.; Reeck, G.; Qiao, C.; Raymond, M.; Kang, L. Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. *Insect Biochem. Mol. Biol.* **2011**, *41*, 1–8. [CrossRef] [PubMed]
106. Bonasio, R.; Zhang, G.; Ye, C.; Mutti, N.S.; Fang, X.; Qin, N.; Donahue, G.; Yang, P.; Li, Q.; Li, C.; *et al*. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **2010**, *329*, 1068–1071. [CrossRef] [PubMed]
107. Mamillapalli, R.; Haimovitz, R.; Ohad, M.; Shinitzky, M. Enhancement and inhibition of snake venom phosphodiesterase activity by lysophospholipids. *FEBS Lett.* **1998**, *436*, 256–258. [CrossRef]
108. Parkinson, N.; Smith, I.; Weaver, R.; Edwards, J.P. A new form of arthropod phenoloxidase is abundant in venom of the parasitoid wasp *Pimpla hypochondriaca*. *Insect Biochem. Mol. Biol.* **2001**, *31*, 57–63. [CrossRef]
109. Danneels, E.L.; Rivers, D.B.; De Graaf, D.C. Venom proteins of the parasitoid wasp *Nasonia vitripennis*: recent discovery of an untapped pharmacopee. *Toxins* **2010**, *2*, 494–516. [CrossRef] [PubMed]
110. Peiren, N.; de Graaf, D.C.; Vanrobaeys, F.; Danneels, E.L.; Devreese, B.; Van Beeumen, J.; Jacobs, F.J. Proteomic analysis of the honey bee worker venom gland focusing on the mechanisms of protection against tissue damage. *Toxicon* **2008**, *52*, 72–83. [CrossRef] [PubMed]
111. Resende, V.M.F.; Vasilj, A.; Santos, K.S.; Palma, M.S.; Shevchenko, A. Proteome and phosphoproteome of Africanized and European honeybee venoms. *Proteomics* **2013**, *13*, 2638–2648. [CrossRef] [PubMed]
112. Lewis, J.C.; Day, A.J.; De la Lande, I.S. Phospholipase A in the venom of the Australian bulldog ant *Myrmecia pyriformis*. *Toxicon* **1968**, *6*, 109–112. [CrossRef]
113. Magalhaes, G.S.; Caporrino, M.C.; Della-Casa, M.S.; Kimura, L.F.; Prezotto-Neto, J.P.; Fukuda, D.A.; Portes-Junior, J.A.; Neves-Ferreira, A.G.; Santoro, M.L.; Barbaro, K.C. Cloning, expression and characterization of a phospholipase D from *Loxosceles gaucho* venom gland. *Biochimie* **2013**, *95*, 1773–1783. [CrossRef] [PubMed]
114. Ramos-Cerrillo, B.; Olvera, A.; Odell, G.V.; Zamudio, F.; Paniagua-Solís, J.; Alagón, A.; Stock, R.P. Genetic and enzymatic characterization of sphingomyelinase D isoforms from the North American fiddleback spiders *Loxosceles boneti* and *Loxosceles reclusa*. *Toxicon* **2004**, *44*, 507–514. [CrossRef] [PubMed]
115. Angulo, Y.; Olamendi-Portugal, T.; Possani, L.D.; Lomonte, B. Isolation and characterization of myotoxin II from *Atropoides* (*Bothrops*) *nummifer* snake venom, a new Lys49 phospholipase A2 homologue. *Int J. Biochem. Cell. Biol.* **2000**, *32*, 63–71. [CrossRef]
116. Nagai, H.; Oshiro, N.; Takuwa-Kuroda, K.; Iwanaga, S.; Nozaki, M.; Nakajima, T. A new polypeptide toxin from the nematocyst venom of an Okinawan sea anemone *Phyllodiscus semoni* (Japanese name "unbachi-isoginchaku"). *Biosci. Biotechnol. Biochem.* **2002**, *66*, 2621–2625. [CrossRef] [PubMed]
117. Alam, M.; Gomes, A. Viper venom-induced inflammation and inhibition of free radical formation by pure compound (2-hydroxy-4-methoxy benzoic acid) isolated and purified from anantamul (*Hemidesmus indicus* R.BR) root extract. *Toxicon* **1998**, *36*, 207–215. [CrossRef]
118. Bull, H.; Murray, P.G.; Thomas, D.; Fraser, A.; Nelson, P.N. Acid phosphatases. *Mol. Pathol.* **2002**, *55*, 65–72. [CrossRef] [PubMed]
119. Hoffman, D.R. Allergens in Hymenoptera venom XXIV: The amino acid sequences of imported fire ant venom allergens Sol i II, Sol i III, and Sol i IV. *J. Allergy Clin. Immunol.* **1993**, *91*, 71–78. [CrossRef]
120. Hoffman, D.R. Hymenoptera venom allergens. *Clin. Rev. Allergy Immunol.* **2006**, *30*, 109–128. [CrossRef]
121. Lockwood, S.A.; HaghiPour-Peasley, J.; Hoffman, D.R.; Deslippe, R.J. Identification, expression, and immuno-reactivity of Sol i 2 & Sol i 4 venom proteins of queen red imported fire ants, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). *Toxicon* **2012**, *60*, 752–759. [PubMed]
122. Borer, A.S.; Wassmann, P.; Schmidt, M.; Hoffman, D.R.; Zhou, J.J.; Wright, C.; Schirmer, T.; Marković-Housley, Z. Crystal structure of Sol i 2: A major allergen from fire ant venom. *J. Mol. Biol.* **2012**, *415*, 635–648. [CrossRef] [PubMed]
123. King, T.P.; Lu, G. Hornet venom allergen antigen 5, Dol m 5: Its T-cell epitopes in mice and its antigenic cross-reactivity with a mammalian testis protein. *J. Allergy Clin. Immunol.* **1997**, *99*, 630–639. [CrossRef]

124. Lee, E.K.; Jeong, K.Y.; Lyu, D.P.; Lee, Y.W.; Sohn, J.H.; Lim, K.J.; Hong, C.S.; Park, J.W. Characterization of the major allergens of *Pachycondyla chinensis* in ant sting anaphylaxis patients. *Clin. Exp. Allergy* **2009**, *39*, 602–607. [CrossRef] [PubMed]
125. Padavattan, S.; Schmidt, M.; Hoffman, D.R.; Marković-Housley, Z. Crystal structure of the major allergen from fire ant venom, Sol i 3. *J. Mol. Biol.* **2008**, *383*, 178–185. [CrossRef] [PubMed]
126. Georgieva, D.; Seifert, J.; Ohler, M.; von Bergen, M.; Spencer, P.; Arni, R.K.; Genov, N.; Betzel, C. Pseudechis australis venomics: Adaptation for a defense against microbial pathogens and recruitment of body transferrin. *J. Proteome Res.* **2011**, *10*, 2440–2464. [CrossRef] [PubMed]
127. Aniszewski, T. *Alkaloids: Chemistry, Biology, Ecology and Applications*; Elsevier: Boston, MA, USA, 2015; p. 481.
128. Wallace, J.; Mansell, R. *Biochemical Interaction between Plants and Insects*; Plenum Press: New York, NY, USA, 1975.
129. Wink, M.; Roberts, M.F. Compartmentation of alkaloid synthesis, transport, and storage. In *Alkaloids: Biochemistry, Ecology and Medicinal Application*; Wink, M., Roberts, M.F., Eds.; Springer: New York, NY, USA, 1998; pp. 239–262.
130. MacConnell, J.G.; Blum, M.S.; Fales, H.M. The chemistry of fire ant venom. *Tetrahedron* **1971**, *27*, 1129–1139. [CrossRef]
131. Numata, A.; Ibuka, T. Alkaloids from ants and other insects. In *The alkaloids*; Brossi, A., Ed.; Academic Press: San Diego, CA, USA, 1987; pp. 193–315.
132. Tumlinson, J.; Silverstein, R.; Moser, J.; Brownlee, R.; Ruth, J. Identification of the trail pheromone of a leaf-cutting ant, *Atta texana*. *Nature* **1971**, *234*, 348–349. [CrossRef] [PubMed]
133. Ali, M.F.; Billen, J.P.; Jackson, B.D.; Morgan, E.D. The Dufour gland contents of three species of Euro-African *Messor* ants and a comparison with those of North American *Pogonomyrmex* (Hymenoptera: Formicidae). *Biochem. Syst. Ecol.* **1989**, *17*, 469–477. [CrossRef]
134. Wheeler, J.; Olubajo, O.; Storm, C.; Duffield, R. Anabaseine: venom alkaloid of *Aphaenogaster* ants. *Science* **1981**, *211*, 1051–1052. [CrossRef] [PubMed]
135. Leclercq, S.; Charles, S.; Daloze, D.; Braekman, J.C.; Aron, S.; Pasteels, J.M. Absolute configuration of anabasine from *Messor* and *Aphaenogaster* ants. *J. Chem. Ecol.* **2001**, *27*, 945–952. [CrossRef] [PubMed]
136. Jones, T.; Blum, M.; Fales, H.; Brandão, C.; Lattke, J. Chemistry of venom alkaloids in the ant genus *Megalomyrmex*. *J. Chem. Ecol.* **1991**, *17*, 1897–1908. [CrossRef] [PubMed]
137. Talman, E.; Ritter, F.; Verwiel, P. Structure elucidation of pheromones produced by the Pharaoh's ant, *Monomorium pharaonis* L. In *Mass spectrometry in biochemistry and medicine*; Castagnoli, A.F.N., Ed.; Raven Press: New York, NY, USA, 1974; pp. 197–217.
138. Francke, W.; Schröder, F.; Walter, F.; Sinnwell, V.; Baumann, H.; Kaib, M. New alkaloids from ants: Identification and synthesis of (3R,5S,9R)-3-butyl-5-(1-oxopropyl) indolizidine and (3R,5R,9R)-3-butyl-5-(1-oxopropyl) indolizidine, constituents of the poison gland secretion in *Myrmicaria eumenoides* (hymenoptera, formicidae). *Liebigs Ann.* **1995**, *1995*, 965–977. [CrossRef]
139. MacConnell, J.G.; Blum, M.S.; Fales, H.M. Alkaloid from fire ant venom: Identification and synthesis. *Science* **1970**, *168*, 840–841. [CrossRef] [PubMed]
140. Blum, M.; Jones, T.; Hölldobler, B.; Fales, H.; Jaouni, T. Alkaloidal venom mace: offensive use by a thief ant. *Naturwissenschaften* **1980**, *67*, 144–145. [CrossRef]
141. Jones, T.; Torres, J.; Spande, T.; Garraffo, H.; Blum, M.; Snelling, R. Chemistry of venom alkaloids in some *Solenopsis* (*Diplorhoptrum*) species from Puerto Rico. *J. Chem. Ecol.* **1996**, *22*, 1221–1236. [CrossRef] [PubMed]
142. Jones, T.H.; Adams, R.M.; Spande, T.F.; Garraffo, H.M.; Kaneko, T.; Schultz, T.R. Histrionicotoxin alkaloids finally detected in an ant. *J. Nat. Prod.* **2012**, *75*, 1930–1936. [CrossRef] [PubMed]
143. Allies, A.B.; Bourke, A.F.; Franks, N.R. Propaganda substances in the cuckoo ant *Leptothorax kutteri* and the slave-maker *Harpagoxenus sublaevis*. *J. Chem. Ecol.* **1986**, *12*, 1285–1293. [CrossRef] [PubMed]
144. Saporito, R.A.; Garraffo, H.M.; Donnelly, M.A.; Edwards, A.L.; Longino, J.T.; Daly, J.W. Formicine ants: An arthropod source for the pumiliotoxin alkaloids of dendrobatid poison frogs. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8045–8050. [CrossRef] [PubMed]
145. Braekman, J.C.; Daloze, D.; Pasteeis, J.; Hecke, P.V.; Declercq, J.P.; Sinnwell, V.; Francke, W. Tetraponerine-8, an alkaloidal contact poison in a neoguinean pseudomyrmecine ant, *Tetraponera* sp. *Z Naturforsch.* **1987**, *42*, 627–630.

146. Pacheco, J.A.; Mackay, W.P.; Lattke, J. *The Systematics and Biology of the New World Thief Ants of the Genus SOLENOPSIS (Hymenoptera: Formicidae)*; Edwin Mellen Press: Ceredigion, UK, 2013.
147. Pankewitz, F. Unusual natural products in insects: Molecular and chemical analyses of anthraquinone origin in galerucini leaf beetles. Ph.D. Thesis, Freie Universität, Berlin, Germany, 2006. Available online: http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_000000002645/00_Pankewitz.pdf?hosts= (accessed on 19 January 2016).
148. Callahan, P.S.; Blum, M.S.; Walker, J.R. Morphology and histology of the poison glands and sting of the imported fire ant (*Solenopsis saevissima* v. richteri Forel). *Ann. Entomol. Soc. Am.* **1959**, *52*, 573–590. [CrossRef]
149. Fox, E.G.P.; Bueno, O.C.; Yabuki, A.T.; de Jesus, C.M.; Solis, D.R.; Rossi, M.L.; de Lima Nogueira, N. General morphology and ultrastructure of the venom apparatus and convoluted gland of the fire ant, *Solenopsis saevissima*. *J. Insect Sci.* **2010**, *10*, 24. [PubMed]
150. Leclercq, S.; Braekman, J.C.; Daloze, D.; Pasteels, J.M.; van der Meer, R.K. Biosynthesis of the solenopsins, venom alkaloids of the fire ants. *Naturwissenschaften* **1996**, *83*, 222–225. [CrossRef]
151. Radwan, M.; Hanora, A.; Khalifa, S.; Abou-El-Ela, S.H. Manzamines: A potential for novel cures. *Cell. Cycle* **2012**, *11*, 1765–1772. [CrossRef] [PubMed]
152. Nikbakhtzadeh, M.R.; Tirgari, S.; Fakoorziba, M.R.; Alipour, H. Two volatiles from the venom gland of the Samsum ant, *Pachycondyla sennaarensis*. *Toxicon* **2009**, *54*, 80–82. [CrossRef] [PubMed]
153. Escoubas, P.; Blum, M. The biological activities of ant-derived alkaloids. In *Applied Myrmecology: A World Perspective*; Vander Meer, R., Jaffe, K., Cedeno, A., Eds.; Westview Press: Boulder, CO, USA, 1990; pp. 482–489.
154. Pelletier, S.W. *Alkaloids: Chemical and Biological Perspectives*; Springer: New York, NY, USA, 1983.
155. Brossi, A. *The alkaloids: Chemistry and Pharmacology*; Academic Press: New York, NY, USA, 1987.
156. Bosque, I.; Gonzalez-Gomez, J.C.; Loza, M.I.; Brea, J. Natural tetraponerines: A general synthesis and antiproliferative activity. *J. Org. Chem.* **2014**, *79*, 3982–3991. [CrossRef] [PubMed]
157. Tschinkel, W.R. *The Fire Ants*; Harvard University Press: Cambridge, MA, USA, 2006.
158. Chen, J.; Cantrell, C.L.; Shang, H.W.; Rojas, M.G. Piperideine alkaloids from the poison gland of the red imported fire ant (Hymenoptera: Formicidae). *J. Agric. Food Chem.* **2009**, *57*, 3128–3133. [CrossRef] [PubMed]
159. Chen, L.; Fadamiro, H.Y. Re-investigation of venom chemistry of *Solenopsis* fire ants. I. Identification of novel alkaloids in *S. richteri*. *Toxicon* **2009**, *53*, 469–478. [CrossRef] [PubMed]
160. Pianaro, A.; Fox, E.G.P.; Bueno, O.C.; Marsaioli, A.J. Rapid configuration analysis of the solenopsins. *Tetrahedron: Asymmetry* **2012**, *23*, 635–642. [CrossRef]
161. Brossi, A. Mammalian alkaloids II. In *The alkaloids*; Cordell, G.A., Ed.; Academic Press: San Diego, CA, USA, 1993; pp. 119–183.
162. Jouvenaz, D.; Blum, M.; MacConnell, J. Antibacterial activity of venom alkaloids from the imported fire ant, *Solenopsis invicta* Buren. *Antimicrob. Agents Chemother* **1972**, *2*, 291–293. [CrossRef] [PubMed]
163. Obin, M.S.; Vander Meer, R.K. Gaster flagging by fire ants (*Solenopsis* spp.): Functional significance of venom dispersal behavior. *J. Chem. Ecol.* **1985**, *11*, 1757–1768. [CrossRef] [PubMed]
164. Vander Meer, R.K.; Morel, L. Ant queens deposit pheromones and antimicrobial agents on eggs. *Naturwissenschaften* **1995**, *82*, 93–95. [CrossRef]
165. Blum, M.S.; Walker, J.R.; Callahan, P.S.; Novak, A.F. Chemical, insecticidal and antibiotic properties of fire ant venom. *Science* **1958**, *128*, 306–307. [CrossRef] [PubMed]
166. Lai, L.C.; Chang, R.; Huang, R.N.; Wu, W.J. Comparative toxicity of two fire ant venoms to *Spodoptera litura*. *Sociobiol* **2010**, *56*, 653–663.
167. Yeh, J.; Narahashi, T.; Almon, R. Characterization of neuromuscular blocking action of piperidine derivatives. *J. Pharmacol. Exp. Ther.* **1975**, *194*, 373–383. [PubMed]
168. Read, G.W.; Lind, N.K.; Oda, C.S. Histamine release by fire ant (*Solenopsis*) venom. *Toxicon* **1978**, *16*, 361–367. [CrossRef]
169. Foster, D.; Ahmed, K. Effect of piperidines and fire ant venom on ATPase activities from brain homogenate fractions and characterization of Na^+-K^+ ATPase inhibition. *Biochem. Pharmacol.* **1977**, *26*, 983–985. [CrossRef]
170. Lind, N.K. Mechanism of action of fire ant (*Solenopsis*) venoms. I. Lytic release of histamine from mast cells. *Toxicon* **1982**, *20*, 831–840. [CrossRef]
171. Javors, M.; Zhou, W.; Maas, J., Jr.; Han, S.; Keenan, R. Effects of fire ant venom alkaloids on platelet and neutrophil function. *Life sci.* **1993**, *53*, 1105–1112. [CrossRef]

172. Yi, G.; Mc Clendon, D.; Desaiah, D.; Goddard, J.; Lister, A.; Moffitt, J.; Vander Meer, R.; de Shazo, R.; Lee, K.; Rockhold, R. Fire ant venom alkaloid, isosolenopsin A, a potent and selective inhibitor of neuronal nitric oxide synthase. *Int J. Toxicol.* **2003**, *22*, 81–86. [CrossRef] [PubMed]
173. Howell, G.; Butler, J.; Farley, J.M.; Liu, H.L.; Nanayakkara, N.; Yates, A.; Gene, B.Y.; Rockhold, R.W. Cardiodepressant and neurologic actions of *Solenopsis invicta* (imported fire ant) venom alkaloids. *Ann. Allergy Asthma Immunol.* **2005**, *94*, 380–386. [CrossRef]
174. Arbiser, J.L.; Kau, T.; Konar, M.; Narra, K.; Ramchandran, R.; Summers, S.A.; Vlahos, C.J.; Ye, K.; Perry, B.N.; Matter, W. Solenopsin, the alkaloidal component of the fire ant (*Solenopsis invicta*), is a naturally occurring inhibitor of phosphatidylinositol-3-kinase signaling and angiogenesis. *Blood* **2007**, *109*, 560–565. [CrossRef] [PubMed]
175. Cavill, G.; Houghton, E. Some pyrazine derivatives from the Argentine ant, *Iridomyrmex humilis*. *Aust J. Chem.* **1974**, *27*, 879–889. [CrossRef]
176. Grünanger, P.; Pavan, A.; Quilico, S. Chimica degli insetti: Sul secreto odorosco del Formicide *Myrmicaria natalensis* (Fred). *Accad. Nazion Lincei* **1960**, *28*, 293–300.
177. Brand, J.; Blum, M.; Lloyd, H.A.; Fletcher, D. Monoterpene hydrocarbons in the poison gland secretion of the ant *Myrmicaria natalensis* (Hymenoptera: Formicidae). *Ann. Entomol. Soc. Am.* **1974**, *67*, 525–526. [CrossRef]
178. Schultz, D.R.; Arnold, P.I.; Wu, M.C.; Lo, T.M.; Volanakis, J.E.; Loos, M. Isolation and partial characterization of a polysaccharide in ant venom (*Pseudomyrmex* sp.) that activates the classical complement pathway. *Mol. Immunol.* **1979**, *16*, 253–264. [CrossRef]
179. Cavill, G.W.; Robertson, P.L.; Whitfield, F.B. Venom and venom apparatus of the Bull ant, *Myrmecia gulosa* (Fabr.). *Science* **1964**, *146*, 79–80. [CrossRef] [PubMed]
180. Billen, J. New structural aspects of the Dufour's and venom glands in social insects. *Naturwissenschaften* **1987**, *74*, 340–341. [CrossRef]
181. Pasteels, J.; Daloze, D.; Boeve, J.L. Aldehydic contact poisons and alarm pheromone of the ant *Crematogaster scutellaris* (Hymenoptera: Myrmicinae). *J. Chem. Ecol.* **1989**, *15*, 1501–1511. [CrossRef] [PubMed]

Chapter Four

Ant Venom Insecticidal and Antibacterial Activity

Chapter Four: Ant Venom Insecticidal and Antibacterial Activity

Compound Abstract

The previous two chapters revealed the complexity of ant venom at the protein and peptide level. They also showed that ant venom does indeed have insecticidal and antibacterial peptides that could be developed into pharmaceutical drugs. The following chapter aimed to establish whether some of the ant venoms investigated in this thesis possess any insecticidal or antibacterial activities. The following chapter confirmed the previously reported insecticidal activity of the ant *Ectatomma tuberculatum* as well as revealed, for the first time, insecticidal activity in the venoms of *Ectatomma brunneum* and *Odontomachus hastatus.* In addition, it was found that *Neoponera commutata* venom was not insecticidal. Interestingly this venom was found to be a potent antibacterial venom whereas that of *O. hastatus* was not found to be antibacterial.

4.1. Introduction

Insect pests are causing major problems in the agricultural and health industries. This is exemplified by their destruction of one fifth of the world's agricultural crops [2] and their transmission of deadly diseases such as malaria, yellow fever and zika virus. Synthetic chemicals such as pyrethroids and organophosphates are the major method of controlling insect pests [6]. However, there are several issues with these, such as increasing resistance and their effects on human health [7, 8]. Therefore, there is currently a necessity to discover novel natural alternatives (bioinsecticides) to help alleviate this problem [9, 10].

A number of peptide neurotoxins from animal venoms have been investigated as potential biopesticides. However, not all have proved effective at inhibiting or killing insect pest populations or suitable for commercialisation as pesticides [11]. This is usually due to the high cost of registering biological products, their low efficacy and the challenges associated with producing and formulating them [10-12]. These neurotoxins can also have an effect on vertebrate tissues, as many arthropod-derived peptides have promiscuous activity and can affect both vertebrate [13] and invertebrates [14, 15]. The present project concentrated on the identification of insecticidal activity of ant venom peptides for use as bioinsecticides. Accordingly, future studies will need to identify the effect of these peptides on a range of vertebrate tissues both *in vitro* and *in vivo*.

This approach has already proven to be successful with examples such as *Bt* toxin and SPEAR, a peptide-based bioinsecticidal toxin derived from the venom of the Australian funnel-web spider (www.vestaron.com). Social insects are an untapped source of potentially insecticidal peptides as they predate on insects. Thus far, the vast majority of research on hymenopteran venoms has focused on the medically important species which can cause serious allergic reactions such as *Solenopsis invicta* (fire ant), *Myrmecia pilosula* (jumper ants), *Apis mellifera* (honey bee) and *Vespula germanica* (European wasp) [16] whilst other hymenopteran species have been largely neglected. Despite the fact that there are over 15,000 ant species, of which 9,000 are venomous and prey predominantly on invertebrates [17], there are almost no discovery programs which focus on the isolation of peptide toxins targeting insects.

There have been neurotoxic peptides previously isolated from ant venoms that are essential for defence and prey capture [18, 19]. One of these is poneratoxin, a 25-residue linear peptide derived from the ant *Paraponera clavata.* It interacts with voltage-activated sodium channels [20, 21] and causes a concentration-dependent block of synaptic transmission in the CNS,

which leads to paralysis of the insect [20-22]. Another example is ectatomin, a disulfide-linked dimeric toxin (two α-helical chains) of 34 and 37 residues derived from the ant *Ectatomma tuberculatum* [23, 24]. This mini-protein is toxic to both humans and insects by causing an irreversible increase in cation leakage across the cell membrane by pore formation, inducing a decrease in membrane resistance and cell death [24]. The pores predominantly affect calcium ion movement, which accounts for its high toxicity, as there are a number of physiological processes that rely on calcium, such as neurotransmitter release and muscle contraction [24].

Consequently, ants are potential sources of neurotoxic peptides that could be used as novel bioinsecticidal compounds. Other strong indicators include the use of their venom for predation and defence against other insects [25], and the presence of disulfide-rich, low molecular weight peptides [21, 26-29] which are readily absorbed, highly stable and resistant to degradative environmental factors [15, 30, 31]. Given the extremely limited investigation of ant venoms, this is likely to generate structurally and pharmacologically novel toxins. As a result, novel insect-selective toxins could be developed as biopesticides to control agricultural pests and invasive insects that transmit diseases.

Antibacterial activity is a commonly reported property of ant venoms, as several ants have been found to have antimicrobial activities against both gram positive and negative bacteria to help prevent infections within colonies [19, 32-38]. These include the venoms of *Neoponera goeldii* [32], *Ectatomma quadridens* [35] and *Dinoponera quadriceps* [34] that have been found to have broad spectrum antibacterial activity. The first ant venom antimicrobial peptides (AMPs) were isolated from *N. goeldii* which has 15 AMPs known as ponericins which display potent activity comparable to that of AMPs from other venoms such as mellitin [32]. Bicarinalin from *Tetramorium bicarinatum* is another AMP with broad spectrum and potent antibacterial activity similar to melittin and other AMPs such as defensins [19]. Additionally, Zhang et al., (2012) identified several AMPs using a genome screening approach that included 26 inhibitor cystine-knot (ICK)-like peptides that were similar in sequence to those found in other hymenopteran venoms [38]. This diverse range of AMPs confirms the presence of peptides within ant venoms to protect against bacterial infections within the colony [32, 38]. Structurally, these peptides are quite distinct given that bicarinalin and mellitin lack disulfide bonds while ICK and defensins are structurally constrained given the presence of disulfide bonds. In addition, the solution structure for bicarinalin is not presently known, although it is believed to form an alpha helix similar to mellitin [19]. Thus structure activity relationships are difficult to predict.

One of the bacteria investigated in the present activity is *Staphylococcus aureus,* which is a Gram positive bacterium that is known to cause many serious hospital-acquired bacterial infections yearly and has been said to have caused the death of over 20,000 people in the last five years [34]. It is of particular importance, as it has evolved many enzymes that has made it resistant to many antibiotics such as beta lactamases and DNAses [34, 39]. *Escherichia coli* is a Gram negative bacterium common in the gut of humans and other animals which has also developed several resistance mechanisms to conventional antibiotics such as extended spectrum beta lactamases [40]. It would therefore be of great benefit to find novel sources of antibacterial compounds for use against these bacteria.

Thus, the goal of this study was to screen the venom of candidate ants for discovery of new insecticidal and antibacterial compounds. Insecticidal activity was determined using bioassays and antibacterial activity was determined using minimum inhibitory concentration (MIC) assays against the Gram positive *S. aureus* and the Gram negative *E. coli*. Finding a new source of antibacterial peptides is of great benefit, as we are currently in an era of high antibiotic resistance, with the World Health Organization (WHO) declaring it as a great threat to public health that requires immediate action [41]. Microbial resistance leads to many consequences such as treatment failures, prolonged illnesses and possible death [42].

4.2. Methods

4.2.1. Supply of ant venoms

Lyophilised ant venoms were provided by VenomeTech (Valbonne, France). Live ants were collected by Dr. Axel Touchard from the following species: *Neoponera commutata*, *Ectatomma tuberculatum*, *Ectatomma brunneum* and *Odontomachus hastatus,* were collected from various regions in French Guiana and stored at –20 °C. The venom glands were removed, dissected and pooled in 10% v/v acetonitrile (ACN) and centrifuged for 5 min at 14,400 rcf. The supernatant was collected, lyophilised and frozen at –20 °C.

4.2.2. Bicinchoninic acid (BCA) assay

The protein content of the crude venom was determined using a QuantiPro BCA Assay Kit (Sigma-Aldrich, USA), with bovine serum albumin (BSA) as the standard. This assay is based on the reduction of Cu^{2+} to Cu^{1+} by proteins in a concentration-dependent manner resulting in a purple colour reaction depending on the protein concentration [43].

The assay was conducted in a 96-well microtitre plate, with each well containing 150 µL of sample and 150 µL of BCA working reagent which consists of 25 parts of reagent A (sodium carbonate, sodium tartrate, and sodium bicarbonate in 0.2 M NaOH, pH 11.25), 25 parts reagent B (4% (w/v) bicinchoninic acid solution, (pH 8.5) and 1 part reagent C (copper (II) sulphate). The plate was then incubated at 37 °C for 2 hours to allow colour development, and absorbances of BSA standards and unknown samples measured at 562 nm using a Titertek Multiscan® PLUS MKII (type 313) plate reader (Flow laboratories Australasia, North Ryde, NSW).

Absorbances of the protein standards and venom samples were entered into Prism v6 (GraphPad Software Inc., La Jolla, CA, USA) and a standard curve was constructed using linear regression analysis. Protein concentrations of unknown samples were interpolated from the standard curve.

4.2.3. Insect toxicity testing

Insecticidal activity of crude venom was tested using direct subcutaneous injections in crickets. Freeze-dried venom fractions or crude venom were resuspended in normal insect saline (NIS) (see Table 4.1 for composition). House crickets (*Acheta domestica),* 3rd–4th instar nymphs, sex not determined) of mass 70–130 mg were then injected with venom in groups of 5–10 crickets.

Table 4.1: Composition of insect saline solution for insect toxicity testing of venom.

Chemical	mM
NaCl	200
KCl	3.1
HEPES*	10
$CaCl_2$	5
$MgCl_2$	4
Sucrose	50
Bovine Serum Albumin (BSA)	0.1% w/v to reduce non-specific binding
pH	7.4

*4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid

Due to the likely breakdown of venom components in the gut, venom is usually inactive via the oral route. Therefore, venom was administered via direct subcutaneous lateroventral thoracic injections between legs two and three. The needle was inserted under the pentagonal sternal plate at a 45° angle, to avoid puncturing the ventral nerve cord or gut (see Figure. 4.1). Solutions of different concentrations were injected at a dose of 5 µL/100 mg of cricket body weight. Control crickets were injected with a solution of insect saline/0.1% BSA.

Figure 4.1: Lateroventral cricket injections of toxin solution.

Image courtesy of Professor Graham Nicholson.

The test equipment consisted of a light source, dissecting microscope and a 0.5 mL insulin syringe with a fixed 29-gauge needle (Terumo Microfine-III) in an Arnold vernier microapplicator (see Figure 4.2 for a picture of the setup) (Burkard Scientific Supply, Rickmansworth, UK). No anaesthesia was performed that may have masked any early symptoms. A set of 10 control crickets was injected with NIS for each bioassay experiment performed. The survival of these crickets was always ~90–100% if sufficient care was exercised.

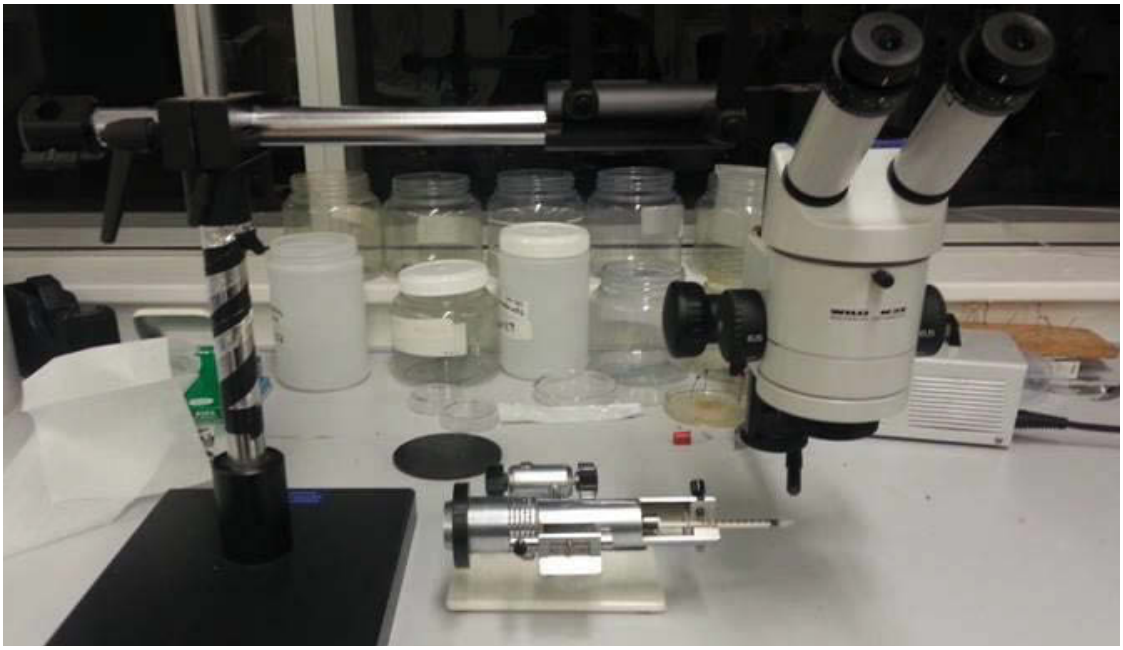


Figure 4.2: Apparatus used for insect toxicity testing.
Figure shows the Arnold-Burkard microapplicator and the stereo microscope (Olympus SZ model).

Following injection, crickets were housed in groups of 5 in plastic petri dishes containing a portion of carrot for food and water-moistened filter paper. Dishes were kept at room temperature and observed at 0, 2, 4, 6, 8, 12, 24, 48 hours post-injection for signs of toxicity (see Table 4.2)

Table 4.2: Categories of neurotoxicity signs in crickets

Category	Observed signs in crickets
Normal	Unaffected, maintain normal activity and posture. Infrequent abdominal pulses
Mild effects	Mostly normal posture, but show some twitches, ataxia, erratic movements or stilting
Knockdown	Unable to maintain normal posture AND display twitches or tremor of appendages
Mortality	No movement when shaken or touched: insects dead or completely paralysed

The median lethal dose (LD_{50}) and median knockdown dose (KD_{50}) was determined for the venom of the ant *E. tuberculatum.* The experiment was performed by injecting whole venom at doses of 1, 5, 10, 20 and 100 µg/g body weight. The LD_{50} and KD_{50} were determined 72 hours post injection by plotting data as dose-response curves. All data shown represent the mean ± S.E. KD_{50} and LD_{50} doses. Mathematical curve fitting was accomplished using GraphPad Prism

version 7 (GraphPad Software, CA, USA). All curve-fitting routines were performed using non-linear regression analysis employing a least squares method and the following form of the logistic equation:

$$y = 100 \times \left(1 - \frac{1}{1 + (x/LD_{50})^{nH}}\right)$$

where *x* is the toxin dose, *n*H the Hill coefficient (slope parameter), and LD_{50} is the median inhibitory dose causing lethality (LD_{50}) or knockdown (substitute KD_{50}).

4.2.4. MALDI-TOF mass spectrometry

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) was employed to determine the molecular mass of venom peptides/proteins. Lyophilised venoms were re-suspended in Milli-Q water and 0.5–0.75 µL samples of each fraction were spotted onto a 384-well Opti-TOF target plate (AB Sciex, USA), overlaid with 0.5–0.75 µL of the MALDI matrix using the dried droplet method [44], and allowed to dry at room temperature. The matrix used was α-cyano-4-hydroxycinnamic acid (CHCA) matrix (5 mg/mL) dissolved in 1% TFA/100% ACN/0.1 M $NH_4H_2PO_4$ (45/45/10). The instrument was calibrated externally using a mixture of peptides with known molecular masses in the same *m*/*z* range (human angiotensin II, *m*/*z* 1046.5423; Synthetic peptide P14R, *m*/*z* 1533.8582; ACTH fragment 18–39, *m*/*z* 2465.1989; bovine insulin oxidised B chain, *m*/*z* 3494.6513), and mass-measurement accuracy was typically ±0.3 Da. 0.5–0.75 µL of the calibrant mix was spotted onto the 'Cal' spots of the stainless steel target plate and overlaid with an equal volume of CHCA for external calibration. Spots were analysed using a 5800 MALDI-TOF/TOF mass spectrometer (AB Sciex, USA). A UV laser beam (349 nm nitrogen laser) of variable intensity was fired at the MALDI target spot using MALDI linear mode, releasing the sample into the gas phase.

Spectra were acquired in linear positive mode, to allow resolution of peptides with masses up to 10,000 Da. Mass spectra were acquired from 1,000–10,000 m/z with 400 laser shots averaged per second using the Dynamic Exit algorithm selected, which monitors spectral quality and stops shot accumulation if a user defined threshold is met. Signals above 10,000 *m*/*z* are usually proteins where the use of CHCA matrix is not as effective for generation of MS signal, while signals below 1,000 m/z correspond to low mass molecules masked by matrix clusters [45]. Spots were sampled several times until an acceptable spectrum was seen based on adequate signal intensity in the target *m*/*z* range. Data Explorer® v4.11 software (AB SCIEX,

CA, USA) was used to analyse spectra. Spectra were subjected to baseline correction with a correlation factor of 0.7 and Gaussian smoothing to reduce noise with a 5 point filter width.

4.2.5. Minimum inhibitory concentration assays

Microdilution growth assays were used to assess the minimum inhibitory concentration (MIC) of the venom of the ants *Pachycondyla commutata* and *Odontomachus hastatus* against *Staphylococcus aureus* (ATCC 25923) and *Escherichia coli* (MG1556) strains. Glycerol stocks of *S. aureus* and *E. coli* strains from a plate stored at 4 °C were streaked onto a 1.5% CAMHB agar plate (Cation adjusted Mueller-Hinton Broth (Oxoid, cat. # CM0405)) and incubated overnight at 37 °C.

For each strain, a colony from the CAMHB plate was incubated overnight in 2 mL of liquid CAMHB at 37 °C in a shaking incubater at 250 rpm. Diluted overnight bacterial culture (10^7 CFU/mL, determined by CFU counting) was used to inoculate wells of a sterile 96-well flat-bottomed plate. Various concentrations of venom or antibiotic controls (0.5 mg/mL amphotericin and 1 mg/mL tetracycline) were added to the designated wells by twofold serial dilutions with TSB growth media to a final volume of 150 µL. Untreated controls were also included. In the static growth assay, the plates were briefly shaken to mix the contents of each well and the optical density (OD) of each well was measured at 595 nm in a Synergy HT BioTek plate reader (BioTek Instruments Inc., USA). The plate was then incubated without shaking in a 37 °C humidified incubator for 24 h, and the OD of each well was measured again at the end of incubation. The OD difference between the two time points was used to measure cell growth. The MIC was defined as the lowest concentration of antimicrobial agent that inhibited 99% growth of *S. aureus* when compared to the untreated control.

4.3. Results

4.3.1. Ant venom insecticidal activity

In order determine the insecticidal activity of the venoms investigated, lyophilised venom was re-constituted in insect saline and injected into house crickets. Following injection, crickets were observed for up to 48 h. The venom with the highest activity was that of *E. tuberculatum,* which at a dose of 100 µg of venom per gram of cricket, caused 90% lethality of crickets within 1 h of injection (Figure. 4.3). The remainder of the crickets were also affected, as they exhibited knockdown effects which included immediate loss of locomotion with no movement except for leg twitching after tapping the dish. Figure 4.4 shows control and *E. tuberculatum* crickets 24 hours post-injection.

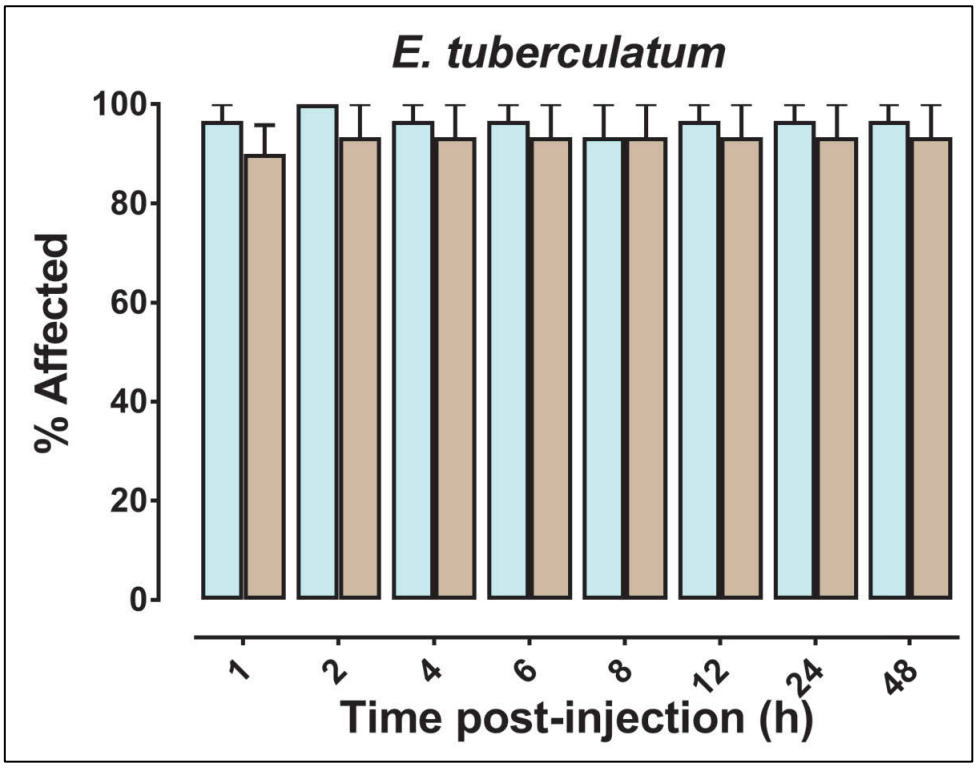


Figure 4.3: Acute toxicity of *E. tuberculatum* whole venom in crickets.
Data shows the percentage of crickets affected by injection of 100 µg/g of whole *E. tuberculatum* venom. Cyan columns indicate percentage knockdown, brown columns indicate percentage lethality. Values represent the mean ± S.E. (*n* = 3)

Figure 4.4: House crickets after an acute toxicity assay.
House crickets following injection with insect saline solution at 48 hours after injection with vehicle (insect saline/0.1% BSA) at 5 µL/100 mg body weight (left) and at 48 hours after injection with 100 µg/g of *E. tuberculatum* whole venom in vehicle at 48 hours (right).

E. tuberculatum venom was then investigated further, given its potent action on insects at 100 µg/g. The venom was subsequently tested at lower concentrations to determine the LD_{50} and KD_{50}. The dose-response curve for the effects at 48 h can be seen in Figure. 4.4. From this graph, the 48 h LD_{50} was calculated to be 9 ± 3 µg/g and the KD_{50} was 6 ± 2 µg/g (*n* = 3). The effect on the crickets was rapid, as seen in Figure 4.5 as crickets were affected at 1 h post-injection. It can also be seen that the LD_{50} did not change significantly 12 h post-exposure.

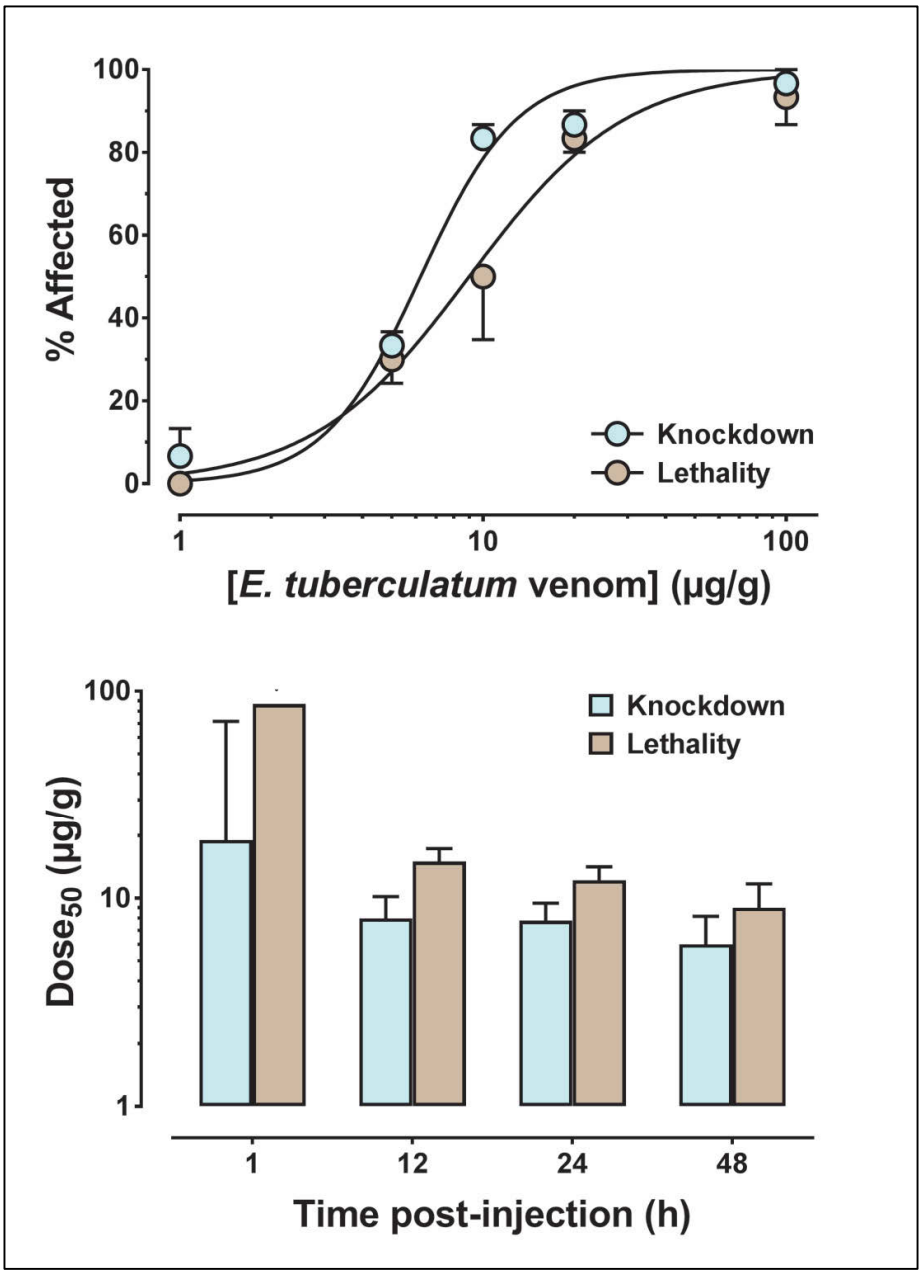


Figure 4.5: Acute toxicity profile of *E. tuberculatum* whole venom in crickets.
(A) Dose response curve at 48 h post intrathoracic injection. Panel B shows the LD_{50} and KD_{50} of the venom at 4 different time points (1, 12, 24 and 48 h post-injection). (*n* = 3)

The venom of *E. brunneum* was also tested for insecticidal activity using the same methods described above. Although *E. brunneum* is of the same genus as *E. tuberculatum*, the insecticidal activity was markedly lower. This is shown in Figure 4.6 where it can be seen that one hour post-injection, only 30% knockdown was seen. However, by 4 hours post-injection, this number increased to 70%. The lethal effects started appearing at 6 hours post-injection by which stage 30% of the injected crickets had died. By the end of the observation period (72 hours), the percentage lethality had reached 70%. This shows that *E. brunneum* is a potential

source of insecticidal peptides, as there is significant insecticidal activity seen in the whole venom assay.

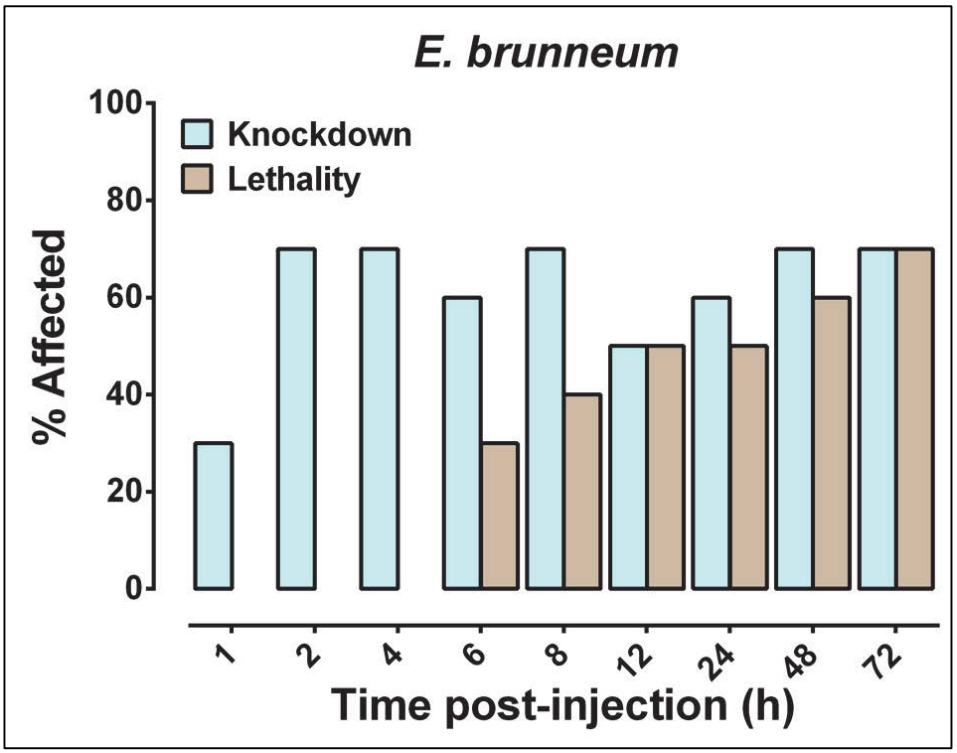


Figure 4.6: Acute toxicity of *E. brunneum* whole venom in crickets.
Data shows the percentage of crickets affected by injection of 100 µg/g of whole *E. brunneum* venom. Cyan columns indicate percentage knockdown, brown columns indicate percentage lethality. Values represent the mean ± S.E. (*n* = 1).

Contrary to the activities seen in the venoms of *E. brunneum* and *E. tuberculatum*, the venoms of *O. hastatus* and *N. commutata* are less potent on house crickets. As seen in Figure 4.7, the highest percentage knockdown seen for *O. hastatus* was at 1 hr post-injection. This number decreased overtime, with the final percentage of crickets affected being 33%. In terms of lethality, no crickets had died until 48 hours post injection where 23% of crickets were affected.

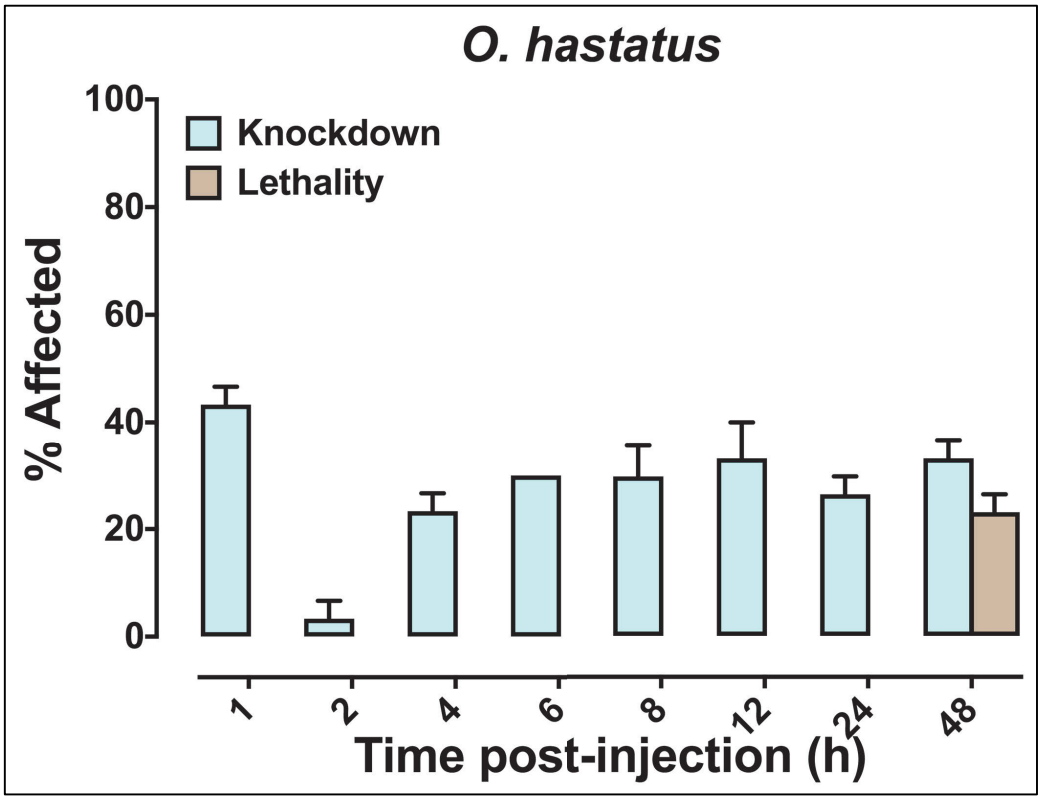


Figure 4.7: Acute toxicity of *O. hastatus* whole venom in crickets.
Data shows the percentage of crickets affected by injection of 100 µg/g of whole *O. hastatus* venom. Cyan columns indicate percentage knockdown, brown columns indicate percentage lethality. Values represent the mean ± S.E. (*n* = 3)

The venom of *N. commutata* seems to be the least potent of all the venoms tested for insecticidal activity. As seen in Figure 4.8, the initially low knockdown effects of 23% seen 1 hour post-injection were lost and then re-appeared after 12 hours and only reached 27% by the end of the 48 hour observation period. Additionally, *N. commutata* had the lowest overall lethality and this was only observed 48 hours post-injection at a level of 17%.

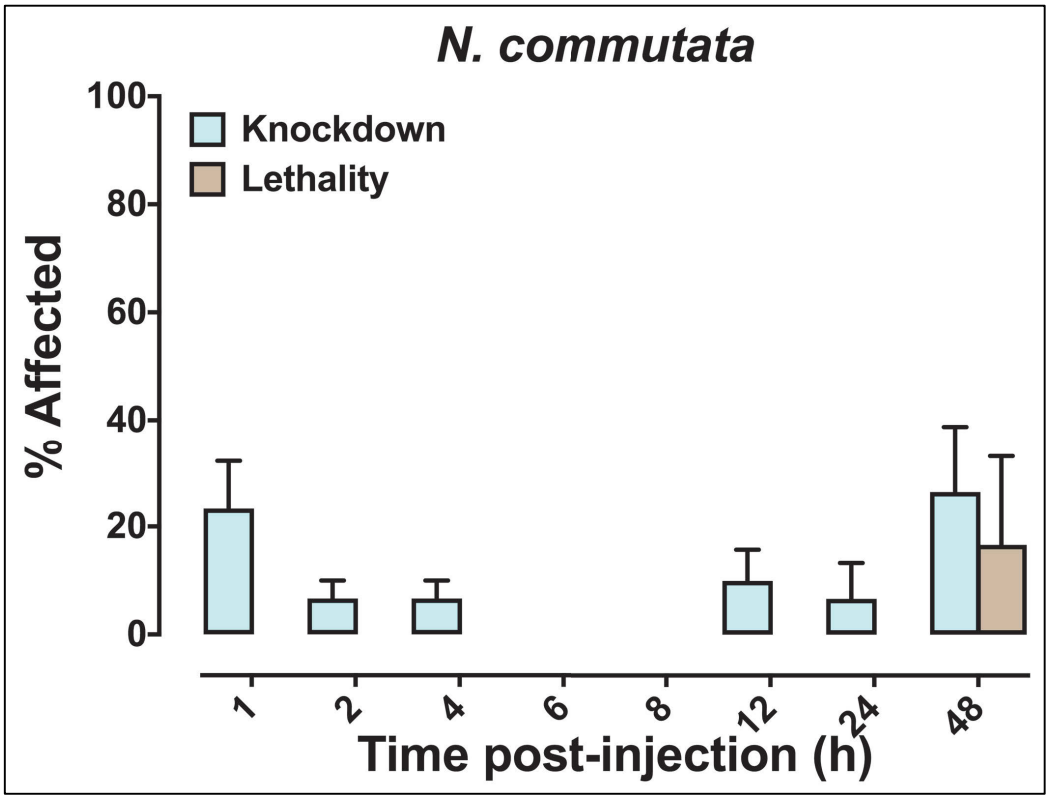


Figure 4.8: Acute toxicity of *N. commutata* whole venom in crickets.
Data shows the percentage of crickets affected by injection of 100 µg/g of whole *N. commutata* venom. Cyan columns indicate percentage knockdown, brown columns indicate percentage lethality. Values represent the mean ± S.E. (*n* = 3).

4.3.2. Minimum Inhibitory Concentration Assays

The *in vitro* antibacterial activity against planktonic growth for the venoms *O. hastatus* and *N. commutata*, and for each of the control antibiotics was established by determining the MICs against *S. aureus* and/or *E. coli*. The venom of *O. hastaus* was tested for activity against the gram positive bacterium *S. aureus* using serial doubling dilutions of the venom at concentrations of 0.0078 mg/mL to 1 mg/mL. As seen in Figure 4.9, the effects were small and not dose-dependent. The results indicate that although some antibacterial activity was seen with the venom of *O. hastatus,* it was not as significant as that seen with *N. commutata* (Figure 4.10).

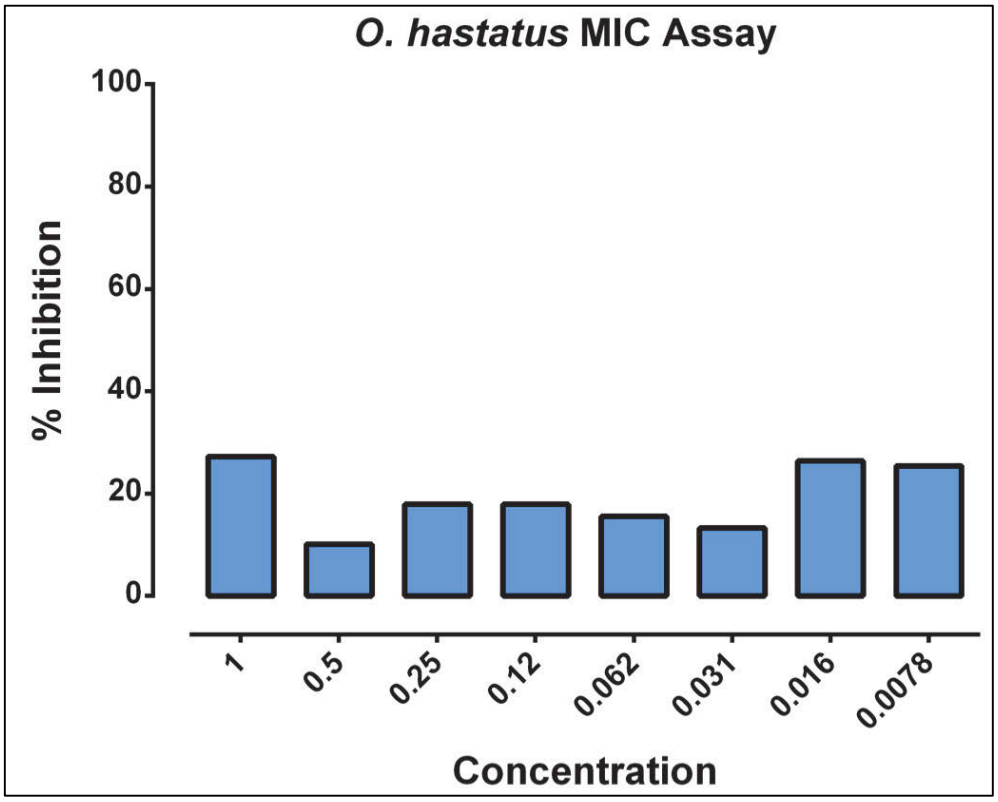


Figure 4.9: MIC results of *O. hastatus* venom against *S. aureus*.
Figure shows *O. hastatus* concentrations tested in mg/mL and resulting percentage inhibition of *S. aureus* (Y-axis). (n=1)

The venom of *N. commutata* was significantly more effective at inhibiting bacterial growth compared to *O. hastatus* venom. This venom was tested against two types of bacteria, *S. aureus* and the gram negative *E. coli* using an MIC assay that showed dose-dependent inhibition. As seen in Figure 4.10, the highest inhibition of *S. aureus* seen was 60% at venom concentrations of 0.25 mg/mL or higher. However, this inhibition was lost at the 0.12 mg/mL venom concentration, where only 50% inhibition was seen. In terms of activity against *E. coli*, *N. commutata* was able to inhibit its growth more effectively than it was able to inhibit *S. aureus*, because 100% inhibition was seen at all concentrations higher than 0.12 mg/mL which inhibited the growth by 50% (see Figure 4.10).

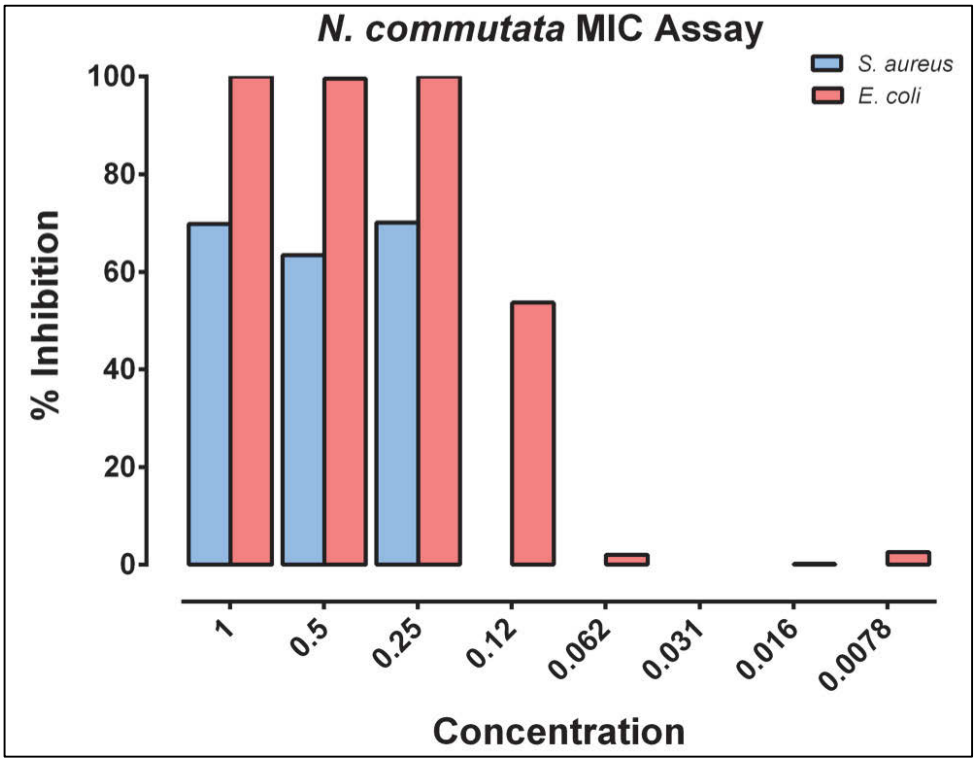


Figure 4.10: MIC results of *N. commutata* venom against the bacteria *S. aureus* and *E. coli*.
Figure shows *N. commutata* concentrations tested in mg/mL and resulting percentage inhibition of *S. aureus* (Y-axis). Blue bars represent *S. aureus* inhibition whilst pink bars represent *E. coli* inhibition.

In order to visualise the EC_{50}, a dose-response curve was constructed utilising the data obtained from the *N. commutata* MIC assay. Figure 4.11 clearly shows how *N. commutata* is more effective at inhibiting *E. coli* than *S. aureus* as the EC_{50} for the two bacteria were 0.3083 mg/mL for *S. aureus* and 0.1221 mg/mL for *E. coli.* Due to limited amount of material, the experiment for Fig 11 could only be repeated once. It clearly shows antimicrobial effects, but lack of 100% antibacterial action in this particular experiment has resulted in an increased hill slope value for *S. aureus,* and an apparent decrease in potency as evidenced by an increased MIC value. Ideally, this would need to be repeated if more material was available to confirm the activity.

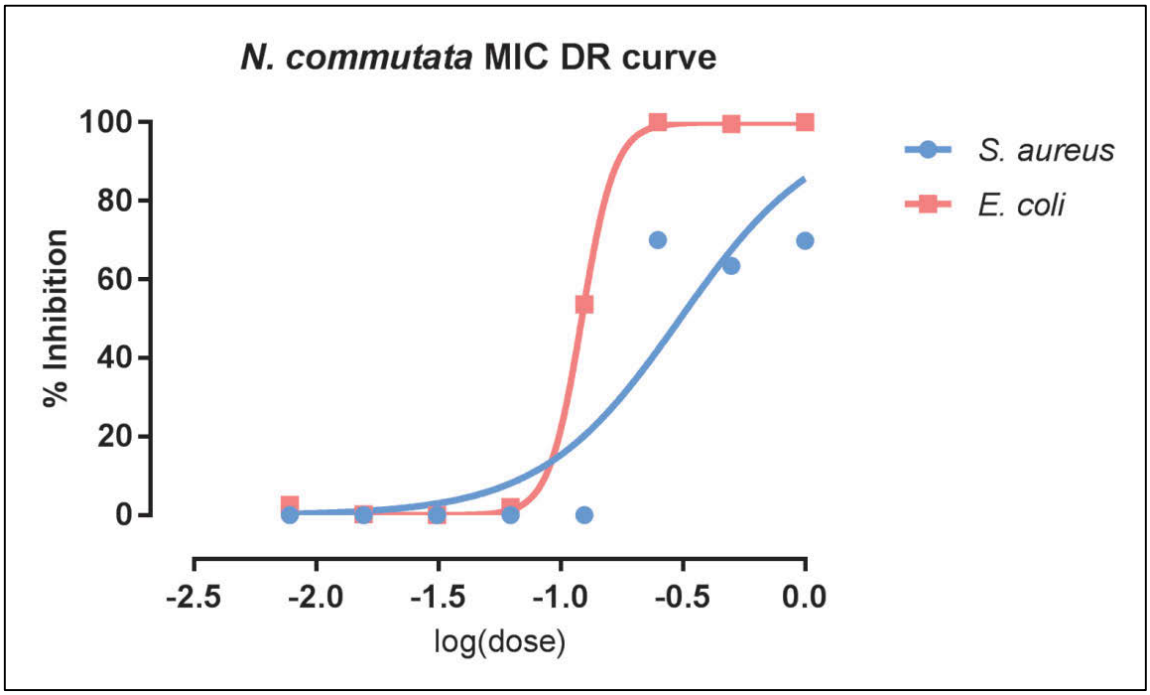


Figure 4.11: *N. commutata* MIC assay dose-response curve.
Figure shows percentage inhibition (Y-axis) at each of the doses tested, shown as log(dose) values (X-axis). Blue curve represents *Staphylococcus aureus* inhibition whilst pink curve represents *Escherichia coli* inhibition by *N. commutata* at the different doses tested.

4.3.3. Whole venom MALDI-TOF MS

After testing the activities of the venoms, we wanted to identify the complexity of the venom peptide components. Therefore, whole venom MALDI-TOF mass spectrometry was performed on all four venom samples to identify the major peptides. The resulting spectra (see Figure 4.12) revealed that there were distinct interspecies differences in abundance and the mass profiles of all four venoms despite the fact that some of the species are closely related. Spectra were acquired in positive linear mode using CHCA matrix, focussing on masses in the range 1,000–10,000 Da.

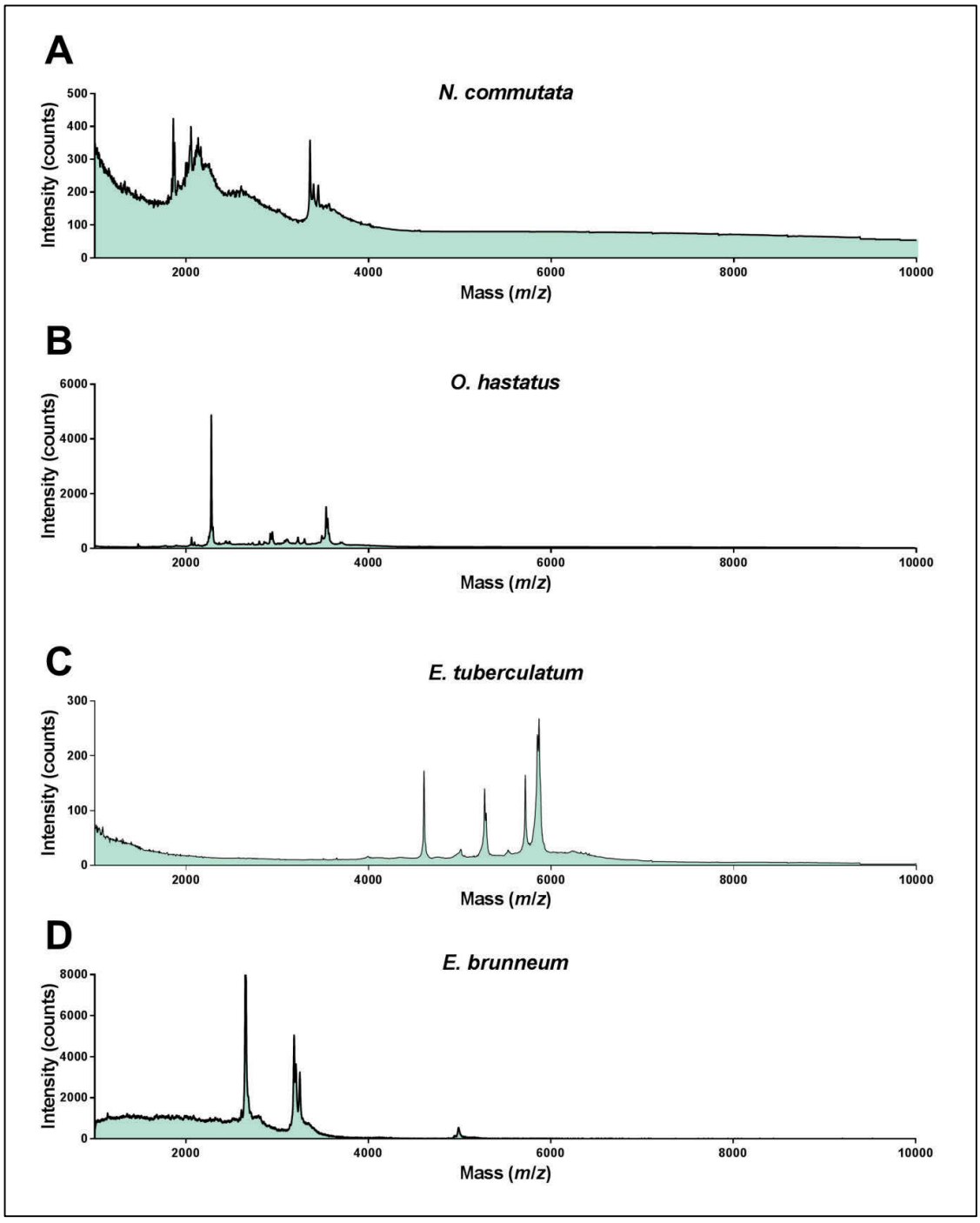


Figure 4.12: MALDI-TOF mass spectra of four crude ant venoms.
Spectra were acquired on linear mode using CHCA matrix, focussing on masses in the range 1,000–10,000 *m/z*. Panel A: *Neoponera commutata*, B: *Odontomachus hastatus,* C: *Ectatomma tuberculatum,* D: *Ectatomma brunneum.*

As seen in Figure 4.12, the two poneroid venoms *N. commutata* and *O. hastatus* appear to have peptides of mass 4000 m/z or smaller. This was also the case with the formicoid venom *E. brunneum,* which also had a peptide at approximately 5000 m/z. *E. tuberculatum* seems to be unique, in that the peptides seen are all between 4000–6000 m/z. As a first assessment of their complexity, all these venoms seem to be less complex than expected in terms of their peptide compositions. However, mass spectrometry limitations such as ion suppression should be taken into account, as it is later shown that there are several more peptides within each venom.

4.4. Discussion

The experiments presented in this chapter show that ant venoms are viable sources of insecticidal and antibacterial peptides. Of all the venoms tested for insecticidal activity, that of *E. tuberculatum* was found to be the most potent. This was not unexpected, as *E. tuberculatum* is one of the few ant venoms investigated for insecticidal activity and has already been shown to be insecticidal [22, 26]. Numerous peptide toxins have been isolated from venomous animals, scorpions, cone snails and spiders in particular. However, not many toxins have been isolated and characterised from ant venoms. Those with insecticidal activity are ectatomins [24-26], ponericins [32] and poneratoxin [22]. Arseniev *et al*. (1994) found the LD_{50} of ectatomin to be 6.8 μg/g [26], and our calculated LD_{50} for *E. tuberculatum* crude venom was 9.0 μg/g, which is in good agreement with the results obtained. The higher LD_{50} may be a result of differences in isoforms of the peptide or differences in the test insect, as the Arseniev (1994) study used cockroaches. The potent insecticidal activity of *E. tuberculatum* venom has been attributed to ectatomin, which is of mass 7928 Da [26]. However, this peptide was not observed in the whole venom MALDI spectrum which suggests that it is not a major component of the venom or has been masked by ion suppression. A bioassay-guided fractionation and characterisation of the individual venom fractions is necessary to isolate ectatomin and identify whether it is identical to that previously described.

In comparison to one of the other characterised insecticidal toxins from ant venom, ponericin, *E. tuberculatum* venom is more potent as the LD_{50} of ponericin is 130 μg/g (~40 nmol/g) [32]. There are other venoms with similar potency, such as that of the spider *Segestris florentina* (LD_{50}: 4–10 μg/g) [46]. Notably, *E. tuberculatum* venom is not as potent as other venom toxins that have been isolated, such as the Australian funnel-web spider toxins ω-hexatoxin-Hv1a and κ-hexatoxin-Hv1c which have LD50 values of 0.36 μg/g (89 pmol/g) and 0.34 μg/g (91 pmol/g), respectively, when injected into *A. domesticus* [47, 48]. However, *E. tuberculatum* venom likely represents a complex mixture containing highly potent toxins.

The next most potent venom was that of *E. brunneum* which caused lethality of 70% of the crickets by 48 hours. This is a significant finding as *E. brunneum* is an ant species with no available data on its venom composition or activity. The only other ant species within the *Ectatomma* genus that has been investigated for biological activity is *E. quadridens* which was found to have significant antibacterial activity against Gram positive and negative bacteria at lower doses than *E. tuberculatum* venom [35]. This confirms that the *Ectatomma* genus

harbours ants that are significantly different in their biological activities. This is consistent with our bioassay results as well as our MALDI-MS results which revealed different peptide profiles.

The insecticidal activity of crude *N. commutata* and *O. hastatus* venom was negligible. This was surprising, especially in the case of *N. commutata,* as insecticidal peptides have been found in the ant *Neoponera goeldii* which is of the same genus [32]. However, complete inactivity shouldn't be assumed as we only tested for activity against the house cricket *A. domesticus* which is taxonomically related to economically important pest insects such as locusts and grasshoppers [49, 50]. It has previously been shown that the venom of the tarantula *Coremiocnemis validus* was only active on crickets, not cockroaches or mice [51]. Therefore, the venoms should be tested in other insect orders to determine their phyla-selectivity such as the mealworm (*Tenebrio molitor*) which has been shown to be 25–100 times more sensitive to the actions of spider venoms [52]. Another way in which the biological targets of these venoms can be better understood is through investigation of the natural predatory behaviour of the ant. A possible reason for the inactivity of *N. commutata* venom might be due to the fact that their sole prey is termites [53, 54], therefore *N. commutata* has not evolved appreciable quantities of toxins that affect other insects such as crickets.

An additional reason for not discounting the presence of insecticidal toxins in the venoms of *N. commutata* and *O. hastatus* is that only crude venom was tested, which has been found to be much lower in activity compared to individual toxins [52]. The reason for this is unclear, but it is speculated that the venomous animal needs to maintain a diverse toxin repertoire in its venom for use against a wide range of predators in order to ensure that all potential prey can be paralysed or killed. In addition, venoms contain a number of other compounds such as processing enzymes and cellular debris from the venom secretion process which decrease the insecticidal potency on a per gram of protein basis [55].

Although *N. commutata* was not insecticidal, it was found to be antibacterial which is consistent with previous work on the ant from the same genus *N. goeldii* [32]. We found that *N. commutata* was more effective at inhibiting the Gram negative *E. coli* than the Gram positive *S. aureus*. This is consistent with other investigations into venoms, such as that of *T. bicarinatum* where its venom peptide bicarinalin was found to be active against Gram negative bacteria at levels lower than control agents but lower activity against gram positive bacteria [19] and crotamine from the rattlesnake which was found to be bactericidal against *E. coli* but only had weak activity against *S. aureus* [42]. However, scorpions have been reported to be the opposite, with activity against *S. aureus* but not *E. coli* [56, 57]. Our results also indicate that *N.*

commutata is less potent than other ant venoms investigated for antibacterial activity as *E. quadridens* had an MIC of 30 µg/mL against *E. coli* whereas the MIC of *N.* commutata venom was 120 µg/mL [35]. It's activity against *S. aureus* is also higher than other reported ant venoms, as the MIC obtained (250 µg/mL) is a lot higher than the lowest reported by *D. quadriceps* (6.5 µg/mL). It must be noted that activity differs significantly depending on the strains tested, as one of the strains tested against D. quadriceps had an MIC of 100 µg/mL.

From the above data it seems that *N. commutata* is antibacterial, which is a significant finding as venom toxins have the potential for development as therapeutics in the face of a diminishing number of effective antimicrobial agents on the market. This is especially true of bee venoms which seem to be more potent antibacterial agents, as MICs of 0.086 µg/mL have been reported against methicillin-resistant *S. aureus* (MRSA) [58].

4.5. Conclusion

After the present testing of ant venoms for insecticidal and antibacterial activities, it has been confirmed that ant venoms are potential candidates for bioinsecticide and antimicrobial drug discovery. It has also been shown through MALDI-TOF MS that ant venoms contain complex peptide profiles that need to further investigation. The remainder of this thesis represents an in-depth proteomic investigation of the peptide and protein profiles of various ant venoms as well as a transcriptomic investigation of one particular ant venom gland.

Chapter Five

Analysis of ant venom collection methods using *Paraponera clavata*

Chapter Five: Peptidomic and Proteomic Analysis of Electrically Stimulated and Manually Dissected *Paraponera clavata* Venom

Compound Abstract

The previous chapter found that not all ant venoms are equal in their pharmacological activities,with two of the four venoms tested for insecticidal activity possessing significant activity whilst the other two were less active. Interestingly, the venom with the lowest insecticidal activity, *N. commutata* venom, had the highest antimicrobial activity. However, due to low venom yields, bioassay-guided fractionation of the venoms was not able to isolate the exact peptide responsible for the observed activities. The small size of ants is the main reason for this problem, therefore this chapter aimed to identify the best collection method for ant venoms.

This chapter was published in the *Journal of Proteome Research* and it investigated two collection methods for ant venoms - manual venom gland dissection and electrical stimulation - using the bullet ant *Paraponera clavata*. It was found that both methods gave different venom profiles, particularly at the peptidomic level. Therefore, for characterisation purposes it is best to use dissected venom as it encompasses a wider range of venom components. Electrical stimulation could be used in investigations that focus on certain components that are known to be present in this collection method as it is easier to perform.

Certificate of authorship and originality

This paper was published in the *Journal of Proteome Research*. The following is a copy reprinted with permission from ACS publications, Copyright © 2016 American Chemical Society. I certify that the work present in chapter five of this thesis has not previously been submitted as part of the requirements for a degree. I also certify that I carried out all the experimental work, analysis and interpretation of the data presented in this paper.

The remaining authors listed on this manuscript contributed in the following way:

- Axel Touchard: Collected all ants and dissected the venom glands
- Frederic Peticlerc: Collected electrically stimulated venom
- Alain Dejean. and Jerome Orivel: Assisted in venom collection
- Matthew Padula: Mass spectrometry support
- Pierre Escoubas: Proof-read the manuscript
- Graham Nicholson: Assisted in experimental design and manuscript direction
- All authors proof-read manuscript

Principal supervisor

Professor Graham Nicholson	Samira Ryma Aili
_____	_____
Signature	Signature
_____	_____
Date	Date

Journal of proteome research

Article

pubs.acs.org/jpr

Combined Peptidomic and Proteomic Analysis of Electrically Stimulated and Manually Dissected Venom from the South American Bullet Ant *Paraponera clavata*

Samira R. Aili,[†] Axel Touchard,[‡] Frédéric Petitclerc,[‡] Alain Dejean,[‡,§] Jérôme Orivel,[‡] Matthew P. Padula,[∥] Pierre Escoubas,[⊥] and Graham M. Nicholson*[,†]

[†]Neurotoxin Research Group, School of Life Sciences, University of Technology Sydney, New South Wales 2007, Australia

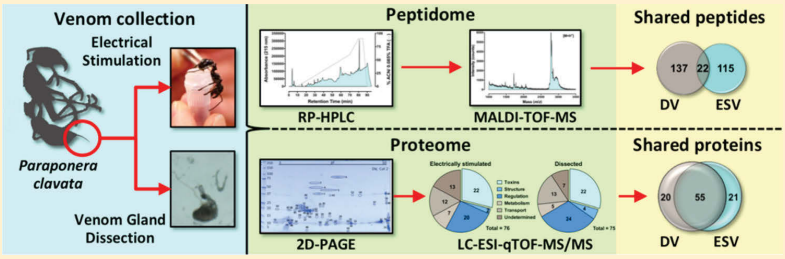
[‡]CNRS, UMR Ecologie des Forêts de Guyane (EcoFoG), AgroParisTech, Cirad, INRA, Université des Antilles, Université de Guyane, Université des Antilles, 97310 Kourou, France

[§]Ecolab, Université de Toulouse, CNRS, INPT, UPS, 31400 Toulouse, France

[∥]Proteomics Core Facility, Faculty of Science, University of Technology Sydney, New South Wales 2007, Australia

[⊥]VenomeTech, 473 Route des Dolines — Villa 3, Valbonne 06560, France

Supporting Information



ABSTRACT: Ants have evolved venoms rich in peptides and proteins used for predation, defense, and communication. However, they remain extremely understudied due to the minimal amount of venom secreted by each ant. The present study investigated the differences in the proteome and peptidome of the venom from the bullet ant, *Paraponera clavata*. Venom samples were collected from a single colony either by manual venom gland dissection or by electrical stimulation and were compared using proteomic methods. Venom proteins were separated by 2D-PAGE and identified by nanoLC-ESI-QTOF MS/MS. Venom peptides were initially separated using C18 reversed-phase high-performance liquid chromatography, then analyzed by MALDI-TOF MS. The proteomic analysis revealed numerous proteins that could be assigned a biological function (total 94), mainly as toxins, or roles in cell regulation and transport. This investigation found that ca. 73% of the proteins were common to venoms collected by the two methods. The peptidomic analysis revealed a large number of peptides (total 309) but with <20% shared by the two collection methods. There was also a marked difference between venoms obtained by venom gland dissection from different ant colonies. These findings demonstrate the rich composition and variability of *P. clavata* venom.

KEYWORDS: *proteome, peptidome, ants, bullet ant, venom, MALDI-TOF MS, nanoLC-ESI-QTOF MS/MS, 2D-PAGE, Paraponera clavata, dissection, electrical stimulation*

INTRODUCTION

Ants of the order Hymenoptera are a diverse group of insects, with ca. 16,000 extant species belonging to 3 clades and 16 subfamilies.[1,2] The number of species within each subfamily ranges from 1 to more than 6000.[2] Of these subfamilies, 13 comprise ants that employ a stinger to inject a peptide-rich venom used for predation, defense, and communication.[3−6] Although hundreds of peptides and proteins, in each venom, bear those functions, little work has been undertaken to characterize these components in detail (for recent reviews, see refs 2 and 7).

Venom proteins and peptides are now routinely being investigated as drug or biopesticide leads, particularly those from bees,[8] spiders,[9,10] and cone snails[11−14] as many of these organisms are of medical importance, which suggests interesting pharmacological properties. *Paraponera clavata* (commonly known as the bullet ant or giant tropical ant) is a stinging ant that belongs to the subfamily Paraponerinae of the poneroid clade.[15,16] It is one of the few ants of medical importance due to its extremely painful sting[16−18] that is used

Received: October 31, 2016
Published: January 24, 2017

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

to deter predators and capture prey.[19] *P. clavata* venom is toxic to both vertebrates and invertebrates and is thought to elicit at least part of its action through the neurotoxic peptide poneratoxin.[20−23] This toxin induces slowly activating $Na_V1.7$-mediated currents at low activation voltages and also sustained sodium currents as a result of a slowing of Na_V channel inactivation,[16] similar to that previously described in rat and frog skeletal muscle.[23] All studies performed on this ant thus far have focused on this peptide because it is seen as a potential peptide for commercialization as an insecticide due to its high activity.[20] An example of other ant-venom-derived peptides that have been thoroughly investigated are ponericins from the ant *Neoponera goeldii* (Ponerinae; formerly *Pachycondyla goeldii*). These were found to have antimicrobial and insecticidal activities.[24] These peptides confirm that the venom of ants can be potential source of bioactive peptide leads that can be developed into therapeutic or insecticidal drugs.

To date the major hurdle for studying ant venoms has been the limited amount of venom present in their venom glands; however, this restriction has been overcome with recent advances in mass spectrometry and DNA sequencing techniques.[25] Ant venom is conventionally obtained by venom gland dissection; however, it can also be collected by electrical stimulation, which can be considered a more efficient method of collection because ants remain alive following venom collection. These two collection methods have been employed with spider,[26] cone snail,[13] wasp,[27] scorpion,[28] and bee venoms,[3] and differences in the venom components have been seen between these two collection methods.[29−31] The dissection technique has been used for many years for a variety of different arthropods after it was first introduced in 1961.[32] However, if isolation of only one certain venom component is desired, then electrical stimulation would be the better collection method, as it increases the likelihood of isolating the potent component due to the absence of contaminants such as structural proteins from the venom gland.[27,33] The other advantage of electrical stimulation, if performed en mass, is that it can give a higher yield of venom, is less time-consuming, and keeps the organism alive for subsequent venom extraction.[34] If mass milking is not possible, as is the case with vespid venoms,[27] then automated methods of milking individual venoms have also been proposed.[35] Previous work suggests that dissected venom contains most of the proteins that are in electrically stimulated venom but also other proteins that are usually contaminants from the venom gland including cellular proteins used in metabolic machinery or toxin maturation and processing.[13,27,36,37] Electrically stimulated venom may therefore give a more genuine representation of the venom components. The present study aimed to compare the proteomic and peptidomic components of *P. clavata* obtained by electrical stimulation and manual dissection. While previous studies have looked at individual components of either the proteome or peptidome of ant venoms, this is the first in-depth study of both proteomic and peptidomic components. It also aimed to identify the presence of intercolony variations to further confirm the diversity of this group of insects.

EXPERIMENTAL SECTION

Venom Collection

Venom samples from *P. clavata* were collected from two separate ant colonies at "la Montagne des Singes" near Kourou, French Guiana. To determine differences in venom profiles using the two collection procedures, ants from one colony (colony 1) were separated into two random groups and venom was extracted by either electrical stimulation or manual dissection of individual workers. To establish the extent of intercolony variations in *P. clavata* venom profiles, venom from another colony (colony 2) was also obtained by venom gland dissection only. The manually dissected venom gland samples were prepared by storing ants at −20 °C prior to the dissection of the venom glands. After dissection, the glands were pooled in 10% (v/v) acetonitrile (ACN)/water. Samples were then centrifuged for 5 min at 14 400 rpm (12 000g_{av}), and the supernatant was collected and lyophilized prior to storage at −20 °C. Electrically stimulated venom was collected by placing individual *P. clavata* ants into a glass insert. Ants were then milked by placing a pair of tweezers attached to electrodes on their abdomen, and a 12 V square wave pulse of 1 ms duration was delivered at 100 Hz. The venom was collected into a glass tube insert and diluted with 100 μL of 10% ACN. The venom was then transferred to an Eppendorf tube prior to being freeze-dried.

Reverse-Phase High Performance Liquid Chromatography (RP-HPLC)

Lyophilized crude electrically stimulated or manually dissected venom was separated on a Shimadzu HPLC system (Shimadzu, Kyoto, Japan) using a Vydac analytical C18 RP-HPLC column (218TP54, 25 cm × 4.6 mm, 5 μm particle size; Grace, MD). Fractionation of venom peptide components was achieved using a gradient of 0.085% (v/v) trifluoroacetic acid (TFA) in ACN (solvent B) in 0.1% (v/v) TFA in water (solvent A). Separation was achieved using a flow rate of 1 mL/min with the following gradient: 0−5 min, 0% solvent B; 5−65 min, 0−60% B; 65−75 min, 60−90% B; 75−80 min, 90% B; 80−85 min, 90−0% B; and 85−90 min, 0% B. Peaks were monitored at absorbances of 280 and 215 nm and manually collected; then, fractions were lyophilized and stored at −20 °C for further analysis.

MALDI-TOF MS Analysis

MS analysis of each HPLC fraction was performed on an AB SCIEX TOF/TOF 5800 mass spectrometer (AB SCIEX, Framingham, MA). Each venom fraction was manually collected and freeze-dried, then reconstituted in 10 μL of water. Subsequently, 1 μL of each fraction was then spotted onto a MALDI plate and overlaid with 0.5 to 0.75 μL of matrix using the dried droplet method.[38] The matrix consisted of 5 mg/mL of α-cyano-4-hydroxycinnamic acid (CHCA) dissolved in 1% (v/v) TFA/100% (v/v) ACN/0.1 M $NH_4H_2PO_4$ (45:45:10). Each spectrum was calibrated externally using a mixture of peptides with known molecular masses in the same m/z range (AB SCIEX). Mass spectra were acquired in positive linear mode over the range 1000−10 000 m/z with 400 laser shots accumulated for each sample based on the acceptance parameters and adequate signal intensity. Signals below 1000 m/z were not recorded as they are comprised of mostly matrix-related ion clusters.[39] Data Explorer v4.11 software (AB SCIEX) was used to analyze spectra from individual RP-HPLC fractions to characterize the number and masses of peptides per venom, as previously reported.[9] In brief, spectra were subjected to baseline correction with a correlation factor of 0.7 and Gaussian smoothing to reduce noise with a 5-point filter width. All mass attributions were verified manually, and a mass list was created for each LC-MALDI-TOF MS run. Potential adducts from oxidation, hydration, sodium, and

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

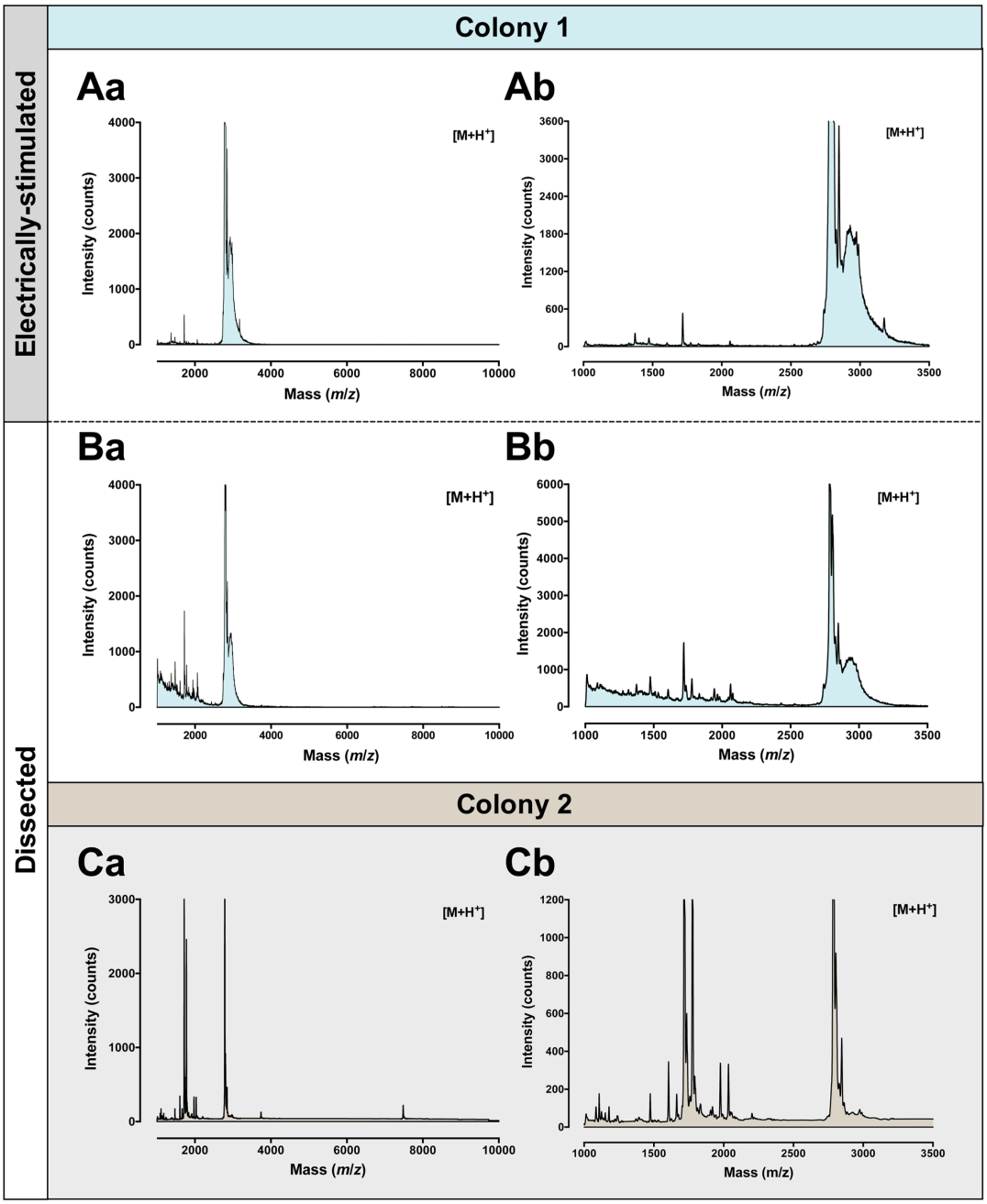


Figure 1. Representative MALDI-TOF MS spectra of *P. clavata* whole venom. Venom was obtained by electrical stimulation (A) and manual dissection (B,C) from colony 1 (A,B) and colony 2 (C). Initial screening was between 1000 and 10 000 m/z (left-hand panels Aa−Ca). Right-hand panels show mass spectra between 1000 and 3500 m/z, where peptide abundance was the highest (panels Ab−Cb).

deamination were manually removed from all mass lists as well as any potential dimers or doubly charged species. Peptides with mass matches within ±1.0 Da in adjoining HPLC fractions were considered identical and were removed from the data set as they likely reflect incomplete separation. For each venom, the peptides from each RP-HPLC fraction were consolidated into one mass list. Mass matching across data sets was performed in Microsoft Excel (Redmond, WA) to identify

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

similar peptides across the different venoms. Two-dimensional scatter plots, termed "2D venom landscapes", were constructed using Prism v6 software (GraphPad, La Jolla, CA). All peptide masses detected in each HPLC fraction were plotted as a function of their m/z values (x axis) and their HPLC retention time reflecting their hydrophobicity (y axis). Area-proportional Euler plots depicting overlapping peptide masses or proteins from different venoms were constructed using eulerAPE software (www.eulerdiagrams.org/eulerAPE/).

Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE)

300 μg samples of venom, collected by electrical stimulation or manual dissection, were resuspended in 100 μL of 7 M urea, 2 M thiourea, 1% (v/v) C7BzO, 50 mM Tris HCl pH 8.8 before reduction and alkylation of disulfide bonds with 5 mM tributylphosphine (TBP) and 20 mM acrylamide monomers for 90 min. The sample was subjected to 2D-PAGE, as previously described.[40] In brief, the sample was fractionated by isoelectric focusing on an 11 cm pH 3−10 IPG strip (Bio-Rad, Hercules CA) for 100 kVh and then further separated by molecular size by SDS-PAGE on a 4−20% Tris-glycine gel (Criterion TGX, Bio-Rad, NSW, Australia). The gel was then placed in a fixing solution containing 40% (v/v) methanol and 10% (v/v) acetic acid for 30 min at room temperature before staining with Flamingo fluorescent stain (Bio-Rad) for 1 h. After staining, the gel was scanned using a Molecular Imager PharosFX Plus system (Bio-Rad) with QuantityOne software. To perform nanoLC-ESI-QTOF MS/MS, gels were additionally stained with Coomassie Blue G250 overnight, then destained with 1% (v/v) acetic acid for 2 h.

NanoLC-ESI-QTOF MS/MS Analysis

Gel spots or bands were excised and digested as previously described.[41] Peptide samples were then fractionated using a Tempo nano LC system (Eksigent, Dublin, CA). Samples were first loaded onto a Michrom reversed-phase trapping cartridge at a rate of 20 μL/min, then eluted onto a 75 μm × 150 mm PicoFrit column (New Objective, MA) packed with Magic C18AQ chromatography resin (Bruker-Michrom Biosciences, Auburn, CA). An increasing gradient of ACN at 300 nL/min eluted the peptides that were ionized at 2300 V by the MicroIonspray II head holding the PicoFrit column with integrated emitter into the source of a QSTAR Elite Quadrupole TOF mass spectrometer (AB SCIEX). An Intelligent Data Acquisition (IDA) experiment was performed with a mass range of 350−1500 Da scanned for peptides of charge state 2^+ to 5^+ with an intensity of more than 30 counts/scan. The selected peptides were then fragmented, and the ion fragment masses were measured over a mass range of 50−1500 Da. The mass of the precursor peptide was then excluded for 15 s.

Protein Identification

The MS/MS data files were searched using Mascot (v2.4.0) against the LudwigNR database. This database comprised the UniProt, plasmoDB, and Ensembl databases (vQ114), provided by the Australian Proteomics Analysis Facility (APAF, Ltd.), hosted by the Walter and Eliza Hall Institute for Medical Research Systems Biology Mascot Server. The database was searched with the following parameter settings: fixed modifications, none; variable modifications, propionamide, oxidized methionine, deamidated asparagine, and glutamine; enzyme, semitrypsin; number of allowed missed cleavages, 3; peptide mass tolerance, 100 ppm; MS/MS mass tolerance, 0.2 Da; and charge state, 2^+, 3^+, and 4^+. PEAKS 7 software (Bioinformatics Solutions, ON, Canada) was used to generate de novo sequences. A database comprising ant protein sequences available on the NCBI database and known contaminant proteins was compiled and uploaded onto PEAKS. De novo sequences were then also searched against this database. Remaining unassigned de novo sequences were searched on the online BLAST search engine MS-BLAST (http://genetics.bwh.harvard.edu/msblast/).[42] Peptide identification used a cutoff score of >20 (−10 lgP). Proteins were assigned as homologous when spots were located at the same position (mass and pI) on at least two 2D gels and the protein was identified from the peptide fragment in at least one of the gel spots. Functions were assigned according to the GO annotation within UniProt, where available. Spectra have been provided for proteins matched with a coverage of <2% and have been included in the Supporting Information (Figures S1−S4).

Chemical Reagents

All chemicals used were of analytical grade and, unless otherwise stated, were sourced from Sigma-Aldrich (NSW, Australia). All buffers were prepared using Milli-Q (18 MΩ/cm^2) water.

RESULTS

Peptidome Analysis

Initially, the venom collected from *P. clavata* obtained by electrical stimulation and venom gland dissection of colony 1 was subjected to MALDI-TOF MS analysis. Venom obtained by either method was dominated by peptides in the mass range 2600−3200 m/z (Figure 1Aa−Ba). It was also found that no peptides appeared to be present at masses greater than 3200 m/z and that manually dissected venom appeared to contain more peptides than electrically stimulated venom, particularly at masses <2100 m/z. Nevertheless, the most intense peptide signal in both electrically stimulated and manually dissected venoms was of mass 2786.3 and 2787.6 m/z, respectively (Figure 1Ab−Bb).

Ants from colony 2 were also dissected and their venom glands pooled and analyzed alongside those from colony 1 (Figure 1Ca). The MALDI-TOF MS spectrum of the manually dissected colony 2 venom sample was similar to the one obtained from colony 1, particularly the dissected venom sample. However, colony 2 venom had differences in the relative intensities of the major peptides as well as the presence of numerous additional peptides below 1500 m/z (Figure 1Cb). Nevertheless, the most abundant peptide 2786.4 m/z matched that seen in colony 1 samples (±1 Da).

To reduce the complexity of the venoms and therefore diminish the effects of ion suppression, all venoms were subsequently separated by C18 RP-HPLC. The HPLC chromatograms of all three venoms are shown in Figure 2A−C. It can be seen that the venoms obtained by venom gland dissection and electrical stimulation from the same colony (colony 1) have similar profiles, each with 39 fractions. However, while the venom elution profile from colony 2 has many similar peaks, overall it has a different profile with only 29 peaks and an additional dominant peak at 50 min (Figure 2C).

Although each RP-HPLC fraction was seemingly homogeneous, mostly with clear sharp peaks, multiple peptides were observed to coelute in the same LC fraction when further resolved by MALDI-TOF MS analysis. This varied from no

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

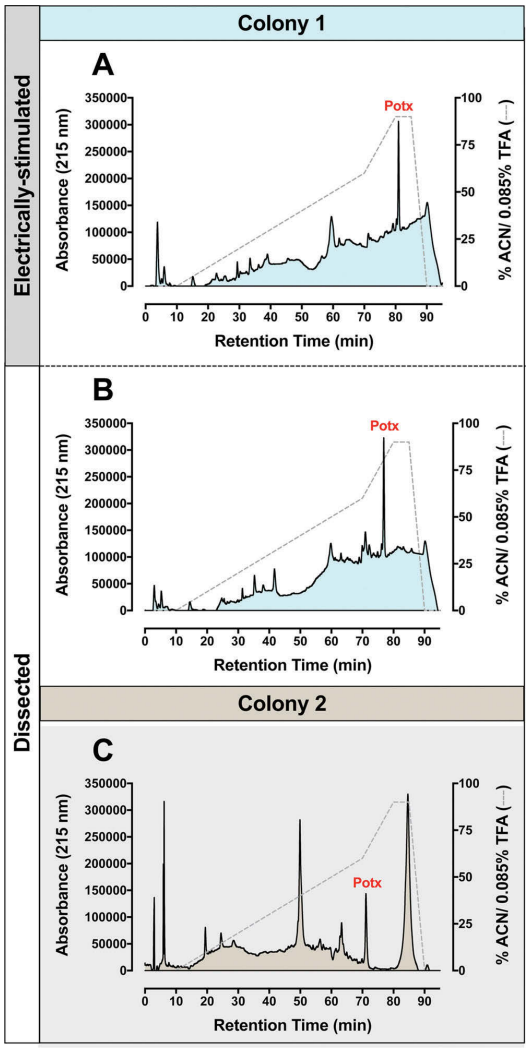


Figure 2. RP-HPLC chromatograms of *P. clavata* venom glands extracted by electrical stimulation or venom gland dissection. Venoms were separated by analytical C18 RP-HPLC using an ACN/0.085% TFA gradient of 1%/min (gray dotted line) at 1 mL/min. Panels show the chromatographic elution profile of *P. clavata* venoms from colony 1 obtained by (A) electrical stimulation and (B) manual dissection. Panel C shows profile of venom from colony 2 obtained by manual dissection. Chromatographic peaks were monitored at 215 and 280 nm (only absorbance at 215 nm is shown for clarity). Peaks in which poneratoxin was found are labeled as "Potx".

peptides in some fractions up to 32 masses in fraction 39 of the dissected venom. Overall, each venom contained a large number of small mass peptides, with cumulative total counts of 115 masses for electrically stimulated venom from colony 1, 137 masses for manually dissected venom from colony 1, and 354 for manually dissected venom from colony 2. This highlights the overall complexity of *P. clavata* venom particularly from colony 2. Although both dissected venoms had a similar mass distribution, with 95% of peptides <2.9 kDa (gray shaded areas in Figure 3B, C), the electrically stimulated venom from colony 1 contained a small number (n = 15; 13%) of higher mass peptides in the range 3−10 kDa that were not present in either of the dissected venoms from colony 1 or 2 (Figure 3A). Venom from colony 2 also had some similar, but not identical, high-mass peptides; however, they had a different

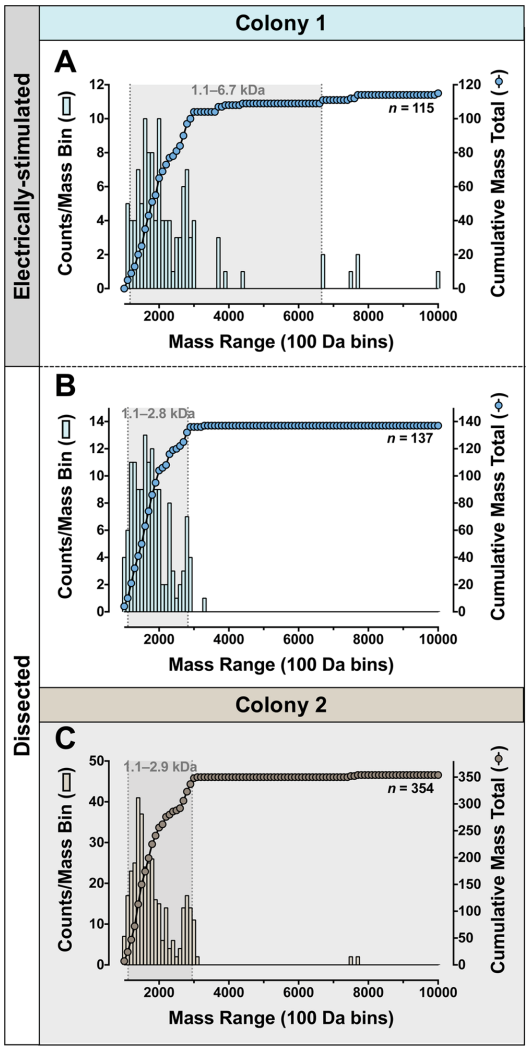
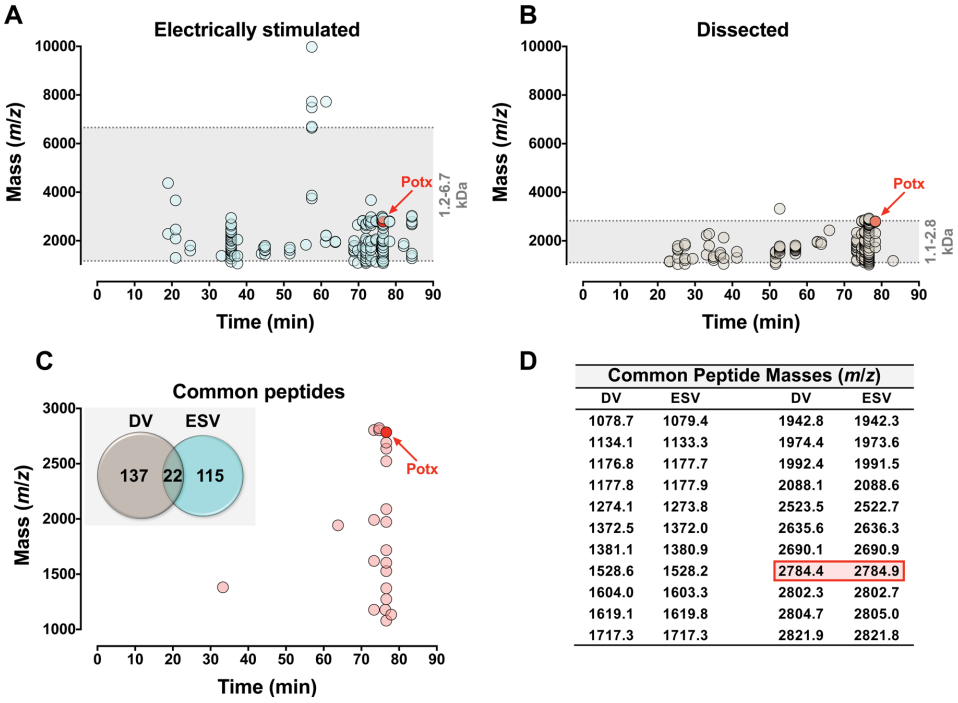


Figure 3. LC-MALDI-TOF mass distribution and total peptide count of *P. clavata* venom obtained using the different venom collection methods. (A−C) Histograms show the abundance of peptide masses in each venom sorted into 100 Da mass bins (columns; left-hand ordinate axis) and the cumulative total of peptide masses in each venom (circles; right-hand ordinate axis). Total peptide counts are indicated in the upper right-hand corner of each panel. (A) *P. clavata* venom obtained by electrical stimulation of ants from colony 1, (B) manual dissection of ants from colony 1, and (C) manually dissection of ants from colony 2. Shaded areas represent 5−95% of the total number of peptides.

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351



D

Common Peptide Masses (*m/z*)			
DV	ESV	DV	ESV
1078.7	1079.4	1942.8	1942.3
1134.1	1133.3	1974.4	1973.6
1176.8	1177.7	1992.4	1991.5
1177.8	1177.9	2088.1	2088.6
1274.1	1273.8	2523.5	2522.7
1372.5	1372.0	2635.6	2636.3
1381.1	1380.9	2690.1	2690.9
1528.6	1528.2	2784.4	2784.9
1604.0	1603.3	2802.3	2802.7
1619.1	1619.8	2804.7	2805.0
1717.3	1717.3	2821.9	2821.8

Figure 4. 2D landscapes of *P. clavata* venom peptides from colony 1 determined using MALDI-TOF MS. Panels show venom peptides collected by (A) electrical stimulation, (B) manual dissection, and (C) peptides common to both collection methods based on a mass tolerance of ±1 Da. Gray shaded areas represent 5−95% of peptide masses. Inset in panel C shows an area-proportional Euler plot depicting variation in numbers of shared and distinct peptides. Potx is represented by red circles. DV, manually dissected venom; ESV, electrically stimulated venom. (D) Mass list for peptides common to manually dissected and electrically stimulated venom. Potx is indicated by a red shaded box.

RP-HPLC retention time compared with those from electrically stimulated venom from colony 1, likely representing unique peptides. This further highlights the complexity of venoms obtained from different colonies of *P. clavata*.

Although there were only subtle differences in the chromatograms and the peptide mass distributions of the three venoms, comparisons between the peptide mass lists from the three venoms revealed large differences between venoms from each of the two collection methods and also between different colonies. Figure 4 shows that venoms obtained by the two different collection methods from ants of colony 1 had only 22 common masses between them (Figure 4C,D). This means that 81 and 84% are peptides unique to electrically stimulated and manually dissected venoms, respectively, despite the fact that the ants belong to the same colony.

Proteome Analysis

2D-PAGE analysis was performed in triplicate for each of the three samples, with representative gels shown in Figure 5. Overall, the resultant gels look similar to one another, however, when comparing the samples from the same colony (Figure 5A,B), 24 spots were common (by mass and p*I*) to both the electrically stimulated venom (total of 53 spots; Figure 5A) and the manually dissected venom (total of 38 spots; Figure 5B) gels. Interestingly, the sample from colony 2 (Figure 5C) contained no unique spots that were not present in either of the other two samples and was not subjected to further analysis.

To identify whether the different collection methods resulted in different venom protein profiles, nanoLC-ESI-QTOF MS/MS analysis was performed on the 2D gel spots from *P. clavata* venom obtained from colony 1 by the different extraction techniques. Because of the absence of any specific proteomic or transcriptomic data for *P. clavata*, several different databases were searched to determine the highest homology matches. Two search engines were employed, Mascot and Peaks, while MS Blast was used to determine proteins with high homology to peptide sequences determined by de novo sequencing of MS/MS spectra. An extensive list of protein matches from each spot is included as Supplementary Tables S1 and S2. A total of 73 and 75 proteins were identified from venom obtained by electrical stimulation and manual dissection, respectively (Figure 6A), of which 54 identified proteins were common to both venoms (Figure 6B). This showed that although there were proteins shared by the two collection methods (Table S3), there still were identified proteins unique to each method: 21 for electrically stimulated venom and 19 for manually dissected venom.

In the 2D gels (Figure 5) it can be seen that the most abundant proteins in all three gels was a train of spots (#12, #40−42, and #72) at ca. 18−21 kDa with an acidic p*I* of ca. 4.5 to 5.5. These spots were all identified as isoforms of the enzyme phospholipase A_2 (PLA_2) in all three venom samples. PLA_2 proteins have been identified in many other hymenopteran venoms.[3,43−46] The high abundance of these proteins,

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

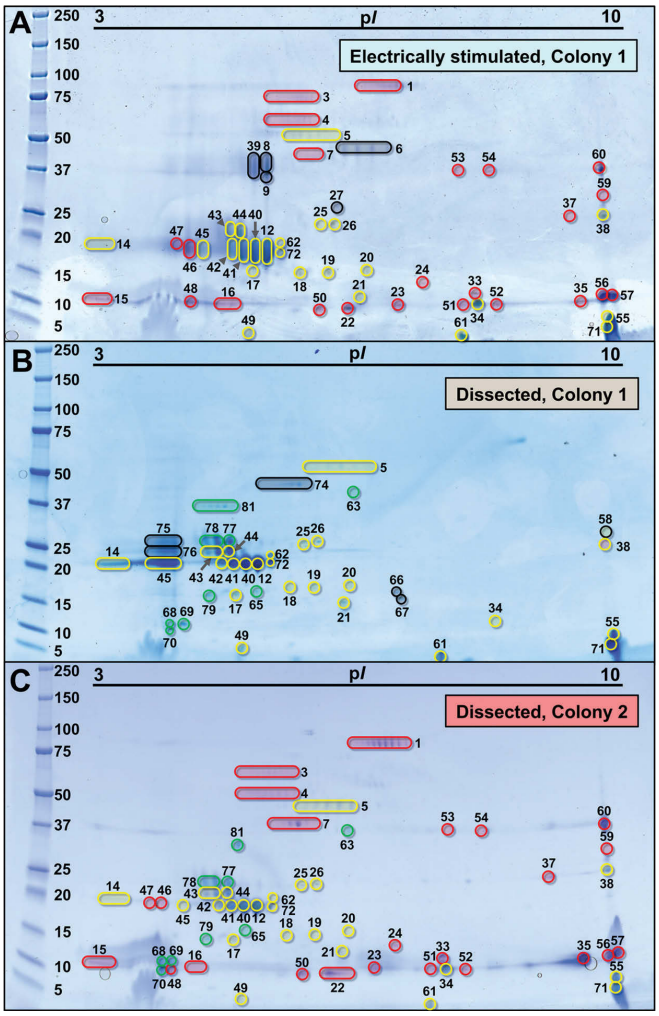


Figure 5. 2D SDS-PAGE of reduced and alkylated *P. clavata* venoms. Venoms were extracted by (A) electrical stimulation of ants from colony 1, (B) dissection of the venom glands of ants from colony 1, and (C) dissection of the venom glands of ants from colony 2. Representative gels from triplicates of each sample are shown. A total of (A) 53 spots and (B) 38 spots (circled) were excised for identification by nanoLC-ESI-QTOF MS/MS. In panel C, 57 spots (circled) were observed, but no spots were unique from panels A and B. Circles/ovals represent shared proteins between A, B, and C (yellow), A and C (red), and B and C (green). Black circles/ovals represent unique proteins not detected in the other gels.

particularly in manually dissected venom from colony 1 (Figure 5B) and the necessity not to overdevelop the gel may have resulted in less abundant proteins being below the limit of detection, particularly with Figure 5B.

Other proteins that were common to all three venoms were spots #5, #25, and #26. MS/MS analysis of spot #5, located at ca. 51 kDa, with a p*I* of ca. 5−6.5, revealed a match to the plasma glutamate carboxypeptidase, a secreted protein involved in peptide metabolism, in both spots. This protein has a reported mass of 53 kDa and p*I* of 5.87, consistent with the spot's position. Spots #25 and #26 (ca. 25 kDa, p*I* 6.5) both matched to the protein gamma-interferon-inducible lysosomal thiol reductase, a 24 kDa protein with a p*I* of 6.5, involved in endosomal disulfide reduction.

As expected, Figure 6 shows that structural proteins are more prominent in manually dissected venom rather than electrically stimulated venom. For example, actin appeared at several spots in the manually dissected venom gel, including spots #5, #63, and #74 (Figure 5B) at a mass of 36−38 kDa and p*I* of ca. 5−7 but only in an isolated spot #8 from electrically stimulated venom (Figure 5A). Another example of a structural protein that was only found in manually dissected venom is flexible cuticle protein, a transmembrane protein, which was identified from spot #69 (Figure 5B) and was not seen in the electrically stimulated venom.

To simplify this data, all protein matches were classified into six different functional categories: toxins, regulation, metabolism, structural, transport, and undetermined. Figure 6A shows

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

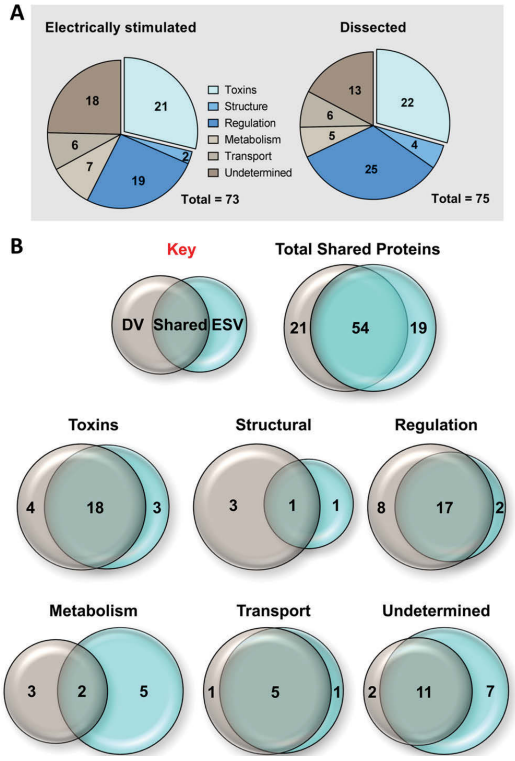


Figure 6. Functions of identified *P. clavata* venom proteins from colony 1. (A) Pie graph depicting classification of venom proteins identified from 2D-PAGE spots using nanoLC-ESI-QTOF MS/MS. Panels show venom proteins obtained by electrical stimulation (left-hand panel) and manually dissected venom (right-hand panel). (B) Area-proportional Euler plots depicting numbers of homologous and distinct venom proteins obtained by electrical stimulation (ESV; cyan circles) and manual dissection (DV; brown circles).

that the two different collection methods result in venoms that contain functionally similar proteins. The main differences were seen in the structural and undetermined protein categories. The electrically stimulated venom had less structural and more undetermined proteins compared with the manually dissected venom (Figure 6A).

As seen in Figure 6B and Table S3, each of the functional categories also had a high proportion of homologous proteins. For example, a regulatory protein identified from spot #19 in both gels was the 26S proteasome non-ATPase regulatory subunit 1 (Tables S1–S3), while an example of a toxin/allergen protein present in both venoms was PLA_2, which was seen in many spots such as in spots #12, #14, #42, and #43 (Tables S1–S3). Another toxin identified was allergen sol i III;[47] however, it was only seen in spot #37 of the electrically stimulated venom.

DISCUSSION

Venom collection is the rate-limiting step in venomic studies of small animals and restricts further investigation for drug and bioinsecticide discovery, allergenicity studies, and antivenom development.[2] The two main methods of collecting venoms are electrostimulation and manual dissection of the venom gland,[48] with the obvious disadvantage of the latter procedure being that the animal must be sacrificed. The present investigation used proteomic and peptidomic techniques to investigate the difference between these two collection methods and also identified differences in venom composition between different *P. clavata* ant colonies. While these two venom collection methods have previously been employed to investigate venom composition from other organisms such as bees,[3] cone snails,[13] and spiders,[25] the present study represents the first systematic investigation of ant venoms obtained using these two methods. Both electrical stimulation[5,49] and manual dissection[50,51] collection methods for ant venoms have been used in separate studies; however, any differences between these two methods were not established. The present investigation represents the first in-depth systematic study of both the peptidome and proteome from a single species of ant and is the first comprehensive study of *P. clavata* venom, an ant that causes one of the most painful insect sting.[48] While there was no major disparity in the numbers of peptides and proteins in venoms collected using the two methods, closer investigation revealed that there were distinct differences in the masses of these peptides and the types of proteins.

Variations in Peptidome

The two collection methods were first compared based on the MALDI-TOF mass profiles of electrically stimulated versus manually dissected whole venom. *P. clavata* venom was found to be rich in peptides with no major differences between venom collection techniques. We therefore performed C18 RP-HPLC separation of the two venoms, followed by offline analysis of the chromatographic fractions (LC-MALDI-TOF MS). This was to avoid the ion suppression effects known to occur in MALDI-TOF MS with complex venoms such as spiders, cone snails, and both poneroid and formicoid ants.[9,52] The RP-HPLC chromatograms were not significantly different, and the large number and mass distribution of the peptides below 4 kDa was similar to that reported for other poneroid ant venoms.[51−54] This mass range of peptides is not unique to ant venoms, as cone snail venom contains mostly low-molecular-weight peptides in their venoms.[55] Nevertheless, other venomous organisms such as spiders and scorpions have peptides in the higher molecular weight range of approximately 3−10 kDa.[9,56] However, comparison of individual masses within ant venom LC fractions found that each collection method resulted in venoms with unique peptide masses. As expected, there were more peptides unique to manually dissected venom compared with electrically stimulated venom (42 vs 31), probably reflecting peptides, or products of proteins, that are part of the venom gland itself or are part of the gland's machinery for toxin maturation and processing.[13,25] Other reasons for this high variability might be attributed to the pooled ants' having different diets[57] or being in different stages of development, which may give rise to ontogenic variation, as shown in several venomous animals such as wasps,[58] spiders,[59] and snakes.[60] The process of electrical stimulation of the abdomen may release not only peptides from the venom gland but also peptides and breakdown products of other glands. These may include antimicrobial and antifungal peptide contaminants from the digestive tract or hemolymph[61] and peptides and proteins that are breakdown products involved in the biosynthesis of trail pheromones and other exocrine secretions.

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

Previous work on the ant *P. clavata* has been focused on poneratoxin (PoTx), a neurotoxic peptide that was identified as the major component of the venom.[19,21,23] Poneratoxin is a 25-residue peptide with an average mass of 2756.4 Da that has been investigated for its potential development as a bioinsecticide.[19,20,62] More recently, an investigation by Johnson et al. (2016) identified three new isoforms of poneratoxin with average masses of 2814.4 Da (glyceryl-PoTx; the presumed precursor of PoTx with an additional glycine residue at position 26), 2814.4 Da (A23E PoTx), and 2783.5 Da (D22N A23V PoTx) that were isolated from ants collected from different geographical locations. Post-translational modifications to peptides are not unique to *P. clavata* venom but are common in many animal venoms such as other ants,[53,63] wasps,[64] spiders,[65] and especially marine cone snails.[53] These modifications have been attributed to the high peptide counts seen in these venoms as well as the variations between species.[29] In support of Johnson et al. (2016) we also identified a peptide of mass ca. 2784.4 Da as the most dominant peptide in venoms obtained by the two collection methods from colony 1 and also from venom from colony 2 (see Figure 2 for elution times). On the basis of mass, we were only able to identify PoTx from venom obtained from colony 2. This further supports the idea that there are several poneratoxin homologues in *P. clavata* geographical variants and within venom from a single colony.

Variations in Proteome

The *P. clavata* proteome seems to be on a similar level of complexity, in terms of spot numbers, as other hymenopteran venoms[3,66] and other poneroid ants.[52,67] However, *P. clavata* venom is less complex, in terms of protein numbers and mass ranges, than that of other venomous organisms such as cone snails,[68] snakes,[69,70] and spiders, which have up to 300 protein spots.[69] Using a 2D-PAGE approach, there was also a large number of protein spots that were shared between the two collection methods; for example, spots #54 and #39 were resolved in the same position (mass and p*I*) in both electrically stimulated and manually dissected venoms. In contrast, there were several spots that were unique to each collection method: spot #29 in electrically stimulated venom and #14 manually dissected venom. While manually dissected venom was expected to have more proteins due to likely contamination of venom gland proteins,[3,29] a number of high-abundance proteins may have obscured low-abundance ones.[66] This included a number of high-intensity spots on the acidic end of the manually dissected venom gel at ∼20 kDa, later identified as PLA_2 homologues.

nanoLC-ESI-QTOF MS-MS analysis of 2D gel spots identified several proteins within the venom using both collection methods. The most abundant proteins from electrically stimulated and manually dissected venoms were identified as PLA_2 homologues. This is a common venom toxin reported in wasp,[71] bee,[3] snake,[74] and ant[44,52,72] venom as well as being a major allergen in hymenopteran[73] venoms. Its activity involves disruption of phospholipid membranes leading to pore formation, cell lysis, cardiac dysfunction, and ultimately death of the organism.[71] Other stinging hymenoptera have also been found to have high levels of PLA_2 that account for at least some their toxic activity.[75] For example, 12% of the dry weight of bumblebee venom is PLA_2.[75] The lipase property of phospholipases may also facilitate further spreading of venom through the host tissue.[13] The allergen sol i III, originally identified in the venom of the ant *Solenopsis invicta*, was also present in *P. clavata* venom, and may contribute to the overall toxicity, particularly allergic responses,[47] despite the fact that these have not been widely reported with *P. clavata* envenomation. Because there were no other proteins with potential toxic activity identified, it appears that the toxic effects of *P. clavata* venom are the result of poneratoxin as suggested by Johnson,[16] perhaps in combination with PLA_2 isoforms. Nevertheless, the sting by *P. clavata* does not just involve a very painful sting but also other systemic effects such as hemolysis[76] and tonic-clonic seizures that have been observed in mice that ultimately lead to death at relatively low doses (6 mg/kg).[77] The venom components that cause these additional activities remain unidentified.

The 2D gels of *P. clavata* venom obtained in this study contained several spots that were identified as PLA_2 proteins, an observation reported for other venoms, including cone snails, snakes, and wasps.[29,73,74] Interestingly, the PLA_2 proteins have a range of different p*I* values and masses. Variance in the p*I* of PLA_2 homologues present in *P. clavata* venom can be explained by differences in the amino acid sequence of PLA_2 subunits, resulting in a range of basic and acidic PLA_2 proteins, as previously described in hymenopterans and snakes.[73,74] The variability in mass can be explained by homology with multimers of PLA_2 subunits (often with different individual subunit masses) or monomer PLA_2 proteins, also with varying degrees of glycosylation on individual subunits.[74] Despite the coverage scores of the matched proteins not being particularly high, it can be safely assumed that the protein matches are correct, as this enzyme is known to be highly conserved within *Hymenoptera*[73] and has also been reported in *P. clavata* using biochemical assays.[76,77]

Although peptides perform a significant role in venom functions, proteins can act as carriers, neurotoxins, or derivative enzymes such as proteases to help in the maturation and function of the peptides.[13] These proteins may also have synergistic effects on the toxic components of venoms.[65] Our study identified several manually dissected venom proteins that are located intracellularly, and we observed that electrically stimulated venom had fewer of these. In particular, these were structural proteins such as actin and cellular transport proteins such as myosin. This has also been reported in previous work that compared these two techniques using other venoms such as that of bees and marine cone snails.[66,78] It has therefore been claimed that electrically stimulated venom is more representative of the injected venom compared with manually dissected venom.[3] The intracellular proteins are usually thought of as artifacts of dissection; however, it has been proposed that they might be cosecreted into the venom gland along with a toxin.[29,65] In support, we saw some structural proteins in electrically stimulated venom as well.

The relative proportion of proteins allocated to each category in this study was not markedly different to those of other venoms, with the majority characterized as intracellular proteins.[3,29] There was also a large number of proteins in the unknown category, a recurring issue in most venomics studies. This is due to the lack of genomic and transcriptomic bioinformatics data available for a number of venoms, especially those of ants, as well as the added problem of low abundance "masking" due to high abundance proteins.[25,65,79]

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

Intercolony Variations

There were several differences between the venom from ants of colony 1 and those of colony 2 in both the peptide and protein numbers and protein composition. This has been previously observed with other ant species such as *Odontomachus haematodus*,[80] *Dinoponera quadriceps*,[53] and other venomous organisms including marine cone snails[11] and wasps.[64] The reason for these marked differences has been attributed to different factors such as age, size, differential gene expression, altered post-translational modifications, and, in the case of ants, genetic polymorphisms that result from the queens mating with different males during their lifetime.[64,81] It has also been previously shown that although the organisms are from the same region, as was the case with these colonies, slight environmental differences between the two colonies may explain the differences seen.[53,64] For example, for a predatory ant like *P. clavata*, prey diversity and abundance, as well as other factors such as peptide composition during venom regeneration and the overall health of individual ants may affect peptide levels and expression.[53]

CONCLUDING REMARKS

The present study has shown that venom obtained by manual venom gland dissection reveals a broader number of peptides and proteins than electrically stimulated *P. clavata* venom. From the 2D-PAGE analysis, approximately 72−74% of the proteins are shared by the two collection methods. It was found that the manually dissected venom contained a number of additional regulatory proteins, while the electrically stimulated venom had a number of undetermined proteins. However, there is no evidence suggesting that these additional undetermined and regulatory proteins are toxic. Unexpectedly, the overlap in peptide mass matches (approximately 16−19%) was far less than what was observed from protein matches. Although extracting venom via electrical stimulation is the preferred method of collection, as it is not destructive to the ants and contains all of the major protein components, this work collected electrically stimulated venom by individually milking ants, which was not an efficient method, and future work should utilize a mass milking technique such as that described by Eskridge et al. (1981).[82] It must be noted that the peptide composition differs markedly from that of manually dissected venom, and both methods may need to be used to obtain a more accurate representation of venom contents. This venom complexity is further increased when venoms from different colonies are compared. This highlights the diversity of ant venom proteins and peptides that can be characterized between colonies, and by using different collection techniques, for the discovery of potential therapeutic or insecticidal leads.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.6b00948.

Supplementary Table S1: Protein assignment of *P. clavata* dissected venom 2D-PAGE spots identified by nanoLC-ESI-QTOF MS/MS. Supplementary Table S2: Protein assignment of *P. clavata* electrically stimulated venom 2D-PAGE spots identified by nanoLC-ESI-QTOF MS/MS. Supplementary Table S3: *P. clavata* proteins identified from 2D-PAGE spots of electrically stimulated and manually dissected venom. Table shows proteins identified by each method as well as proteins unique to each collection method. Supplementary Figure S1: Annotated MS spectra of peptides from manually dissected P. clavata 2D gel spot 22 used to match against hymenopteran proteins. Supplementary Figure S2: Annotated MS spectra of peptides from manually dissected P. clavata 2D gel spot 58 used to match against hymenopteran proteins. Supplementary Figure S3: Annotated MS spectra of peptides from electrically stimulated P. clavata 2D gel spot 26 used to match against hymenopteran proteins. Supplementary Figure S4: Annotated MS spectra of peptides from electrically stimulated P. clavata 2D gel spot 45 used to match against hymenopteran proteins. (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel: (+61 2) 9514 2230. Fax: (+61 2) 9514 1656. E-mail: Graham.Nicholson@uts.edu.au.

ORCID

Graham M. Nicholson: 0000-0002-4277-4296

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support for this study was provided by an Australian Postgraduate Award to S.R.A. and the programme convergence 2007−2013, *Région Guyane* from the European community (Bi-Appli, 115/SGAR-DE-2011/052274) to A.D. This work has also benefited from an "Investissement d'Avenir" grant managed by the *Agence Nationale de la Recherche* (CEBA, ref ANR-10-LABX-25-01).

ABBREVIATIONS

ACN, acetonitrile; C7BzO, 3-(4-heptyl)phenyl-3-hydroxypropyl) dimethylammoniopropanesulfonate; 1,5-DAN, 1,5-diaminonapthalene; CHCA, α-cyano-4-hydroxycinnamic acid; FA, ferulic acid; HGD, Hymenopteran Genome Database; IEF, isoelectric focusing; ionic CHCA, *N-tert*-butyl-*N*-isopropyl-*N*-methylammonium α-cyano-4-hydroxycinnamate; IPG, immobilized pH gradient; LC-ESI-QTOF MS, liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry; MALDI-TOF MS, matrix-assisted laser-desorption ionization time-of-flight mass spectrometry; PLA_2, phospholipase A_2; RP-HPLC, reverse-phase high-pressure liquid chromatography; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid

REFERENCES

(1) *AntWeb*, 2014. http://www.antweb.org (accessed on 26 August 2016).

(2) Aili, S. R.; Touchard, A.; Escoubas, P.; Padula, M. P.; Orivel, J.; Dejean, A.; Nicholson, G. M. Diversity of peptide toxins from stinging ant venoms. *Toxicon* **2014**, *92*, 166−78.

(3) Li, R.; Zhang, L.; Fang, Y.; Han, B.; Lu, X.; Zhou, T.; Feng, M.; Li, J. Proteome and phosphoproteome analysis of honeybee (Apis mellifera) venom collected from electrical stimulation and manual extraction of the venom gland. *BMC Genomics* **2013**, *14*, 766.

(4) Ward, P. S. Taxonomy, phylogenetics, and evolution. *Ant Ecology* **2009**, *1*, 3−17.

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

(5) Pluzhnikov, K. A.; Kozlov, S. A.; Vassilevski, A. A.; Vorontsova, O. V.; Feofanov, A. V.; Grishin, E. V. Linear antimicrobial peptides from Ectatomma quadridens ant venom. *Biochimie* **2014**, *107* (Pt B), 211−5.
(6) Touchard, A.; Koh, J.; Aili, S. R.; Dejean, A.; Nicholson, G. M.; Orivel, J.; Escoubas, P. The complexity and structural diversity of ant venom peptidomes is revealed by mass spectrometry profiling. *Rapid Commun. Mass Spectrom.* **2015**, *29* (5), 385−96.
(7) Touchard, A.; Aili, S.; Fox, E.; Escoubas, P.; Orivel, J.; Nicholson, G.; Dejean, A. The biochemical toxin arsenal from ant venoms. *Toxins* **2016**, *8* (1), 30.
(8) Wood, C. L.; Hoffman, D. R. Two-dimensional polyacrylamide gel electrophoresis of hymenoptera venom and venom sac extracts. *Toxicon* **1983**, *21* (2), 291−9.
(9) Palagi, A.; Koh, J. M.; Leblanc, M.; Wilson, D.; Dutertre, S.; King, G. F.; Nicholson, G. M.; Escoubas, P. Unravelling the complex venom landscapes of lethal Australian funnel-web spiders (Hexathelidae: Atracinae) using LC-MALDI-TOF mass spectrometry. *J. Proteomics* **2013**, *80*, 292−310.
(10) Windley, M. J.; Herzig, V.; Dziemborowicz, S. A.; Hardy, M. C.; King, G. F.; Nicholson, G. M. Spider-venom peptides as bioinsecticides. *Toxins* **2012**, *4* (3), 191−227.
(11) Dutertre, S.; Biass, D.; Stöcklin, R.; Favreau, P. Dramatic intraspecimen variations within the injected venom of Conus consors: An unsuspected contribution to venom diversity. *Toxicon* **2010**, *55* (8), 1453−62.
(12) Dutertre, S.; Jin, A. H.; Kaas, Q.; Jones, A.; Alewood, P. F.; Lewis, R. J. Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. *Mol. Cell. Proteomics* **2013**, *12* (2), 312−29.
(13) Möller, C.; Vanderweit, N.; Bubis, J.; Marí, F. Comparative analysis of proteases in the injected and dissected venom of cone snail species. *Toxicon* **2013**, *65* (0), 59−67.
(14) Peigneur, S.; Van Der Haegen, A.; Möller, C.; Waelkens, E.; Diego-García, E.; Marí, F.; Naudé, R.; Tytgat, J. Unraveling the peptidome of the South African cone snails Conus pictus and Conus natalis. *Peptides* **2013**, *41*, 8−16.
(15) Larson, H. K.; Goffredi, S. K.; Parra, E. L.; Vargas, O.; Pinto-Tomas, A. A.; McGlynn, T. P. Distribution and dietary regulation of an associated facultative Rhizobiales-related bacterium in the omnivorous giant tropical ant, *Paraponera clavata*. *Naturwissenschaften* **2014**, *101* (5), 397−406.
(16) Johnson, S. R.; Rikli, H. G.; Schmidt, J. O.; Evans, M. S. A reexamination of poneratoxin from the venom of the bullet ant Paraponera clavata. *Peptides* **2016**, DOI: 10.1016/j.peptides.2016.05.012.
(17) Bosmia, A. N.; Griessenauer, C. J.; Haddad, V.; Shane Tubbs, R. Ritualistic envenomation by bullet ants among the Sateré-Mawé Indians in the Brazilian Amazon. *Wilderness Environ. Med.* **2015**, *26* (2), 271−273.
(18) Torres, A.; Quinet, Y.; Havt, A.; Rádis-Baptista, G.; Martins, A. Molecular Pharmacology and Toxinology of Venom from Ants. *An Integrated View of the Molecular Recognition and Toxinology−From Analytical Procedures to Biomedical Applications*; InTech: Croatia, 2013; pp 207−222.
(19) Piek, T.; Hue, B.; Mantel, P.; Terumi, N.; Schmidt, J. O. Pharmacological characterization and chemical fractionation of the venom of the ponerine ant, Paraponera clavata (F.). *Comp. Biochem. Physiol., C: Comp. Pharmacol.* **1991**, *99* (3), 481−6.
(20) Szolajska, E.; Poznanski, J.; Ferber, M. L.; Michalik, J.; Gout, E.; Fender, P.; Bailly, I.; Dublet, B.; Chroboczek, J. Poneratoxin, a neurotoxin from ant venom. Structure and expression in insect cells and construction of a bio-insecticide. *Eur. J. Biochem.* **2004**, *271* (11), 2127−36.
(21) Piek, T.; Duval, A.; Hue, B.; Karst, H.; Lapied, B.; Mantel, P.; Nakajima, T.; Pelhate, M.; Schmidt, J. O. Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata*. *Comp. Biochem. Physiol., C: Comp. Pharmacol.* **1991**, *99* (3), 487−95.
(22) Piek, T. Neurotoxins from venoms of the Hymenoptera–twenty-five years of research in Amsterdam. *Comp. Biochem. Physiol., C: Comp. Pharmacol.* **1990**, *96* (2), 223−33.
(23) Duval, A.; Malécot, C. O.; Pelhate, M.; Piek, T. Poneratoxin, a new toxin from an ant venom, reveals an interconversion between two gating modes of the Na channels in frog skeletal muscle fibres. *Pfluegers Arch.* **1992**, *420* (3−4), 239−47.
(24) Orivel, J.; Redeker, V.; Le Caer, J. P.; Krier, F.; Revol-Junelles, A. M.; Longeon, A.; Chaffotte, A.; Dejean, A.; Rossier, J. Ponericins, new antibacterial and insecticidal peptides from the venom of the ant Pachycondyla goeldii. *J. Biol. Chem.* **2001**, *276* (21), 17823−29.
(25) Garb, J. E. Extraction of venom and venom gland micro-dissections from spiders for proteomic and transcriptomic analyses. *J. Visualized Exp.* **2014**, *93*, e51618.
(26) da Silveira, R. B.; Chaim, O. M.; Mangili, O. C.; Gremski, W.; Dietrich, C. P.; Nader, H. B.; Veiga, S. S. Hyaluronidases in Loxosceles intermedia (Brown spider) venom are endo-β-N-acetyl-d-hexosaminidases hydrolases. *Toxicon* **2007**, *49* (6), 758−768.
(27) Mueller, U.; Reisman, R.; Wypych, J.; Elliott, W.; Steger, R.; Walsh, S.; Arbesman, C. Comparison of vespid venoms collected by electrostimulation and by venom sac extraction. *J. Allergy Clin. Immunol.* **1981**, *68* (4), 254−61.
(28) Inceoglu, B.; Lango, J.; Jing, J.; Chen, L.; Doymaz, F.; Pessah, I. N.; Hammock, B. D. One scorpion, two venoms: Prevenom of Parabuthus transvaalicus acts as an alternative type of venom with distinct mechanism of action. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (3), 922−927.
(29) Leonardi, A.; Biass, D.; Kordiš, D.; Stöcklin, R.; Favreau, P.; Križaj, I. Conus consors snail venom proteomics proposes functions, pathways, and novel families involved in its venomic system. *J. Proteome Res.* **2012**, *11* (10), 5046−58.
(30) Gopalakrishnakone, P.; Cheah, J.; Gwee, M. C. Black scorpion (Heterometrus longimanus) as a laboratory animal: maintenance of a colony of scorpion for milking of venom for research, using a restraining device. *Lab. Anim.* **1995**, *29* (4), 456−8.
(31) Oukkache, N.; Chgoury, F.; Lalaoui, M.; Cano, A. A.; Ghalim, N. Comparison between two methods of scorpion venom milking in Morocco. *J. Venomous Anim. Toxins Incl. Trop. Dis.* **2013**, *19* (1), 1−5.
(32) Shulman, S.; Langlois, C.; Arbesman, C. E. The allergic response to stinging insects: I. Preparation of extracts and their biochemical characterization. *J. Allergy* **1964**, *35* (5), 446−63.
(33) O'Connor, R.; Rosenbrook, W., Jr.; Erickson, R. Disc electrophoresis of hymenoptera venoms and body proteins. *Science* **1964**, *145* (3638), 1320−21.
(34) Benton, A. W.; Morse, R. A.; Stewart, J. D. Venom collection from honey bees. *Science* **1963**, *142* (3589), 228−30.
(35) O'Connor, R.; Rosenbrook, W.; Erickson, R. Hymenoptera: pure venom from bees, wasps, and hornets. *Science* **1963**, *139* (3553), 420.
(36) Hsiang, H. K.; Elliott, W. B. Differences in honey bee (Apis mellifera) venom obtained by venom sac extraction and electrical milking. *Toxicon* **1975**, *13* (2), 145−48.
(37) Wood, C. L.; Timmons, B. E. t.; Hoffman, D. R. Allergens in hymenoptera venoms. X. Vespid venoms versus venom sac extracts: comparison by two-dimensional polyacrylamide gel electrophoresis. *Ann. Allergy* **1983**, *51* (4), 441−45.
(38) Jaskolla, T. W.; Karas, M.; Roth, U.; Steinert, K.; Menzel, C.; Reihs, K. Comparison between vacuum sublimed matrices and conventional dried droplet preparation in MALDI-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1104−14.
(39) Escoubas, P.; Sollod, B.; King, G. F. Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon* **2006**, *47* (6), 650−63.
(40) Jobbins, S. E.; Hill, C. J.; D'Souza-Basseal, J. M.; Padula, M. P.; Herbert, B. R.; Krockenberger, M. B. Immunoproteomic approach to elucidating the pathogenesis of cryptococcosis caused by Cryptococcus gattii. *J. Proteome Res.* **2010**, *9* (8), 3832−41.
(41) Raymond, B. B. A.; Tacchi, J. L.; Jarocki, V. M.; Minion, F. C.; Padula, M. P.; Djordjevic, S. P. P159 from Mycoplasma hyopneumo-

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

niae binds porcine cilia and heparin and is cleaved in a manner akin to ectodomain shedding. *J. Proteome Res.* **2013**, *12* (12), 5891−903.
(42) Shevchenko, A.; Sunyaev, S.; Loboda, A.; Shevchenko, A.; Bork, P.; Ens, W.; Standing, K. G. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Anal. Chem.* **2001**, *73* (9), 1917−26.
(43) Hoffman, D. R. Allergens in bee venom: III. Identification of allergen B of bee venom as an acid phosphatase. *J. Allergy Clin. Immunol.* **1977**, *59* (5), 364−66.
(44) Hoffman, D. R.; Sakell, R. H.; Schmidt, M. Sol i 1, the phospholipase allergen of imported fire ant venom. *J. Allergy Clin. Immunol.* **2005**, *115* (3), 611−16.
(45) Lewis, J. C.; Day, A. J.; De la Lande, I. S. Phospholipase A in the venom of the Australian bulldog ant Myrmecia pyriformis. *Toxicon* **1968**, *6* (2), 109−12.
(46) Zalat, S.; Schmidt, J.; Moawad, T. I. Lipase and phospholipase activities of Hymenoptera venoms (wasps and ants). *Egyptian J. Biol.* **2003**, *5*, 138−47.
(47) Schmidt, M.; McConnell, T. J.; Hoffman, D. R. Immunologic characterization of the recombinant fire ant venom allergen Sol i 3. *Allergy* **2003**, *58* (4), 342−49.
(48) Dossey, A. Chemical Defenses of Insects: A Rich Resource for Chemical Biology in the Tropics. In *Chemical Biology of the Tropics*; Vivanco, J. M., Weir, T., Eds.; Springer: Berlin, 2011; Vol. *8*, pp 27−57.
(49) Hink, W. F.; Pappas, P. W.; Jaworski, D. C. Partial biochemical characterization of venom from the ant, Pseudomyrmex triplarinus. *Toxicon* **1994**, *32* (7), 763−72.
(50) Touchard, A.; Labrière, N.; Roux, O.; Petitclerc, F.; Orivel, J.; Escoubas, P.; Koh, J. M.; Nicholson, G. M.; Dejean, A. Venom toxicity and composition in three Pseudomyrmex ant species having different nesting modes. *Toxicon* **2014**, *88*, 67−76.
(51) Touchard, A.; Dauvois, M.; Arguel, M.-J.; Petitclerc, F.; Leblanc, M.; Dejean, A.; Orivel, J.; Nicholson, G. M.; Escoubas, P. Elucidation of the unexplored biodiversity of ant venom peptidomes via MALDI-TOF mass spectrometry and its application for chemotaxonomy. *J. Proteomics* **2014**, *105*, 217−31.
(52) Aili, S. R.; Touchard, A.; Koh, J. M. S.; Dejean, A.; Orivel, J.; Padula, M. P.; Escoubas, P.; Nicholson, G. M. Comparisons of protein and peptide complexity in poneroid and formicoid ant venoms. *J. Proteome Res.* **2016**, *15* (9), 3039−54.
(53) Cologna, C. T.; Cardoso, J. d. S.; Jourdan, E.; Degueldre, M.; Upert, G.; Gilles, N.; Uetanabaro, A. P.; Costa Neto, E. M.; Thonart, P.; de Pauw, E.; Quinton, L. Peptidomic comparison and characterization of the major components of the venom of the giant ant Dinoponera quadriceps collected in four different areas of Brazil. *J. Proteomics* **2013**, *94*, 413−22.
(54) Johnson, S. R.; Copello, J. A.; Evans, M. S.; Suarez, A. V. A biochemical characterization of the major peptides from the venom of the giant neotropical hunting ant Dinoponera australis. *Toxicon* **2010**, *55* (4), 702−10.
(55) Davis, J.; Jones, A.; Lewis, R. J. Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides* **2009**, *30* (7), 1222−27.
(56) Abdel-Rahman, M. A.; Harrison, P. L.; Strong, P. N. Snapshots of scorpion venomics. *J. Arid Environ.* **2015**, *112*, 170−176.
(57) Pucca, M. B.; Amorim, F. G.; Cerni, F. A.; Bordon, K. d. C. F.; Cardoso, I. A.; Anjolette, F. A. P.; Arantes, E. C. Influence of post-starvation extraction time and prey-specific diet in Tityus serrulatus scorpion venom composition and hyaluronidase activity. *Toxicon* **2014**, *90*, 326−36.
(58) Uçkan, F.; Ergin, E.; Rivers, D. B.; Gençer, N. Age and diet influence the composition of venom from the endoparasitic wasp Pimpla turionellae L.(Hymenoptera: Ichneumonidae). *Arch. Insect Biochem. Physiol.* **2006**, *63* (4), 177−87.
(59) Herzig, V.; Ward, R. J.; Dos Santos, W. F. Ontogenetic changes in Phoneutria nigriventer (Araneae, Ctenidae) spider venom. *Toxicon* **2004**, *44*, 635−40.
(60) Alape-Giron, A.; Sanz, L.; Escolano, J.; Flores-Diaz, M.; Madrigal, M.; Sasa, M.; Calvete, J. J. Snake venomics of the lancehead pitviper Bothrops asper: geographic, individual, and ontogenetic variations. *J. Proteome Res.* **2008**, *7* (8), 3556−71.
(61) Bulet, P.; Stocklin, R. Insect antimicrobial peptides: structures, properties and gene regulation. *Protein Pept. Lett.* **2005**, *12* (1), 3−11.
(62) Hendrich, A. B.; Mozrzymas, J. W.; Konopińska, D.; Scuka, M. The effect of poneratoxin on neuromuscular transmission in the rat diaphragm. *Cell. Mol. Biol. Lett.* **2001**, *7* (2), 195−202.
(63) Wiese, M. D.; Chataway, T. K.; Davies, N. W.; Milne, R. W.; Brown, S. G.; Gai, W. P.; Heddle, R. J. Proteomic analysis of Myrmecia pilosula (jack jumper) ant venom. *Toxicon* **2006**, *47* (2), 208−17.
(64) Dias, N. B.; de Souza, B. M.; Gomes, P. C.; Palma, M. S. Peptide diversity in the venom of the social wasp Polybia paulista (Hymenoptera): a comparison of the intra- and inter-colony compositions. *Peptides* **2014**, *51*, 122−30.
(65) Yuan, C.; Jin, Q.; Tang, X.; Hu, W.; Cao, R.; Yang, S.; Xiong, J.; Xie, C.; Xie, J.; Liang, S. Proteomic and peptidomic characterization of the venom from the Chinese bird spider, Ornithoctonus huwena Wang. *J. Proteome Res.* **2007**, *6* (7), 2792−801.
(66) Sookrung, N.; Wong-din-Dam, S.; Tungtrongchitr, A.; Reamtong, O.; Indrawattana, N.; Sakolvaree, Y.; Visitsunthorn, N.; Manuyakorn, W.; Chaicumpa, W. Proteome and allergenome of Asian wasp, Vespa affinis, venom and IgE reactivity of the venom components. *J. Proteome Res.* **2014**, *13* (3), 1336−44.
(67) Pessoa, W. F. B.; Silva, L. C. C.; de Oliveira Dias, L.; Delabie, J. H. C.; Costa, H.; Romano, C. C. Analysis of protein composition and bioactivity of Neoponera villosa venom (Hymenoptera: Formicidae). *Int. J. Mol. Sci.* **2016**, *17* (4), 513.
(68) Safavi-Hemami, H.; Siero, W. A.; Gorasia, D. G.; Young, N. D.; MacMillan, D.; Williamson, N. A.; Purcell, A. W. Specialisation of the venom gland proteome in predatory cone snails reveals functional diversification of the conotoxin biosynthetic pathway. *J. Proteome Res.* **2011**, *10* (9), 3904−3919.
(69) Igci, N.; Demiralp, D. O. A preliminary investigation into the venom proteome of Macrovipera lebetina obtusa (Dwigubsky, 1832) from Southeastern Anatolia by MALDI-TOF mass spectrometry and comparison of venom protein profiles with Macrovipera lebetina lebetina (Linnaeus, 1758) from Cyprus by 2D-PAGE. *Arch. Toxicol.* **2012**, *86* (3), 441−451.
(70) Serrano, S. M.; Shannon, J. D.; Wang, D.; Camargo, A.; Fox, J. W. A multifaceted analysis of viperid snake venoms by two-dimensional gel electrophoresis: An approach to understanding venom proteomics. *Proteomics* **2005**, *5* (2), 501−510.
(71) Ho, C.-L.; Ko, J.-L. Purification and characterization of a lethal protein with phospholipase A1 activity from the hornet (Vespa basalis) venom. *Biochim. Biophys. Acta, Lipids Lipid Metab.* **1988**, *963* (3), 414−22.
(72) Torres, A. F. C.; Huang, C.; Chong, C.-M.; Leung, S. W.; Prieto-da-Silva, A. R. B.; Havt, A.; Quinet, Y. P.; Martins, A. M. C.; Lee, S. M. Y.; Rádis-Baptista, G. Transcriptome analysis in venom gland of the predatory giant ant Dinoponera quadriceps: insights into the polypeptide toxin arsenal of Hymenopterans. *PLoS One* **2014**, *9* (1), e87556.
(73) dos Santos, L. D.; da Silva Menegasso, A. R.; dos Santos Pinto, J. R.; Santos, K. S.; Castro, F. M.; Kalil, J. E.; Palma, M. S. Proteomic characterization of the multiple forms of the PLAs from the venom of the social wasp Polybia paulista. *Proteomics* **2011**, *11* (8), 1403−12.
(74) Manjunatha Kini, R. Excitement ahead: structure, function and mechanism of snake venom phospholipase A 2 enzymes. *Toxicon* **2003**, *42* (8), 827−840.
(75) Xin, Y.; Choo, Y. M.; Hu, Z.; Lee, K. S.; Yoon, H. J.; Cui, Z.; Sohn, H. D.; Jin, B. R. Molecular cloning and characterization of a venom phospholipase A2 from the bumblebee *Bombus ignitus*. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **2009**, *154* (2), 195−202.
(76) Schmidt, J. O.; Blum, M. S.; Overal, W. L. Hemolytic activities of stinging insect venoms. *Arch. Insect Biochem. Physiol.* **1983**, *1* (2), 155−60.

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339−1351

(77) Schmidt, J. O.; Blum, M. S.; Overal, W. L. Comparative lethality of venoms from stinging Hymenoptera. *Toxicon* **1980**, *18* (4), 469–74.
(78) Safavi-Hemami, H.; Young, N. D.; Williamson, N. A.; Purcell, A. W. Proteomic interrogation of venom delivery in marine cone snails: novel insights into the role of the venom bulb. *J. Proteome Res.* **2010**, *9* (11), 5610–19.
(79) Bouzid, W.; Verdenaud, M.; Klopp, C.; Ducancel, F.; Noirot, C.; Vétillard, A. De Novo sequencing and transcriptome analysis for Tetramorium bicarinatum: a comprehensive venom gland transcriptome analysis from an ant species. *BMC Genomics* **2014**, *15* (1), 987.
(80) Touchard, A.; Dejean, A.; Escoubas, P.; Orivel, J. Intraspecific variations in the venom peptidome of the ant Odontomachus haematodus (Formicidae: Ponerinae) from French Guiana. *Journal of Hymenoptera Research* **2015**, *47*, 87.
(81) Prator, C. A.; Murayama, K. M.; Schulz, J. R. Venom variation during prey capture by the cone snail, Conus textile. *PLoS One* **2014**, *9* (6), e98991.
(82) Eskridge, E. M.; Elliott, W. B.; Elliott, A. H.; Eskridge, P. B.; Doerr, J. C.; Schneller, N.; Reisman, R. E. Adaptation of the electrical stimulation procedure for the collection of vespid venoms. *Toxicon* **1981**, *19* (6), 893–97.

DOI: 10.1021/acs.jproteome.6b00948
J. Proteome Res. 2017, 16, 1339–1351

Supporting Information for Publication

Title of Paper

Combined peptidomic and proteomic analysis of electrically stimulated and manually dissected venom from the South American bullet ant *Paraponera clavata*

Authors

Samira R. Aili[a], Axel Touchard[b], Frédéric Petitclerc[b], Alain Dejean[b,c], Jérôme Orivel[b], Matthew P. Padula[d], Pierre Escoubas[e], Graham M. Nicholson[a]*

Affiliations

[a]Neurotoxin Research Group, School of Life Sciences, University of Technology Sydney, NSW 2007, Australia

[b]CNRS, UMR Ecologie des Forêts de Guyane (EcoFoG), AgroParisTech, Cirad, INRA, Université des Antilles, Université de Guyane, Université des Antilles, 97310 Kourou, France

[c] Ecolab, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France

[d]Proteomics Core Facility, Faculty of Science, University of Technology Sydney, NSW 2007, Australia

[e]VenomeTech, 473 Route des Dolines — Villa 3, Valbonne 06560, France

Corresponding Author

*Prof. Graham M. Nicholson, Neurotoxin Research Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW 2067, Australia. Phone: (+61) 9514 2230. Fax: (+61) 9514 1656. E-mail: Graham.Nicholson@uts.edu.au

Table of Contents

Table S1. Protein assignment of *P. clavata* dissected venom 2D-PAGE spots identified by nanoLC-ESI-QTOF MS/MS

Spot N°	Protein ID	Organism	Accession N°	Predicted MW (Da) & pI	Observed MW (Da) & pI	Score	Coverage	Peptides used for Identification
5	Plasma glutamate carboxypeptidase	*Camponotus floridanus*	EFN65651.1	53800 5.87	51000 5.5–6.5	262#	8%	ETVEYR RAAELGAVAVLLR ESATLLQPR IVVYNQK
5	Actin, muscle	*Camponotus floridanus*	EFN60576.1	41700 5.29	51000 5.5–6.5	159*	32%	AGFAGDDAPRAVFPSIVGRPR DSYVGDEAQSKRGILTLK VAPEEHPVLLTEAPLNPK DLTDYLMK GYSFTTTAER SYELPDGQVITIGNER EITALAPSTIK IIAPPER QEYDESGPGIVHR
63					17000 5.0	130*	27%	AGFAGDDAPRAVFPSIVGRPR VAPEEHPVLLTEAPLNPK DLTDYLMK GYSFTTTAER SYELPDGQVITIGNER EITALAPSTIK IIAPPER
74					48500 5.0–6.0	159*	32%	AGFAGDDAPRAVFPSIVGRPR DSYVGDEAQSKRGILTLK VAPEEHPVLLTEAPLNPK DLTDYLMK GYSFTTTAER SYELPDGQVITIGNER EITALAPSTIK IIAPPER QEYDESGPGIVHR

1

12	PREDICTED: phospholipase A2-like	*Megachile rotundata*	XP_003706900.1	27400 6.62	20000 5.0	93*	9%	GGLLNNGIFTR TYFNILKPQCFK
14					20000 3.0–3.5	58*	8%	LLNNGIFTR TYFNILKPQCFK
17					16000 4.8	30.1^	5%	TYFNILKPQCFK
40					20000 4.95	53*	9%	GGLLNNGIFTR TYFNILKPQCFK
41					20000 4.7	50*	8%	LLNNGIFTR TYFNILKPQCFK
42					20000 4.6	108*	12%	GGLLNNGIFTR TYFNILKPQCFK
43					24000 4.4–4.6	90*	8%	LLNNGIFTR TYFNILKPQCFK
44					24000 4.8	45*	9%	LLNNGIFTR TYFNILKPQCFK
45					20000 3.5–4.1	52*	9%	GGLLNNGIFTR TYFNILKPQCFK
77					26000 4.8	37*	8%	LLNNGIFTR TYFNILKPQCFK
76					24000 3.5–4.1	53*	8%	LLNNGIFTR TYFNILKPQCFK
78					26000 4.3–4.5	78*	8%	LLNNGIFTR TYFNILKPQCFK
44	Phospholipase A2, partial	*Harpegnathos saltator*	EFN81056.1	12000 8.01	24000 4.8	33.8^	11%	TYFNILQPQCFK SACSCDAEFYNCLK
62					22500 5.1	41.1^	24%	TYFNILQPQCFK SACSCDAEFYNCLK
72					21000 5.1	25.9^	11%	TYFNILQPQCFK LLNNGLFTR
75					26000 3.5–4.1	37.7^	24%	TYFNILQPQCFK SACSCDAEFYNCLK
77					26000	31.4^	24%	TYFNILQPQCFK

2

					4.8			SACSCDAEFYNCLK
19	PREDICTED: uncharacterized protein	*Apis mellifera*	XP_016770221	29500 9.75	16500 5.9	69*	2%	IKTLQDALKKAK
40					20000 4.95	65*	4%	IKTLQDALKKAK
61					3000 7.6	66*	4%	IKTLQDALKKAK
72					21000 5.1	74*	4%	IKTLQDALKKAK
76					24000 3.5–4.1	73*	4%	IKTLQDALKKAK
81					37000 4.2–4.8	69*	4%	IKTLQDALKKAK
61	TP53-regulating kinase	*Acromyrmex echinatior*	EGI66742.1	27100 9.36	3000 7.6	68*	3%	KQYEQILAK
40					20000 4.95	55*	3%	KQYEQILAK
81					37000 4.2–4.8	59*	3%	KQYEQILAK
17	PREDICTED: TRAF-interacting protein-like	*Bombus impatiens*	XP_003492946.1	48900 9.09	16000 4.8	64#	2%	KLFSSKPK GNGLDDPK
62	PREDICTED: sperm flagellar protein 1-like	*Bombus impatiens*	XP_003486146.1	25700 8.94	22500 5.1	53*	3%	QISFSKPK
40	PREDICTED: apolipoprotein D-like	*Megachile rotundata*	XP_003699771.1	21000 6.29	20000 4.95	48*	5%	KPIGKSDDPK
49	Uncharacterized protein GJ16099	*Drosophila virilis*	XP_002060462.1	8003 12.18	5000 4.9	170#	35%	VVTVLSPR VPGVSLPR VATVSNPR
79	Myosin light chain alkali	*Apis cerana*	AEY59302.1	17000, 4.47	15500 4.4	77*	8%	ALNLNPTNATIEK
19	26S proteasome non-ATPase regulatory subunit 1	*Camponotus floridanus*	EFN63483.1	11100, 9.5	16500 5.9	50*	8%	VSDVNDDVR

22	T-complex protein 11 domain-containing protein	*Microplitis demolitor*	EZA44954.1	58300 5.34	15000 6.2	52*	1%	IQADGLQYTR
22	hypothetical protein X777_04060	*Cerapachys biroi*	EZA55713.1	15400 7.48	15000 6.2	22.84^	12%	GETFCQQQGTSCCSIVK
22	PREDICTED: ES1 protein homolog, mitochondrial-like	*Solenopsis invicta*	XP_011158080.1	29100 8.67	15000 6.2	20.56^	3%	SCASLLQR
25 26	Gamma-interferon-inducible lysosomal thiol reductase	*Harpegnathos saltator*	EFN80944.1	24000 6.51	26000 5.8 26500 6.0	131# 218#	9% 17%	KVDVYYESLC KTAALR GQSEVAALVVDDGSGMFR SLVHPTAK
26	PREDICTED: DNA polymerase eta-like	*Apis florea*	XP_003691455.1	80200 8.9	26500 6.0	59*	3%	NKTIVSQSR SNSNENNNLLNNK
58	Sodium-independent sulfate anion transporter	*Acromyrmex echinatior*	EGI62238.1	62800 8.56	30000 10.0	53*	1%	MTINDFR
55 65	NADPH: adrenodoxin oxidoreductase, mitochondrial	*Harpegnathos saltator*	EFN88437.1	51700 9.29	10000 10.0 17000 5.0	32* 32*	4% 4%	VCIIGAGPAGFYAAQQMLKGLN VCIIGAGPAGFYAAQQMLKGLN
70	Uncharacterized protein	*Nasonia vitripennis*	XP_008213255.1	45800 5.43	10000 3.9	52*	2%	TLIEGEEDK
66	PREDICTED: superoxide dismutase [Cu-Zn]	*Acromyrmex echinatior*	XP_011068977.1	12300 6.2	17000 7.0	36.7^	18%	AVCVLQGEPVK LACGVIGITK
67	Phosphatidylinositol N-acetylglucosaminyltransferase subunit H, partial	*Camponotus floridanus*	EFN72393.1	22100 9.28	16000 7.1	22.1^	4%	CLEYIYR
69	Flexible cuticle protein 12	*Camponotus floridanus*	EFN65074.1	16200 5.09	11000 4.2	68*	10%	QGNPNDITIVK
74	PREDICTED: elongation factor Tu-like	*Bombus impatiens*	XP_003494245	47200 5.11	48500 5.0–6.0	377#	29%	TTDVTGTIELPEGVEMVMPGDNIK GITINTSHVEYDTPTR ELLSQYDFPGDDTPVIR QVGVPYIIVFLNK AGENVGVLLR QVGVPYIIVFLNK

								HTPFFK TVGAGVVAK FESEVYILSK EHILLGR
75 76	Heat shock protein beta-1	*Acromyrmex echinatior*	EGI64873.1	20700 6.8	26000 3.5–4.1 24000 3.5–4.1	42.9^ 64*	9% 9%	DGVLTVEAPLPAIGSGEK DGVLTVEAPLPAIGSGEK
75 77 78	Myosin regulatory light chain 2	*Acromyrmex echinatior*	EGI59363.1	26400 4.88	26000 3.5–4.1 26000 4.8 26000 4.3–4.5	122* 165* 127*	23% 24% 23%	EAFQLMDADKDGIIGK AAFDSVGR MSGSGADDDDVVINAFK EVNDAFDQMYIDDK MSGSGADDDDVVINAFK EAFQLMDADKDGIIGK AAFDSVGR MSGSGADDDDVVINAFK EVNDAFDQMYIDDK
78	PREDICTED: calcium-binding mitochondrial carrier protein	*Bombus impatiens*	XP_00349163 2.1	37800 9.18	26000 4.3–4.5	21*	5%	YEQIKRTIK GKPNTMVAVFK

Symbols in the score column indicate which search engine the results were obtained from (*mascot, #MS blast, ^PEAKS)

5

Table S2. Protein assignment of *P. clavata* electrically stimulated venom 2D-PAGE spots identified by nanoLC-ESI-QTOF MS/MS.

Spot N°	Protein ID	Organism	Accession N°	Predicted MW (Da) & pI	Obs MW (Da) &pI	Score	Coverage	Peptides used for Identification
3	Angiotensin-converting enzyme	*Bombus impatiens*	XP_003486575.1	74400 5.1	75000 5.3–6.0	58*	4%	YGEDIVSKDGPIPAHLLGN YHIIADVEYLR
4	Uncharacterized protein	*Acromyrmex echinatior*	EGI66098.1	51900 6.2	61000 5.3–6.0	35*	3%	QAYKSSTACAGCIVSSA
4	hypothetical protein SINV_06838, partial	*Solenopsis invicta*	EFZ19030.1	57600 6.1	61000 5.3–6.0	248#	7%	YVSSGVGKALEEC
4	Sodium- and chloride-dependent glycine transporter	*Camponotus floridanus*	E2A7A3	163377 8.9	61000 5.3–6.0	51*	2%	LKFQMKENLR HIIGCLRVEN
5	Glucosylceramidase	*Acromyrmex echinatior*	EGI59542.1	55300 9.1	51000 5.5–6.3	100#	4%	LLGFGGAVTDSASFGADK
5	EAI_04268 SET and MYND domain-containing protein 4	*Harpegnathos saltator*	EFN86371	50400 9.0	51000 5.5–6.3	59*	4%	EGQVGTMKSR KICLEDRLAC
5	Plasma glutamate carboxypeptidase	*Camponotus floridanus*	EFN65651.1	53800 5.9	51000 5.5–6.3	262#	8%	ESATLLQPR SRLPAACL RLVECVLRK VTVPRWV
8	Actin, muscle	*Camponotus floridanus*	EFN60576.1	41700 5.3	38000 5.2	65*	27%	AGFAGDDAPRAVFPSIVGRPR VAPEEHPVLLTEAPLNPK DLTDYLMK GYSFTTTAER SYELPDGQVITIGNER EITALAPSTIK IIAPPER
6	Enolase	*Camponotus floridanus*	EFN71651.1	47000 5.9	48500 6.4–7.0	45*	4%	AAVPSGASTGVHEALELR
7	Transmembrane protease, serine 9	*Acromyrmex echinatior*	EGI59119.1	44200 5.5	40000 5.5–6.0	106#	5%	MAGLVDT NTLGVLVGEF TCENY

8 12 14 23 42 43 46	PREDICTED: phospholipase A2-like	*Megachile rotundata*	XP_003706900.1	27400 6.6	38000 5.2 19000 5.35 19000 3.1 10000 7.2 19000 4.9 22000 4.9 22000 4.3	57* 93* 58* 57* 108* 90* 98*	8% 8% 8% 19% 12% 8% 8%	LLNNGIFTR TYFNILKPQCFK LLNNGIFTR TYFNILKPQCFK LLNNGIFTR TYFNILKPQCFK LLNNGIFTR TYFNILKPQCFK ISILKKKTG LLNNGIFTR TYFNILKPQCFK LLNNGIFTR TYFNILKPQCFK LLNNGIFTR TYFNILKPQCFK
9 17 18 19	Phospholipase A2, partial	*Harpegnathos saltator*	EFN81056.1	12000 8.0	35000 5.2 15200 5.1 15200 5.9 15200 6.1	36.86^ 43.93^ 33.02^ 38.78^	24% 11% 11% 11%	TYFNILQPQCFK SACSCDAEFYNCLK TYFNILQPQCFK TYFNILQPQCFK SACSCDAEFYNCLK TYFNILQPQCFK
8 47	Uncharacterized protein	*Atta Cephalotes*	W4W624	45000 9.4	38000 5.2 19000 4.1	53* 55*	2% 2%	LFSLLLT LFSLLLT
8 39	Pancreatic lipase-related protein 2	*Acromyrmex echinatior*	EGI60946.1	39900 5.6	38000 5.2 37000 5.1	168#, 77#	4% 2%	LGHADFYPD KEYLTSVQL GHADFYPDGGT
27	PREDICTED: uncharacterized protein	*Apis mellifera*	XP_016770221	29500 9.8	25500 6.1	46*	4%	IKTLQDALKKAK

7

9	Uncharacterized protein	*Atta Cephalotes*	W4W555	53000 9.3	35000 5.2	50*	3%	SSSSSSSSSSSSSGQQR
12	PREDICTED: sperm flagellar protein 1-like	*Bombus impatiens*	XP_003486146.1	25700 8.9	19000 5.35	48*	3%	QISFSKPK
14	PREDICTED: apolipoprotein D-like	*Megachile rotundata*	XP_003699771.1	21000 6.3	19000 3.1	46*	5%	KPIGKSDDPK
15	Heat shock 70 kDa protein cognate 4	*Acromyrmex echinatior*	EGI62314.1	70300 5.4	11000 3.1	62*	7%	TTPSYVAFTDTER IINEPTAAAIAYGLDKK MVNHFVQEFK LSSNKRALR
15	ATP synthase subunit beta , mitochondrial	*Harpegnathos saltator*	EFN77326.1	55200 5.3	11000 3.1	60*	4%	TIAMDGTEGLVR IINVIGEPIDER
16	Uncharacterized protein GJ16099	*Drosophila virilis*	XP_002060462.1	8030 12.2	9000 4.6–5.0	170#	35%	VVTVLSPR VPGVSLPR VATVSNPR
17	PREDICTED: E3 ubiquitin-protein ligase TRAIP-like	*Bombus impatiens*	XP_003492946.1	48900 9.1	15200 5.1	64#	2%	VDEMLSR
18 19 20	26S proteasome non-ATPase regulatory subunit 1	*Camponotus floridanus*	EFN63483.1	11100 9.5	15200 5.9 15200 6.1 15200 6.9	58* 50* 57*	8% 8% 8%	VSDVNDDVR VSDVNDDVR VSDVNDDVR
25	PREDICTED: carotenoid isomerooxygenase isoform X1	*Megachile rotundata*	XP_003702796.1	72600 5.8	24000 6.0	52*	2%	NKAAQRIVVT LGLSVTPR
26	Gamma-interferon-inducible lysosomal thiol reductase	*Harpegnathos saltator*	EFN80944.1	24000 6.5	24000 6.1	218#	17%	RVDVYYQSLC CALHALK RFQCQHG LSLLPYGKA
26	Uncharacterized protein	*Atta cephalotes*	W4VZM2	56000 6.2	24000 6.1	59*	1%	FQALYGTK

8

26	DNA-directed RNA polymerases I and III subunit RPAC1	*Acromyrmex echinatior*	EGI70238.1	41800 7.9	24000 6.1	118#	5%	VYYQSLCPDSRT LLPYGKAK
33	Dehydrogenase/reductase SDR family member 11, partial	*Fopius arisanus*	JAG77010.1	29700 6.4	12500 8.0	49*	4%	KGLIVVGLAR
37	Allergen Sol i III	*Solenopsis invicta*	P35778.2	26400 8.1	25000 9.5	101#	7%	TDYCNVESCK
45	PREDICTED: ftsJ methyltransferase domain-containing protein 1-like	*Apis florea*	XP_003691009.1	78400 7.5	22000 4.5	21.8^	1%	KTKRLDNPQ
56	Uncharacterised protein	*Nasonia vitripennis*	XP_008213255.1	45800 5.4	10000 10.0	52*	2%	TLIEGEEDK

Symbols in the score column indicate which search engine the results were obtained from (*mascot, #MS blast, ^PEAKS)

9

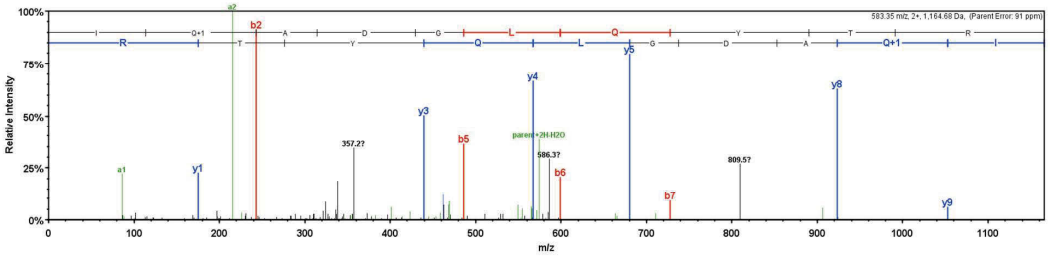
Table S3. *P. clavata* proteins identified from 2D-PAGE spots of electrically stimulated and manually dissected venom. Table shows proteins identified by each method as well as proteins unique to each collection method

Protein Name	Spot N°	Diss	ESV	Function
Glucosylceramidase	5	X	✓	Metabolism
Enolase	6	–	✓	Metabolism
Pancreatic lipase-related protein 2	8	–	✓	Metabolism
Pancreatic lipase-related protein 3	39	–	✓	Metabolism
ATP synthase subunit beta	15	–	✓	Metabolism
Myosin light chain alkali	17	X	✓	Metabolism
Myosin light chain alkali	79	✓	–	Metabolism
Dehydrogenase/reductase SDR family member 11, partial	33	–	✓	Metabolism
PREDICTED: superoxide dismutase [Cu-Zn]	66	✓	–	Metabolism
Phosphatidylinositol N-acetylglucosaminyltransferase subunit H, partial	67	✓	–	Metabolism
Agiotensin-converting enzyme	3	–	✓	Regulation
EAI_04268 SET and MYND domain-containing protein 4	5	X	✓	Regulation
Plasma glutamate carboxypeptidase	5	✓	✓	Regulation
Transmembrane protease, serine 9	7	–	✓	Regulation
TP53-regulating kinase	61	✓	X	Regulation
TP53-regulating kinase	40	✓	X	Regulation
TP53-regulating kinase	81	✓	–	Regulation
Heat shock 70 kDa protein cognate 4	15	–	✓	Regulation
PREDICTED: TRAF-interacting protein-like	17	✓	X	Regulation
26S proteasome non-ATPase regulatory subunit 1	18	X	✓	Regulation
26S proteasome non-ATPase regulatory subunit 2	19	✓	✓	Regulation
26S proteasome non-ATPase regulatory subunit 3	20	X	✓	Regulation
T-complex protein 11 domain-containing protein	22	✓	X	Regulation
PREDICTED: beta,beta-carotene 15,15'-monooxygenase-like	25	X	✓	Regulation
Gamma-interferon-inducible lysosomal thiol reductase	25	✓	X	Regulation
Gamma-interferon-inducible lysosomal thiol reductase	26	✓	✓	Regulation
DNA-directed RNA polymerases I and III subunit RPAC1	26	X	✓	Regulation
PREDICTED: DNA polymerase eta-like	26	✓	X	Regulation
PREDICTED: ftsJ methyltransferase domain-containing protein 1-like	45	X	✓	Regulation
NADPH:adrenodoxin oxidoreductase, mitochondrial	55	✓	X	Regulation
NADPH:adrenodoxin oxidoreductase, mitochondrial	65	✓	–	Regulation
PREDICTED: elongation factor Tu-like	74	✓	–	Regulation
Heat shock protein beta-1	75	✓	–	Regulation
Heat shock protein beta-2	76	✓	X	Regulation
Myosin regulatory light chain 2	75	✓	–	Regulation
Myosin regulatory light chain 3	77	✓	–	Regulation
Myosin regulatory light chain 4	78	✓	–	Regulation
Actin, muscle	5	✓	X	Structure
Actin, muscle	8	–	✓	Structure

Actin, muscle	63	✓	–	Structure
Actin, muscle	74	✓	–	Structure
Flexible cuticle protein 12	69	✓	–	Structure
PREDICTED: phospholipase A2-like	8	–	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	12	✓	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	14	✓	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	17	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	23	X	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	40	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	41	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	42	✓	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	43	✓	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	44	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	45	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	46	X	✓	Toxin/ Allergen
PREDICTED: phospholipase A2-like	77	✓	–	Toxin/ Allergen
PREDICTED: phospholipase A2-like	76	✓	X	Toxin/ Allergen
PREDICTED: phospholipase A2-like	78	✓	–	Toxin/ Allergen
Phospholipase A2, partial	9	–	✓	Toxin/ Allergen
Phospholipase A2, partial	17	X	✓	Toxin/ Allergen
Phospholipase A2, partial	18	X	✓	Toxin/ Allergen
Phospholipase A2, partial	19	X	✓	Toxin/ Allergen
Phospholipase A2, partial	44	✓	X	Toxin/ Allergen
Phospholipase A2, partial	62	✓	X	Toxin/ Allergen
Phospholipase A2, partial	72	✓	X	Toxin/ Allergen
Phospholipase A2, partial	75	✓	–	Toxin/ Allergen
Phospholipase A2, partial	77	✓	–	Toxin/ Allergen
Allergen Sol i III	37	–	✓	Toxin/ Allergen
EAG_03789 Transporter	4	–	✓	Transport
PREDICTED: sperm flagellar protein 1-like	12	X	✓	Transport
PREDICTED: sperm flagellar protein 1-like	62	✓	X	Transport
PREDICTED: apolipoprotein D-like	14	X	✓	Transport
PREDICTED: apolipoprotein D-like	40	✓	X	Transport
Sodium-independent sulfate anion transporter	58	✓	X	Transport
PREDICTED: calcium-binding mitochondrial carrier protein	78	✓	–	Transport
Uncharacterized protein	4	–	✓	Undetermined
hypothetical protein SINV_06838, partial	4	–	✓	Undetermined
Uncharacterized protein	8	–	✓	Undetermined
Uncharacterized protein	47	X	✓	Undetermined
Uncharacterized protein	19	✓	X	Undetermined
Uncharacterized protein	27	–	✓	Undetermined
Uncharacterized protein	40	✓	X	Undetermined
Uncharacterized protein	61	✓	X	Undetermined

Uncharacterized protein	72	✓	X	Undetermined
Uncharacterized protein	76	✓	X	Undetermined
Uncharacterized protein	81	✓	–	Undetermined
Uncharacterized protein	9	–	✓	Undetermined
Uncharacterized protein GJ16099	16	–	✓	Undetermined
Uncharacterized protein GJ16100	49	✓	X	Undetermined
hypothetical protein X777_04060	22	✓	X	Undetermined
PREDICTED: ES1 protein homolog, mitochondrial-like	22	✓	X	Undetermined
Uncharacterized protein	26	X	✓	Undetermined
Uncharacterized protein	56	✓	✓	Undetermined
Uncharacterized protein	70	✓	–	Undetermined

12



Sequence: (E)IqADGLQYTR(R)

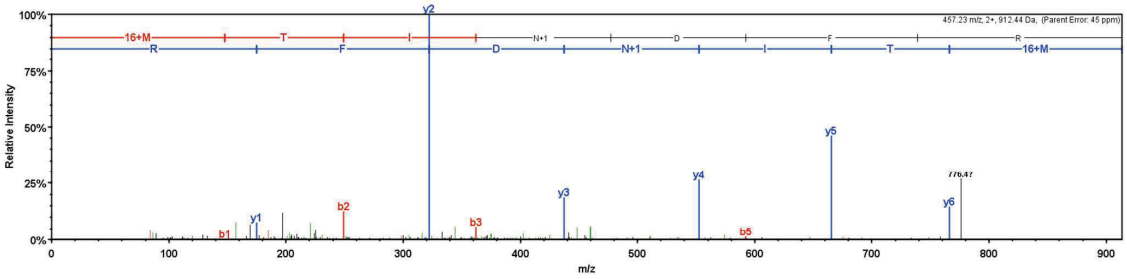
Modifications: Deamidated q (+1)

Probability: 97%

Precursor m/z: 1164.68

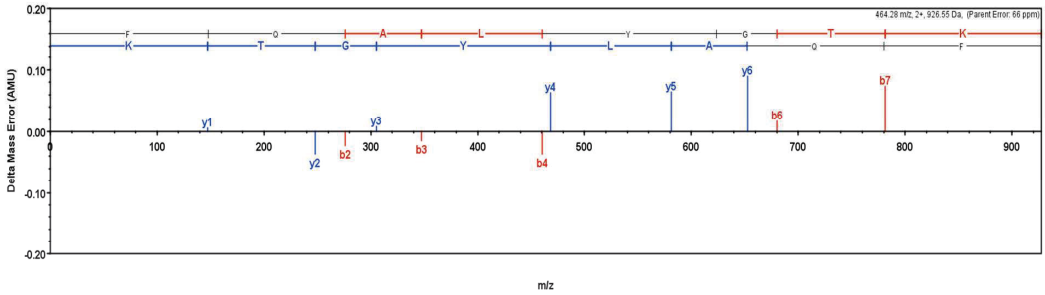
Charge state: 2

Supplementary Figure S1. Annotated MS spectra of peptides from manually dissected *P. clavata* 2D gel spot 22 used to match against hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide matched to T-complex protein 11 domain-containing protein, as well as the peptide's precursor *m*/*z*, charge state and modifications obtained from Scaffold (Proteome Software).



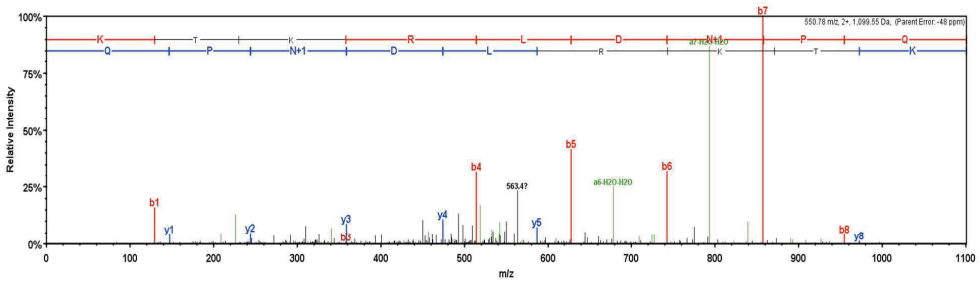
Sequence: (N)mTInDFR(K)
Modifications: Oxidated m (+16), Deamidated n (+1)
Probability: 90%
Precursor m/z: 912.4426
Charge state: 2

Supplementary Figure S2. Annotated MS spectra of peptides from manually dissected *P. clavata* 2D gel spot 58 used to match against hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide matched to sodium-independent sulfate anion transporter protein, as well as the peptide's precursor *m*/*z*, charge state and modifications obtained from Scaffold (Proteome Software).



Sequence: (S)FQALYGTK(T)
Modifications: None
Probability: 96%
Precursor m/z: 926.5476
Charge state: 2

Supplementary Figure S3. Annotated MS spectra of peptides from electrically stimulated *P. clavata* 2D gel spot 26 used to match against hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide matched to an uncharacterized protein, as well as the peptide's precursor *m*/*z*, charge state and modifications obtained from Scaffold (Proteome Software).



Sequence: (K)KTKRLDnPQ(N)
Modifications: Deamidated n (+1)
Probability: 100%
Precursor m/z: 1099.55
Charge state: 2

Supplementary Figure S4. Annotated MS spectra of peptides from electrically stimulated *P. clavata* 2D gel spot 45 used to match against hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide matched to PREDICTED: ftsJ methyltransferase domain-containing protein 1-like protein, as well as the peptide's precursor *m*/*z*, charge state and modifications obtained from Scaffold (Proteome Software).

Chapter Six

Comparisons of poneroid and formicoid ant venoms

Chapter Six: Comparisons of Protein and Peptide Complexity in Poneroid and Formicoid Ant Venoms

Compound Abstract

The previous chapter displayed the remarkable difference in the peptidomes of ant venom collected by electrical stimulation compared to ant venom collected by venom gland dissection in that the latter contained more, and a wider variety of peptides. Therefore, for the purposes of peptide drug discovery and understanding ant venom peptide diversity, venom gland dissection was deemed as the most viable collection method. This chapter, published in the *Journal of Proteome Research,* aimed to identify the true diversity of ant venom peptides in terms of peptide counts and mass using an offline RP-HPLC followed by MALDI-TOF-MS/MS approach that is known to reveal the true peptide complexity of venoms. Ants from the two major ant phylogenetic clades, poneroids and formicoids, had distinct peptide profiles in terms of peptide number and peptide mass distribution, with poneroids having higher peptide counts whereas formicoid venoms contained peptides with wider mass distributions. Closer investigation of the mass of the peptides from the same clades found that there were very few peptides in common.

Certificate of authorship and originality

This paper was published in the *Journal of Proteome Research*. The following is a copy reprinted with permission from ACS publications, Copyright © 2016 American Chemical Society. I certify that the work present in chapter six of this thesis has not previously been submitted as part of the requirements for a degree. I also certify that I carried out all the experimental work, analysis and interpretation of the data presented in this paper.

The authors listed on this manuscript contributed in the following way:

- Axel Touchard: Provided the venom
- Jennifer Koh: Ran the *E. tuberculatum* 2D-PAGE
- Alain Dejean. and Jerome Orivel: Assisted in venom collection
- Matthew Padula: Mass spectrometry support
- Pierre Escoubas: Proof-read manuscript
- Graham Nicholson: Assisted in experimental design and manuscript direction
- All authors proof-read manuscript

Principal supervisor

Professor Graham Nicholson

Samira Ryma Aili

_____

Signature

_____

Signature

_____

Date

_____

Date

Journal of proteome research

Article

pubs.acs.org/jpr

Comparisons of Protein and Peptide Complexity in Poneroid and Formicoid Ant Venoms

Samira R. Aili,[†] Axel Touchard,[‡] Jennifer M. S. Koh,[†] Alain Dejean,[‡,§] Jérôme Orivel,[‡] Matthew P. Padula,[∥] Pierre Escoubas,[⊥] and Graham M. Nicholson*[,†]

[†]Neurotoxin Research Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, Sydney, New South Wales 2007, Australia

[‡]CNRS, UMR Ecologie des Forêts de Guyane (EcoFoG), Campus Agronomique, BP 316, 97379 Kourou Cedex, France

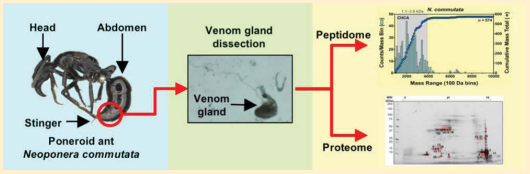
[§]Laboratoire Écologie Fonctionnelle et Environnement, CNRS, UMR 5245, Ecolab, 118 route de Narbonne, 31062 Toulouse, France

[∥]Proteomics Core Facility, Faculty of Science, University of Technology Sydney, Sydney, New South Wales 2007, Australia

[⊥]VenomeTech, 473 Route des Dolines - Villa 3, Valbonne 06560, France

Supporting Information

ABSTRACT: Animal venom peptides are currently being developed as novel drugs and bioinsecticides. Because ants use venoms for defense and predation, venomous ants represent an untapped source of potential bioactive toxins. This study compared the protein and peptide components of the poneroid ants *Neoponera commutata*, *Neoponera apicalis*, and *Odontomachus hastatus* and the formicoid ants *Ectatomma tuberculatum*, *Ectatomma brunneum*, and *Myrmecia gulosa*. 1D and 2D PAGE revealed venom proteins in the mass range <10 to >250 kDa. NanoLC-ESI-QTOF MS/MS analysis of tryptic peptides revealed the presence of common venom proteins and also many undescribed proteins. RP-HPLC separation followed by MALDI-TOF MS of the venom peptides also revealed considerable heterogeneity. It was found that the venoms contained between 144 and 1032 peptides with 5−95% of peptides in the ranges 1−4 and 1−8 kDa for poneroid and formicoid ants, respectively. By employing the reducing MALDI matrix 1,5-diaminonapthalene, up to 28 disulfide-bonded peptides were also identified in each of the venoms. In particular, the mass range of peptides from poneroid ants is lower than peptides from other venoms, indicating possible novel structures and pharmacologies. These results indicate that ant venoms represent an enormous, untapped source of novel therapeutic and bioinsecticide leads.



KEYWORDS: *Hymenoptera, toxin, proteomic analysis, mass spectrometry, peptidome, ant venom, LC-MALDI-TOF MS, nanoLC-ESI-QTOF MS/MS*

INTRODUCTION

Arthropod venoms contain a vast number of insecticidal peptide toxins that are yet to be characterized. The current estimate for the numbers of uncharacterized bioactive peptides within spider venoms alone is in the millions,[1] from a predicted total of 170 000 species.[2] Ant venoms are even less characterized; out of a total of 9000 stinging species, only 11 species have been investigated.[3,4]

Ants and other stinging insects, such as Apoidea (bees) and Vespoidea (wasps), belong to the family Formicidae of the order Hymenoptera.[5] Ants represent a diverse group of arthropods,[6] with approximately 13 000 extant species[7] and an estimated total of 25 000 described and undescribed species.[8,9] Ants are considered leading invertebrate predators, with some adapted to prey upon certain groups of insects using their venom.[10] However, their venom is also used for other purposes including defense against predators and social communication.[11] The success of ants is reflected by their representation of 15−20% of the rainforest animal biomass.[12]

The limited work done thus far reveals that ant venoms show enormous chemical diversity comprising alkaloids, hydrocarbons, proteins, and peptides.[3,4] Most studies on ant venoms have focused on the alkaloidal components, as these are a major cause of allergic reactions due to their histamine-releasing properties.[13,14] However, more recent studies have shown the involvement of the proteome in eliciting these allergic[15,16] and anti-inflammatory[17,18] properties. Additional studies have also revealed that the peptidome is a source of antimicrobial activity against both Gram-positive and Gram-negative bacteria.[6,11] Neurotoxic ant venom peptides have also been identified and appear to be important for defense and prey capture.[19] This is an extremely understudied area of research, as only two neurotoxic ant peptides have been extensively characterized: poneratoxin[20,21] and ectatomin.[22]

The limited number of studies on ant peptides reveals that they are either linear (e.g., pilosulin), homo-, or heterodimeric

Received: February 28, 2016

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

(e.g., ectatomins) or contain an inhibitor cystine knot structural motif.[3,23] The number studies on ant venoms has been restricted by the small amount of venom secreted by ants, which makes it difficult to extract sufficient material for proper purification and characterization;[24] however, the previous lack of high sensitivity analytical tools[6] has recently been overcome by improvements in mass spectrometric instrumentation.[25,26]

While the number of biologically active peptides from all described venomous animal species is estimated to approach 20 million, to date, only 75 peptides are described from ant venoms.[3,4,25] Consequently, ants are an understudied source of neurotoxic peptides that could be used as leads for novel bioinsecticidal compounds. Strong indicators include the use of their venom for predation and defense against other insects[27] and the presence of disulfide-rich, low-molecular-weight peptides.[3,4] These types of peptides are readily absorbed, highly stable, and resistant to degradative environmental factors.[28,29] Given the limited investigation of ant venoms, this study aimed to investigate the complexity of the protein and peptide components of six different ant venoms from the two major ant clades: formicoids and poneroids. The present study found that ant venoms are, indeed, rich in both peptides and proteins that could serve as potential sources of structurally and pharmacologically novel neurotoxins. These could be mined for potential biopesticides to control agricultural pests and insect vectors that transmit diseases.

EXPERIMENTAL SECTION

Ant Collection and Taxonomy

Live specimens of worker ants from the poneroid species *Neoponera commutata*, *Neoponera apicalis*, and *Odontomachus hastatus* (Ponerinae) and the formicoid species *Ectatomma tuberculatum* and *Ectatomma brunneum* (Ectatomminae) were collected from several regions of French Guiana. Collected ants were stored at −20 °C prior to dissection of the venom glands. Dissected glands from an individual species were pooled in 10% acetonitrile (ACN) in water (v/v). Samples were centrifuged for 5 min at 14 400 rpm (12 000g_{av}). The supernatant was then collected and lyophilized prior to storage at −20 °C. Lyophilized *Myrmecia gulosa* (formicoid clade; Myrmeciinae) venom was purchased from Southwestern Biological Institute (Flagstaff, AZ). All chemicals used were of analytical grade and sourced from Sigma-Aldrich (Castle Hill, NSW, Australia). All buffers were prepared using milli-Q water.

Protein Assay

Protein quantification was performed using a QuantiPro BCA Assay Kit (Sigma-Aldrich) or Qubit Protein Assay Kit (Invitrogen, Mulgrave, Victoria, Australia) with the Qubit 2.0 fluorometer as per the manufacturer's instructions using bovine serum albumin (BSA) as the standard.

One-Dimensional Gel Electrophoresis

100 μg of crude venom samples was mixed with 12 μL of SDS sample buffer (15 mM Tris, pH 6.8; 0.25% (w/v) SDS; 0.25% (v/v) β-mercaptoethanol; 2.5% (v/v) glycerol, and 0.0025% (w/v) bromophenol blue) and boiled as previously described.[30] Samples were then centrifuged at 16 873g_{av} for 5 min before being loaded into a 4−20% Criterion TGX gel (Bio-Rad, Gladesville, NSW, Australia) and electrophoresed in Tris-glycine-SDS running buffer (Bio-Rad) alongside Precision Plus Protein standards (Bio-Rad). The gel was run for ∼30 min at 300 V until the dye front reached the bottom of the gel. The gel was then placed in a fixing solution containing 40% (v/v) methanol and 10% (v/v) acetic acid for 30 min at room temperature before staining with Flamingo fluorescent stain (Bio-Rad) for 1 h. After staining, the gel was scanned using a Molecular Imager PharosFX Plus system (Bio-Rad) with QuantityOne software. To perform nanoLC-ESI-QTOF MS/MS, we additionally stained gels with Coomassie Blue G250 overnight then destained them with 1% (v/v) acetic acid for 2 h.

Two-Dimensional Gel Electrophoresis

300 μg samples of *N. commutata* and *E. tuberculatum* venom were resuspended in 100 μL of 7 M urea, 2 M thiourea, 1% (v/v) C7BzO, 50 mM Tris HCl pH 8.8 before reduction and alkylation of disulfide bonds with 5 mM tributylphosphine and 20 mM acrylamide monomers for 90 min. The venom was then desalted using a BioSpin column (Bio-Rad) equilibrated in 7 M urea, 2 M thiourea, 1% (v/v) C7BzO. The sample was subjected to 2D-PAGE as previously described.[31] In brief, sample was fractionated by isoelectric focusing on an 11 cm pH 3−10 IPG strip (Bio-Rad) for 100 kVh and then separated on a second dimension on a 4−20% gel. The gels were subsequently visualized by staining with Flamingo Gel Stain or Coomassie Blue G250 as per the previous section.

Protein Identification

Gel spots or bands were excised, digested as previously described,[32] and subjected to nanoLC-ESI-QTOF MS/MS. In brief, tryptic digests of proteins in gel spots or bands were placed on a Tempo nanoLC system (Eksigent, Waverley, Victoria, Australia) and loaded onto a Michrom reversed-phase trapping cartridge at a rate of 20 μL/min with MS loading solvent, consisting of 2% ACN and 0.2% trifluoroacetic acid (TFA), onto a C8 trap column (Michrom Biosciences, USA). The samples were eluted at 300 nL/min onto a 75 μm × 150 mm PicoFrit column (New Objective, USA) packed with Magic C18AQ chromatography resin (Bruker-Michrom Biosciences, Auburn, CA). Peptides were eluted from the column and into the source of a QSTAR Elite hybrid quadrupole-time-of-flight mass spectrometer (AB SCIEX, Framingham, MA) using the following program: 5−50% MS solvent B (98% ACN + 0.2% formic acid) over 8 min, 50−80% MS buffer B over 5 min, 80% MS buffer B for 2 min, and 80−5% for 3 min. MS solvent A was 2% ACN + 0.2% formic acid. The eluting peptides were ionized with a 75 μm ID emitter tip that tapered to 15 μm (New Objective) at 2300 V. An Intelligent Data Acquisition (IDA) experiment was performed, with a mass range of 375−1500 Da continuously scanned for peptides of charge state 2^+−5^+ with an intensity of more than 30 counts/s. Selected peptides were fragmented and the product ion fragment masses were measured over a mass range of 100−1500 Da. The mass of the precursor peptide was then excluded for 15 s.

The MS/MS data files produced by the QSTAR were searched using Mascot Daemon (version 2.4, provided by the Walter and Elisa Hall Institute[33]) and searched against the LudwigNR database (composed of the UniProtKB, plasmoDB, and Ensembl databases vQ215; 19 375 804 sequences with 6 797 271 065 residues) with the following parameter settings. Fixed modifications: none. Variable modifications: propionamide C, oxidized M, and deamidation NQ. Enzyme: semi-trypsin. Number of allowed missed cleavages: 3. Peptide mass tolerance: 100 ppm. MS/MS mass tolerance: 0.2 Da. Charge state: 2^+, 3^+, and 4^+. MS/MS data files were also searched against Hymenopteran proteins in the NCBInr protein database using PEAKS 6 software (Bioinformatics Solutions,

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

Waterloo, ON, Canada). The results from both searches were then filtered by including only protein hits with at least one unique peptide. Peptides identified by PEAKS 6 were further validated by manual inspection of the MS/MS spectra for the peptide to ensure the b- and y-ion series were sufficiently extensive for an accurate identification.

C18 Crude Venom RP-HPLC Fractionation

Lyophilized crude venom was separated using a Vydac analytical C18 RP-HPLC column (218TP54, 25 cm × 4.6 mm, 5 μm pore size) on a Shimadzu HPLC system. Fractionation of venom peptide components was achieved using a linear gradient of two mobile phases: 0.1% (v/v) trifluoroacetic acid (TFA) in water (solvent A) and 0.085% (v/v) TFA in ACN (solvent B). For all venoms except *E. brunneum*, separation was achieved using a flow rate of 1 mL/min with the following gradient of: 0−5 min, 0% solvent B; 5−65 min, 0−60% B; 65−75 min, 60−90% B; 75−80 min, 90% B; 80−85 min, 90−0% B; and 85−90 min, 0% B. For *E. brunneum*, venom separation used a slightly longer gradient to facilitate the isolation of the more hydrophobic peptides comprising: 0−10 min, 5% solvent B; 10−80 min, 5−70% B; 80−90 min, 70−95% B; 90−95 min, 95% B; 95−100 min, 95−5% B; and 100−105 min, 5% B. Peaks were monitored at absorbances of 280 and 215 nm and manually collected, then lyophilized and stored at −20 °C for further use.

MALDI-TOF MS Analysis

MS analysis of the HPLC fractions was performed on an AB SCIEX 5800 MALDI-TOF/TOF mass spectrometer. Initial experiments were performed using several MALDI matrices on *E. tuberculatum* HPLC fractions to test for optimum peptide ionization and crystallization. The four matrices tested were: 5 mg/mL of α-cyano-4-hydroxycinnamic acid (CHCA) dissolved in 1% (v/v) TFA/100% (v/v) ACN/0.1 M $NH_4H_2PO_4$ (45/45/10); 10 mg/mL ferulic acid (FA) in 40% (v/v) ACN/10 mM serine; 10 mg/mL *N*-*tert*-butyl-*N*-isopropyl-*N*-methylammonium α-cyano-4-hydroxycinnamate (ionic CHCA) in 0.1% (v/v) formic acid/50% (v/v) ACN; and a 1:1 mixture of CHCA and FA. All venom fractions were reconstituted with 10 μL of water. Then, 1 μL of each fraction was overlaid with 0.5 to 0.75 μL of the matrix using the dried droplet method.[34] Each spectrum was calibrated externally using a mixture of peptides with known molecular masses in the same m/z range (AB SCIEX). Mass spectra were acquired in linear mode over the range 1000−10 000 m/z with 400 laser shots accumulated for each sample, based on the acceptance parameters and adequate signal intensity. Mass spectra were collected in positive ion mode, and signals below 1000 m/z were not recorded as they were composed of mostly matrix-related ion clusters.[35]

Identification of Disulfide-Bonded Peptides

Disulfide-bonded peptides were identified by chemical reduction of HPLC fractions. RP-HPLC fractions were spotted onto a MALDI plate, as detailed in the previous section and overlaid with the reducing matrix 1,5-diaminonapthalene (1,5-DAN; 10 mg/mL) in 0.05% (v/v) formic acid/50% (v/v) ACN.[36] The matrix was always prepared immediately before use due to the instability of 1,5-DAN in ACN. All MALDI-TOF data collected using 1,5-DAN or CHCA matrix were acquired in reflector mode to isotopically resolve ions for comparison purposes. Chemical reduction results in a mass increase of 2 Da for each disulfide bond. Thus, by comparing the mass spectra of native and reduced samples the number of disulfide bonds in ant venom peptides could be determined.

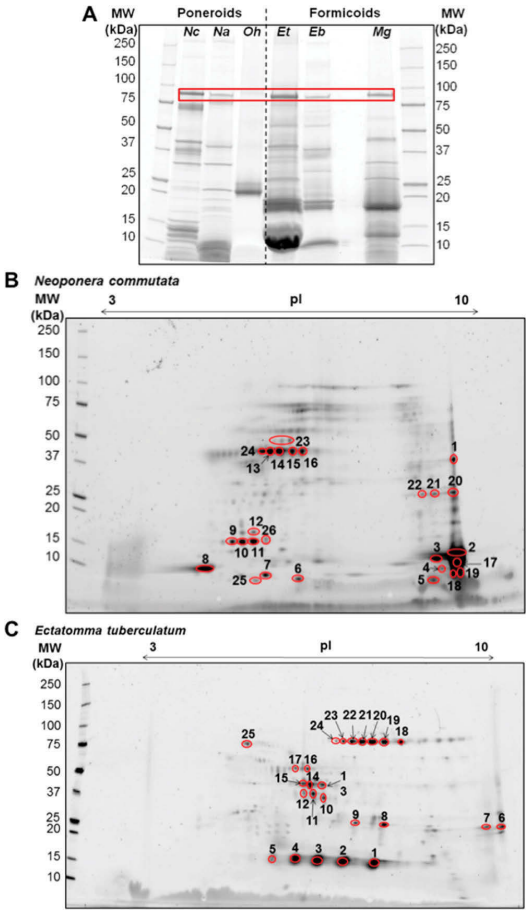


Figure 1. 1D and 2D SDS-PAGE gels of representative ant venoms. (A) 1D gel showing molecular weight markers (MW) and reduced and alkylated 100 μg samples of *N. commutata* (*Nc*), *N. apicalis* (*Na*), *O. hastatus* (*Oh*), *E. tuberculatum* (*Et*), *E. brunneum* (*Eb*), and *M. gulosa* (*Mg*). The red box highlights a conserved protein band across the six species. (B−C) 2D SDS-PAGE of reduced and alkylated *N. commutata* (B) and *E. tuberculatum* (C) whole venom. A 300 μg sample of crude venom was subjected to isoelectric focusing on an 11 cm pH 3−10 IEF strip. All gels were separated by a Tris-glycine (4−20%) SDS-PAGE and stained with Flamingo. Protein spots (circled) were excised for identification by nanoLC-ESI-QTOF MS/MS and de novo peptide sequencing.

Data Analysis

Data Explorer v4.11 software (AB SCIEX) was used to analyze spectra from individual RP-HPLC fractions to characterize the number and masses of peptides per venom, as previously reported.[37] In brief, spectra were subjected to baseline correction with a correlation factor of 0.7 and Gaussian smoothing to reduce noise with a five-point filter width. All mass attributions were verified manually, and a mass list was created for each LC-MALDI-TOF MS run. Potential adducts from oxidation, hydration, sodium, and deamination were manually

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 1. Protein Assignment of *N. commutata* Venom 2D-PAGE Spots Identified by nanoLC-ESI-QTOF MS/MS

2D gel spot	protein ID[a]	organism	UniProt accession number	theoretical MW (Da) and pI	observed MW (Da) and pI	coverage and score (−10lgP)	peptide (*m/z*)	peptide sequence
1	actin-5C	*Acromyrmex echinatior*	EGI67274.1	41822	38000	12%	1108.04	DLYANTVLSGGTTMYPGIADR
				5.33	9.8	90.4	855.43	LC(+71.04)YVALDFEQEMA-TAASSSSLEK
2	venom acid phosphatase	*Camponotus flor-idanus*	EFN60591	41821.8		4%	911.52	LYLYSGHETNLASLLK
				5.30	11500	23.6	443.80	LFAGPLIR
	ponericin-G3	*Neoponera goeldii*	P82416.1	3383.0	9.5	27%	523.80	GWKDWLNK[b]
				10.22		26.66		
5	ponericin-G3	*Neoponera goeldii*	P82416	3383.0	<5000	27%	523.78	GWKDWLNK[b]
				10.22	9.2	25.8		
8	TP53-regulating kinase	*Acromyrmex echinatior*	EGI66742.1	27120.4	10000	4%	561.32	KQ(+.98)YEQILAK[b]
				9.36	5.0	27.5		
10	venom acid phosphatase	*Camponotus flor-idanus*	EFN60591	41821.8	15000	4%	911.52	LYLYSGHETNLASLLK
				5.30		23.6%	443.79	LFAGPLIR
	putative uncharacterized pro-tein	*Acromyrmex echinatior*	F4WHI8	18400.6	5.5	4%	421.76	TLENLPR[b]
				6.06		23.2%		
13	actin-5, muscle-specific	*Harpegnathos saltator*	EFN78406	41782.7	39000	26%	977.57	VAPEEHPVLLTEAPLNPK
							895.96	SYELPDGQVITIGNER
							1114.54	DLYANNVLSGGTTMYPGIADR
				5.30	6.1	138.9	566.77	GYSFTTTAER
							758.41	IWHHTFYNELR
							599.87	AVFPSIVGRPR
							572.35	EITALAPSTIK
14	testicular acid phosphatase-like protein	*Harpegnathos saltator*	EFN78070	46695.4	39000	7%	430.79	LINVVFR
				5.75	6.3	38.3	1258.11	VDDKLLLSDEC(+71.04) PEYLN(+0.98)EYER
15	actin	*Acromyrmex echinatior*	EGI64684.1	41773.9	39000	3%	599.89	AVFPSIVGRPR[b]
				5.29	6.5	43.9		
17	ponericin	*Neoponera goeldii*	P82416	3383.0	10000	27%	523.80	GWKDWLNK[b]
				10.22	9.7	26.7%		
18	ponericin	*Neoponera goeldii*	P82416	3383.0	8000	27%	523.79	GWKDWLNK[b]
				10.22	9.4	29.3		
	uracil phosphoribosyltrans-ferase	*Nasonia vitripen-nis*	K7IRF5	29771.4		4%	490.26	MGSADAATKK[b]
				7.53		21.3		
19	ponericin	*Neoponera goeldii*	P82416	3383.0	9000	27%	523.78	GWKDWLNK[b]
				10.22	9.6	33%		
20	PLA_2 2.2	*Apis mellifera*	H9K6W6	23582.7	26000	11%	877.94	QMSSNLVADLEETCK
				6.96	9.6	41.4	590.34	SVLVADTTMSR
23	trehalase	*Microplitis demo-litor*	W4VV04	67322.0	49000	2%	676.41	SQPPLLIPMVEK[b]
				5.50	6.3	41.8		
24	arylsulfatase B	*Harpegnathos saltator*	EFN75189	65573.1	39000	2%	577.87	GVAAIWSPLIK[b]
				5.86		48.6	977.56	VAPEEHPVLLTEAPLNPK
	actin-5, muscle-specific	*Harpegnathos saltator*	EFN78406	41782.7		12%	895.97	SYELPDGQVITIGNER
				5.30	5.9	121.2	599.87	AVFPSIVGRPR
	testicular acid phosphatase-like protein	*Camponotus flor-idanus*	EFN63831	148847.4		5%	1258.08	VDDNLILTDEC(+71.04)PQ(+0.98) YLDEYER
				6.61		41.4	443.78	LFAGPLIR

[a]MS/MS peptide sequences were searched using PEAKS software against the NCBInr database restricted to Hymenopterans. Protein matches from the NCBInr database with the highest coverages are shown. Theoretical mass and pI were determined using Protparam in ExPASy.[76] [b]Spectra for these sequences have been included in the Supporting Information as Supplementary Figures S2−S11. Spots with no matches: 3, 4, 6, 7, 9, 11, 12, 16, 21, 22, 25, and 26.

removed from all mass lists as well as any potential dimers or doubly charged species. Peptides with mass matches within ±1.0 Da in adjoining HPLC fractions were considered identical and were removed from the data set because they likely reflect incomplete separation. For each venom, the peptides from each RP-HPLC fraction were consolidated into one mass list. Mass matching across data sets was performed in Microsoft Excel (Redmond, WA) to identify similar peptides across the different venoms or between different data sets of the same venom when comparing matrices. Two-dimensional scatter plots, termed "2D venom landscapes", were constructed using Prism v6 software (GraphPad, La Jolla, CA). All peptide masses detected in each HPLC fraction were plotted as a function of their *m/z* values (*x* axis) and their HPLC retention time reflecting their hydrophobicity (*y* axis).

RESULTS

Venom Protein Characterization

Initially, a 1D-PAGE of all venoms was performed to determine overall venom protein profiles. It was found that the six ant

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 2. Protein Assignment of *E. tuberculatum* Venom 2D-PAGE Spots Identified by nanoLC-ESI-QTOF MS/MS

2D gel spot	protein ID[a]	organism	UniProt accession number	theoretical MW (Da) and pI	observed MW (Da) and pI	coverage and score (−10lgP)	peptide (*m*/*z*)	peptide sequence
1	uncharacterized protein	*Apis mellifera*	A0A088A5D5	90811.8	14500	1%	578.34	IVNKVEVINK[b]
				8.91	7.5	24.4		
2	phospholipase A2 (fragment)	*Harpegnathos saltator*	E2BTD5	12034.7	14800	13%	883.9	SACSCDAEFYNCLK[b]
				8.01	6.9	20.3		
7	hypothetical protein EAI_10007	*Harpegnathos saltator*	E2BJ69	26958.7	22000	10%	640.83	YPLPLADGSGYK
				6.7	9.9	56.3	670.8	GAMGCGPQETFR
13	glutathione S-transferase	*Acromyrmex echinatior*	F4W5B2	25278.2	38000	6%	772.43	IRALPEIKAYIEK[b]
				6.85		22.1		
	venom acid phosphatase	*Cerapachys biroi*	EZA53020.1	47041.5	6.5	2%	423.77	LVNVVFR[b]
				7.60		30.4		
14	venom acid phosphatase	*Cerapachys biroi*	EZA53020.1	47041.5	38000	3%	423.77	LVNVVFR
				7.60	6.1	38.9	364.21	LLYYR
15	venom acid phosphatase	*Cerapachys biroi*	EZA53020.1	47041.5	39000	2%	423.77	LVNVVFR[b]
				7.60	6.0	33.2		
17	endochitinase	*Harpegnathos saltator*	E2BB48	52819.7	50000	6%	739.85	NNFFYFVEELR
				6.08	5.9	75.3	543.84	LVVGIPLYGR
18	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	5%	783.42	VYYLATAPGEPTQR
				6.24	8.0	71.9	832.08	C(+71.04) GISVAPVTSWIYYDSIYTER
19	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	2%	783.41	VYYLATAPGEPTQR[b]
				6.24	7.8	56.6		
20	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	5%	783.41	VYYLATAPGEPTQR
				6.24	7.5	51.6	832.08	C(+71.04) GISVAPVTSWIYYDSIYTER
21	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	5%	783.41	VYYLATAPGEPTQR
				6.24	7.2	72.5	1247.61	CGISVAPVTSWIYYDSIYTER
22	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	5%	832.01	C(+71.04) GISVAPVTSWIYYDSIYTER
				6.24	7..0	80.1	783.41	VYYLATAPGEPTQR
23	dipeptidyl peptidase 4	*Camponotus floridanus*	E2AF09	86583.9	75000	5%	783.41	VYYLATAPGEPTQR
				6.24	6.9	90.6	661.82	IYYDSIYTER
							832.08	C(+71.04) GISVAPVTSWIYYDSIYTER
25	angiotensin-converting enzyme	*Camponotus floridanus*	E2B155	75684.5	75000	6%	696.39	YHIIADVEYIR
				5.30	5.0	59.2	554.29	YVELANTAAR
							646.65	SM(+15.99) LELGSSKPWPDAM(+15.99)EK

[a]MS/MS peptide sequences were searched using PEAKS software against the NCBInr database restricted to Hymenopterans. Protein matches from the NCBInr database with the highest coverages are shown. Theoretical mass and pI were determined using Protparam in ExPASy.[75] [b]Spectra for these sequences have been included in the Supporting Information as Supplementary Figure S12−S16. Spots with no matches: 3, 4, 5, 6, 8−12, 16, and 24.

venoms all had unique protein compositions (Figure 1A). *N. commutata* venom contained the highest number of proteins in the range 10−250 kDa, while *E. tuberculatum* venom also had a high number of proteins mainly in the range 10−80 kDa. The remaining venoms all had diverse and often complex protein profiles apart from *O. hastatus*, with only three distinct protein bands. Despite this clear heterogeneity in protein composition, the 1D-PAGE electrophoretic protein profiles revealed similarities among all six venoms (red box in Figure 1A). For example, using nanoLC-ESI-QTOF MS/MS the ca. 90 kDa band seen in all venoms was identified as dipeptidyl peptidase IV in all cases except *O. hastatus*. Another common feature of most venoms was the presence of bands below ca. 13 kDa.

To identify these proteins, we performed 2D-PAGE on the venoms of *N. commutata* and *E. tuberculatum*. The remaining venoms were not investigated due to the limited amount of material. The 2D-PAGE spot patterns obtained after flamingo staining are depicted in Figure 1B−C. The 2D gel of *N. commutata* venom had a large cluster of low molecular mass proteins (spots 2−5, 17−19; corresponding to ca. 10−15 kDa) present in the lower right corner of the 2D gel at ca. pH 10. This was not seen with *E. tuberculatum* venom that had a greater number of higher molecular weight proteins, particularly ca. 90 kDa (Figure 1C). Nevertheless, there were several similarities between the two gels, such as very few acidic proteins. Both gels also had several spots with isoforms, such as those at 15 kDa in both venoms (e.g., *N. commutata* spots 9−11, *E. tuberculatum* spots 1−5), ca. 40 kDa for *N. commutata* (spots 13−16, 24) and ca. 80 kDa for *E. tuberculatum* at neutral pH.

Because of the limited number of sequenced ant genomes, the generated nanoLC-ESI-QTOF MS/MS data were initially searched with Mascot; however, no significant hits were seen. Therefore, all data, including peptides that were identified by de novo sequencing in PEAKS 6, were searched against the NCBInr database using PEAKS 6. The results of this search are summarized in Tables 1 and 2. From all of the spots subjected

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

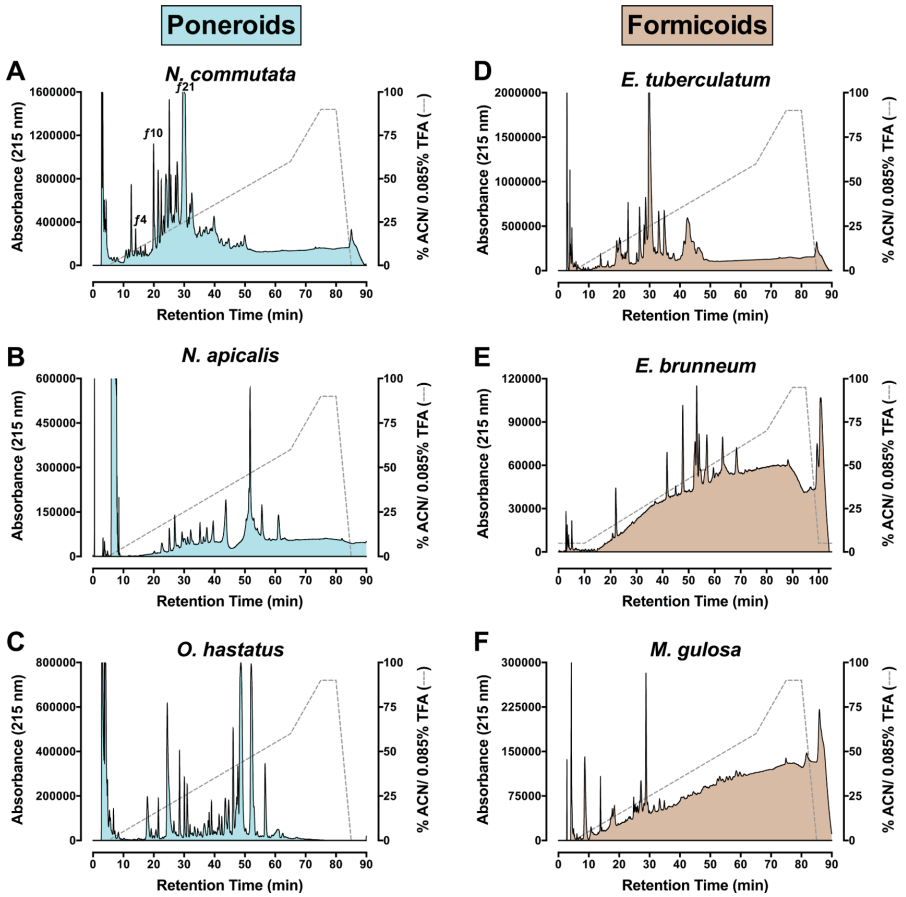


Figure 2. RP-HPLC chromatograms of poneroid and formicoid ant venoms. Venoms were separated by analytical C18 RP-HPLC using an ACN/0.085% TFA gradient of 1%/min (gray dotted line) at 1 mL/min. Panels show the chromatographic elution profile of venoms from the poneroid ants (A) *N. commutata*, (B) *N. apicalis*, and (C) *O. hastatus* and the formicoid ants (D) *E. tuberculatum*, (E) *E. brunneum*, and (F) *M. gulosa*. Note the different gradient in panel E for *E. brunneum* venom. Chromatographic peaks were monitored at 215 and 280 nm (only absorbance at 215 nm is shown for clarity). MALDI spectra of the numbered HPLC fractions (*f*) from *N. commutata* (panel A) are shown in Figure 4.

to MS/MS analysis, 14 of the 26 *N. commutata* and 14 of the 25 *E. tuberculatum* 2D gel spots analyzed matched to proteins. None of these were proteins from the ants investigated due to no entries being present for those species in databases. The remaining spots could not be matched to any known ant or other hymenopteran proteins despite the numerous high-quality MS/MS spectra and de novo peptide sequences generated, including 136 spectra from spot 8 of the *N. commutata* gel. Several of the spots were matched to proteins using a single peptide; the MS/MS spectra for these spots have been included in the Supporting Information as Supplementary Figures S2−S16.

Of all of the matches, three proteins were common to both gels. First, phospholipase A_2 (PLA_2) proteins were identified, although in completely different positions in the gels from the two different species. The *E. tuberculatum* PLA_2 was identified at ca. pI 7 and mass of ca. 15 kDa (Figure 1C; spot 2). However, the *N. commutata* PLA_2 was identified at a pI of 10 and mass of ca. 26 kDa (Figure 1B; spot 20), potentially as a result of dimerization. The second protein common to both species was venom acid phosphatase; in *N. commutata* it was found in both spots 2 and 10 and in *E. tuberculatum* it was found in spots 13−15. The third protein common to both venoms was identified as dipeptidyl peptidase. Although this protein was not identified in the 2D gel of *N. commutata*, it was detected from the 1D gel slice (Figure 1A). The mass of this protein was ca. 90 kDa in the 1D gels from both *E. tuberculatum* and *N. commutata* venom (Figure 1A) as well as in the 2D gel of *E. tuberculatum* venom (Spot 18−23, Figure 1C).

In both venoms, the mass of some spots on the gel did not correspond to the exact theoretical mass of their assigned protein. For example, *N. commutata* spot 10 is at ca. 15 kDa and pI of 6, which is ca. 3 kDa lighter than the predicted protein, although a similar pI. One spot that seems to be far from the predicted position of the protein is *E. tuberculatum* spot 1, as seen in Figure 1C. It is at ∼15 kDa on the gel and at pI ∼8. This is significantly different to the assigned protein's calculated mass of 91 kDa and pI of 9, as indicated in Table 2. This is most likely the result of post-translational modification, where

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

the produced protein has undergone proteolysis, for activation of either function or degradation.

Ant Venom Peptide Characterization

Whole Venom C18 RP-HPLC Separation. The venom of each ant was separated using RP-HPLC to reveal the complexity of the venoms while avoiding possible ion suppression effects known to occur in electrospray when using complex mixtures of peptides such as venoms. This initial separation step revealed complex chromatographic profiles (Figure 2) with each venom yielding a number of venom fractions. All venoms, except that of *E. brunneum*, appeared to contain quite hydrophilic peptides, as most of the peptides eluted early in the chromatogram at ACN concentrations <30%.

MALDI-TOF MS Matrix Optimization. Our recent work has shown that DHB, sinapinic acid, caffeic acid, and 2,4,6-trihydroxyacetophenone (THAP) are poor MALDI matrices for the automated detection of ant venom peptides and that FA and CHCA allow for the detection of the greatest number of peptides with the highest intensities.[38] Accordingly, initial experiments tested four different combinations of these two matrices to determine the best one to use for the remainder of the study: CHCA alone, ionic CHCA, FA alone, and FA in combination with CHCA. Of these matrices, the matrix that covered the largest mass range was FA (Figure 3A). However, the matrix yielding the highest peptide count was CHCA (Figure 3B); therefore, all remaining venoms investigated were analyzed with both CHCA and FA.

Number of Ant Venom Peptides and Mass Distribution. Each HPLC fraction from the six venoms was spotted onto a MALDI target with either CHCA or FA matrix to determine the number of peptides per fraction and their mass distribution. Although the ant venom RP-HPLC fractions were seemingly homogeneous, with sharp peaks, several peptides were often observed in the same fraction when further resolved by MALDI-TOF MS. There were clear differences in the abundance of masses from each chromatographic peak as well as differences in the mass range of peptides detected and intensities between the two matrices. A single RP-HPLC fraction mostly contained multiple signal peaks that corresponded to different masses. For example, 37 masses were observed in *f*10 from *N. commutata* venom using FA and 26 masses using CHCA (Figure 4A). Furthermore, many low UV absorbance peaks from the RP-HPLC displayed significant numbers of peptides (e.g., 35 peptides in *N. commutata f*4; Figure 4Bb). Conversely, fractions with high UV absorbances sometimes only had low numbers of peptides in their MALDI spectrum using either CHCA or FA. For example, *N. commutata f*21 contained only a single high intensity mass at 3364.9 *m*/*z* (Figure 4C). Overall, FA appeared to be better at detecting peptides in the poneroid venoms and was more efficient at detecting lower molecular weight peptides below 1800 Da in the formicoid venoms. However, CHCA was better at detecting peptides in formicoid venoms above 1800 Da (Figure 3C and shaded area of Figure 8Ab).

Frequency histograms in 100 Da bins were constructed to show the full peptide mass distribution of all masses detected. The data showed that, in general, ponerine ant venoms have peptides with similar masses, with most of the peptides below 4 kDa (Figure 5). In contrast, the formicoid venoms appear to have a broader peptide mass range with a significant number of peptides in the range 4–9 kDa (Figure 6). Peptide peak assignment resulted in between 54 (*E. brunneum* venom using

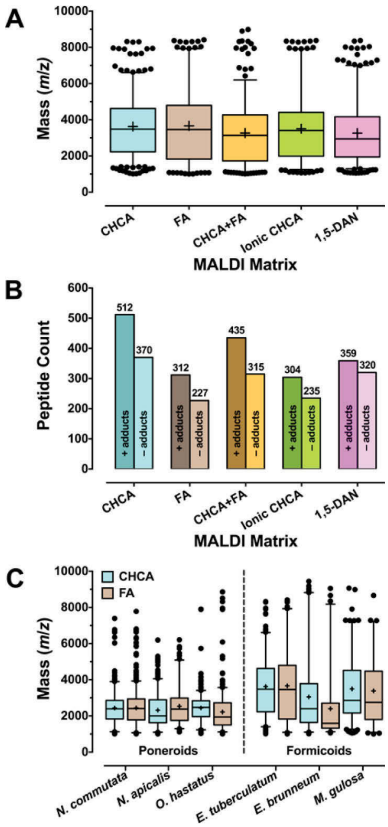


Figure 3. Number and distribution of *E. tuberculatum* venom peptides detected with different MALDI matrices. (A) Box-and-whisker plot depicting the mass distribution of peptides obtained with the different MALDI matrices. (B) Total peptide counts from all RP-HPLC fractions using the matrices CHCA, FA, a combination of CHCA and FA, ionic CHCA, and 1,5-DAN before and after removal of adducts, dimers, and doubly charged species. (C) Box-and-whisker plot showing ranges of peptide masses determined by CHCA (cyan boxes) and FA (beige boxes) matrices from the six venoms. The bottom and top of each box in panels A and C represent the first and third quartiles, respectively, while the band and cross inside each box represent the median and mean masses, respectively. The ends of the whiskers represent the 5–95 percentile range, while the closed circles represent masses outside the 5th to 95th percentile range.

FA matrix) and 1032 (*N. commutata* venom using FA matrix) distinguishable peptide masses for the six ant venoms. The two matrices gave similar numbers of peptides, with *N. commutata* being the only venom to have markedly different peptide counts using the two matrices.

Despite the relatively high peptide counts, there exists the possibility that the two matrices were detecting different peptides. Using the separate mass lists created using CHCA and FA matrices, masses common to both data sets (±1 Da) were identified using an in-house mass-matching Microsoft Excel spread sheet. This analysis revealed a large number of peptides that were only detected using one of the two matrices. For example, with *N. commutata* venom, only 245 peptides were detected using both FA and CHCA from a total of 1032

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX–XXX

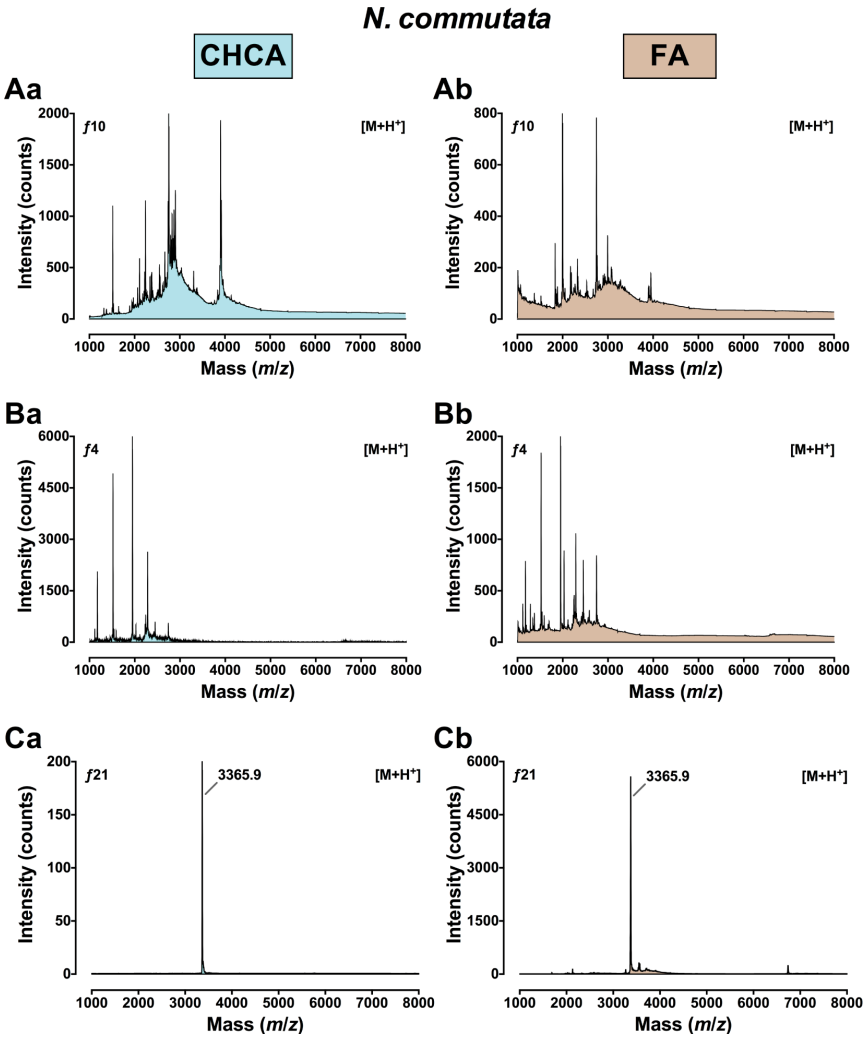


Figure 4. Representative MALDI-TOF MS spectra of *N. commutata* RP-HPLC fractions. MS Spectra were obtained in positive linear mode and fractions are labeled according to the RP-HPLC chromatogram shown in Figure 2A as *f*10 (A), *f*4 (B), and *f*21 (C). Left-hand panels labeled "a" with cyan shading were obtained using CHCA matrix, while right-hand panels labeled "b" with beige shading were obtained using FA matrix.

peptides detected using FA alone (24%) and 574 peptides using CHCA alone (43%). This leaves 787 unique peptides detected by FA alone and 329 unique peptides detected by CHCA alone. Together, the total mass count for *N. commutata* was therefore 1361 peptides (245 + 787 + 329). The data for all six venoms are presented in the form of area-proportional Euler plots in Figure 7.

Inter-Species Variations in Peptide Masses. If ant venoms are to be exploited as chemical libraries of unique peptides, it is important to determine if there is sufficient peptide heterogeneity between various ant venoms both across and within ant subfamilies. First, LC-MALDI-TOF mass spectrometry revealed that all ant venoms investigated in the present study showed differences in the broad distribution of masses within the range 1−10 kDa, with limited overlap (see Figures 5−7). Using the same mass lists generated from LC-MALDI-TOF MS, an analysis of masses common to different venoms was undertaken. Comparisons were made between ants belonging to the poneroid clade and those belonging to the formicoid clade. Analyses were carried out using both CHCA and FA as MALDI matrices to negate any bias in peptide identification.

Area-proportional Euler plots (Figure 8) revealed only limited overlap of masses between ants from the same clades, as only 16 peptides were common to all three poneroid ants and only 2 peptides were common between the three formicoid ants using CHCA matrix. These numbers were similar to those obtained using FA. In general, the poneroid ants appeared to have more peptides in common compared with formicoid ants, with *O. hastatus* and *N. commutata* sharing the highest number of peptides.

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

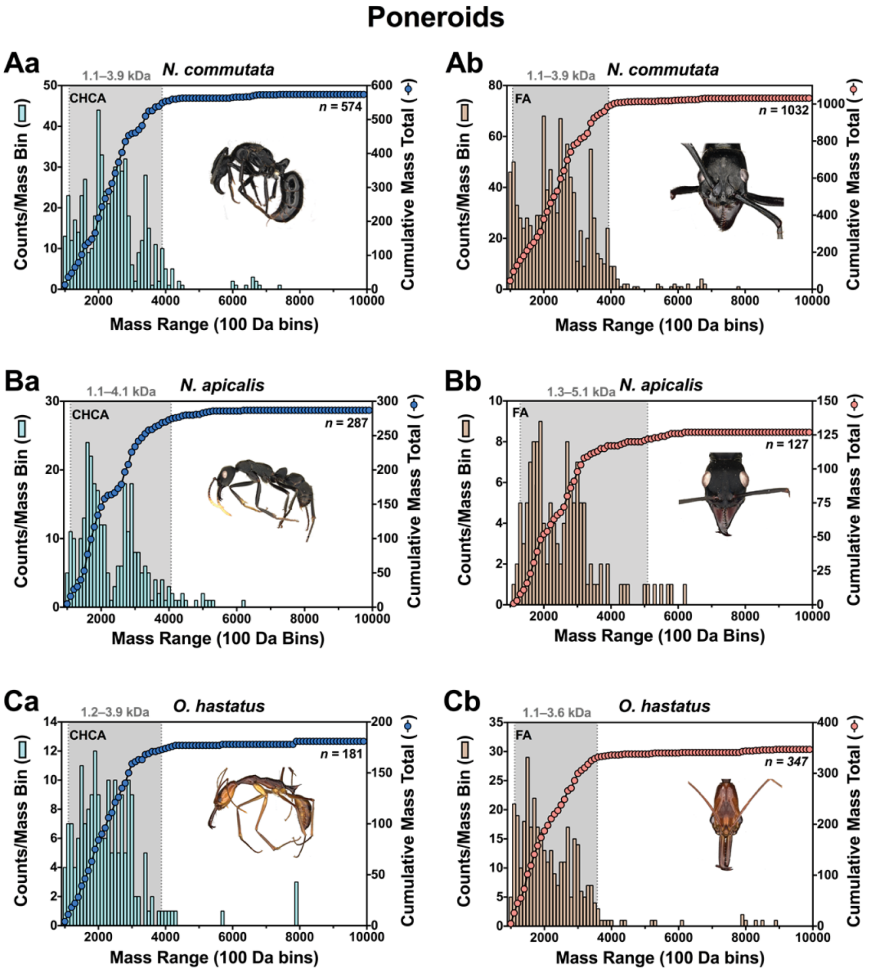


Figure 5. LC-MALDI-TOF mass profile of the poneroid ants *N. commutata*, *N. apicalis*, and *O. hastatus*. Histograms show the abundance of peptide masses in each venom sorted into 100 Da mass bins (columns; left-hand ordinate axis) and the cumulative total of peptide masses in each venom (circles; right-hand ordinate axis). Shaded areas represent 5−95% of the total number of peptides. (Aa−Ca) Peptide masses obtained using CHCA matrix (Aa) *N. commutata*, (Ba), *N. apicalis*, and (Ca) *O. hastatus*. (Ab−Cb) Peptide masses obtained using FA matrix (Ab) *N. commutata*, (Bb), *N. apicalis*, and (Cb) *O. hastatus*. Total peptide counts are indicated in the upper right-hand corner of each panel. All lateral and front images of the ants investigated were taken by Ryan Perry and can be found at www.AntWeb.org.

Disulfide Bond Identification. Peptides with disulfide bonds are highly stable and inherently more resistant to proteases.[29] Individual LC fractions were spotted on a MALDI target plate along with the reducing matrix 1,5-DAN and subjected to MALDI-TOF MS in reflector mode to identify disulfide-bonded peptides present in each ant venom. Fractions were also analyzed using the nonreducing matrix CHCA in reflector mode for comparison purposes. Each peptide peak obtained with the 1,5-DAN matrix was then analyzed for a +2 Da or greater shift from the native mass obtained using CHCA matrix (see Supplementary Figure S1).

The six venoms were found to contain variable numbers of disulfide-bonded peptides. This ranged from 2 disulfide-bonded peptides in *M. gulosa* venom to 28 disulfide-bonded peptides in *E. tuberculatum* venom (Figure 9A−F). These peptides were usually of low mass (typically <3 kDa) and distributed across the full HPLC chromatogram. The one exception was *E. tuberculatum* venom, where disulfide-bonded peptides were detected in two distinct groups in the ranges 1 to 2 kDa and 3 to 4 kDa (Figure 9G). As a percentage of the total number of peptides, *N. apicalis* had the greatest proportion of disulfide-bonded peptides (10.2%), followed by *O. hastatus* (9.0%), *E. tuberculatum* (7.6%), *E. brunneum* (5.0%), *N. commutata* (3.7%), and *M. gulosa* (1.4%). What can also be seen in the 2D landscapes in Figure 9A−F is that multiple peptides coelute in each LC fraction, with some fractions containing more than 50 peptides. It can also be seen that peptide mass distribution is independent of the RP-HPLC retention time (hydrophobicity).

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

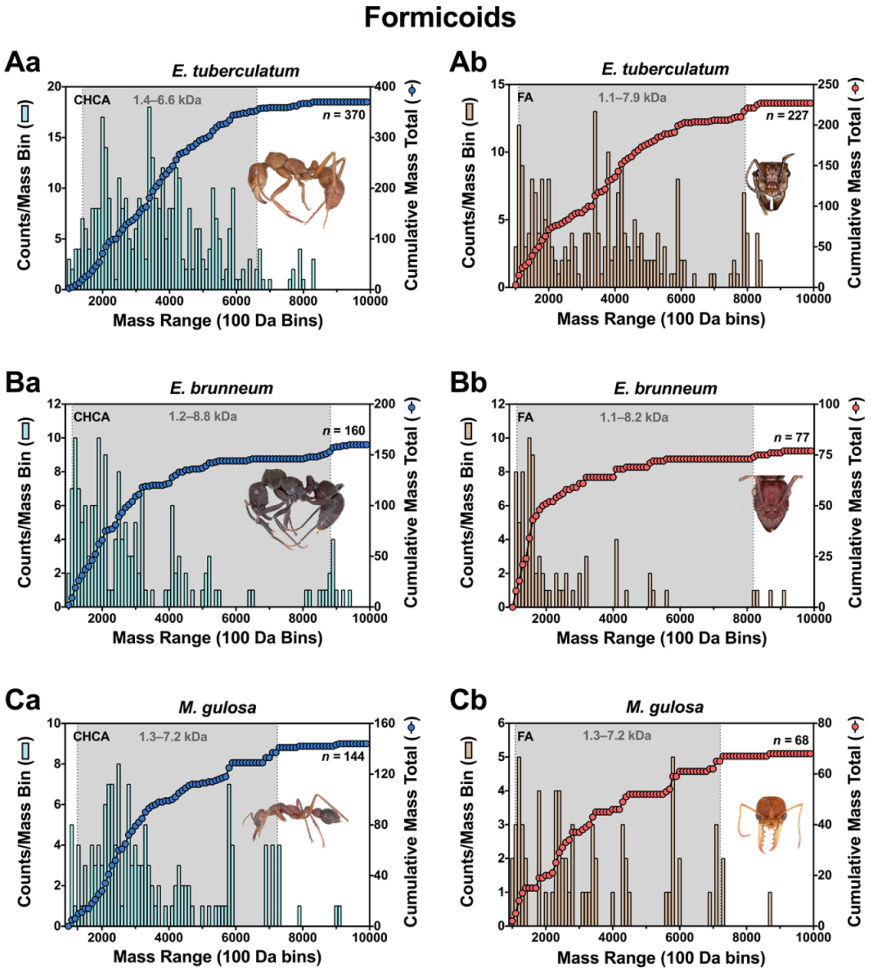


Figure 6. LC-MALDI-TOF mass profile of the formicoid ants *E. tuberculatum*, *E. brunneum*, and *M. gulosa*. Histograms show the abundance of peptide masses in each venom sorted into 100 Da mass bins (columns; left-hand ordinate axis) and the cumulative total of peptide masses in each venom (circles; right-hand ordinate axis). Shaded areas represent 5−95% of the total number of peptides. (Aa−Ca) Peptide masses obtained using CHCA matrix (Aa) *E. tuberculatum*, (Ba) *E. brunneum*, and (Ca) *M. gulosa*. (Ab−Cb) Peptide masses obtained using FA matrix (Ab) *E. tuberculatum*, (Bb) *E. brunneum*, and (Cb) *M. gulosa*. Total peptide counts are indicated in the upper right-hand corner of each panel. Lateral and front images of the ants investigated were taken by April Nobile, Estella Ortega, and Zach Lieberman and can be found at www.AntWeb.org.

DISCUSSION

Until recently, knowledge of the composition and properties of ant venoms was limited despite the fact that ants are among the most widespread of all social insects.[39] The present study constitutes the first characterization and comparison of the proteome and peptidome of ants belonging to the poneroid and formicoid clades.

Ant Venom Proteomes

The 1D- and 2D-PAGE gels confirmed the complexity of ant venom proteomes as well as the broad interspecies variation in protein expression. The 2D-PAGE separation of both *N. commutata* and *E. tuberculatum* resulted in over 70 spots being resolved, significantly higher than the 25 spots previously described from *S. invicta* venom, one of the only ant venoms on which a 2D-PAGE has been performed.[16,24] Not surprisingly, there were some proteins that were shared between *N. commutata* and *E. tuberculatum* venoms after examination by 2D-PAGE. The 90 kDa protein band present in all six venoms from the 1D-PAGE and confirmed in *E. tuberculatum* by 2D-PAGE was identified as dipeptidyl peptidase IV. There was also a band common to all six venoms at ca. 24 kDa in the 1D-PAGE gels, identified as a PLA_2.[17,40] This was confirmed by the nanoLC-ESI-QTOF MS/MS identification of PLA_2 in both 2D-PAGE gels around this molecular weight range. This is a common allergenic protein of ant venoms[17] and other hymenopterans,[41] although the PLA_2 from the fire ant *Solenopsis invicta* is not considered as an allergenic protein.[42]

Overall, the majority of spots were assigned to proteins from hymenopteran species, where there is a high conservation of

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

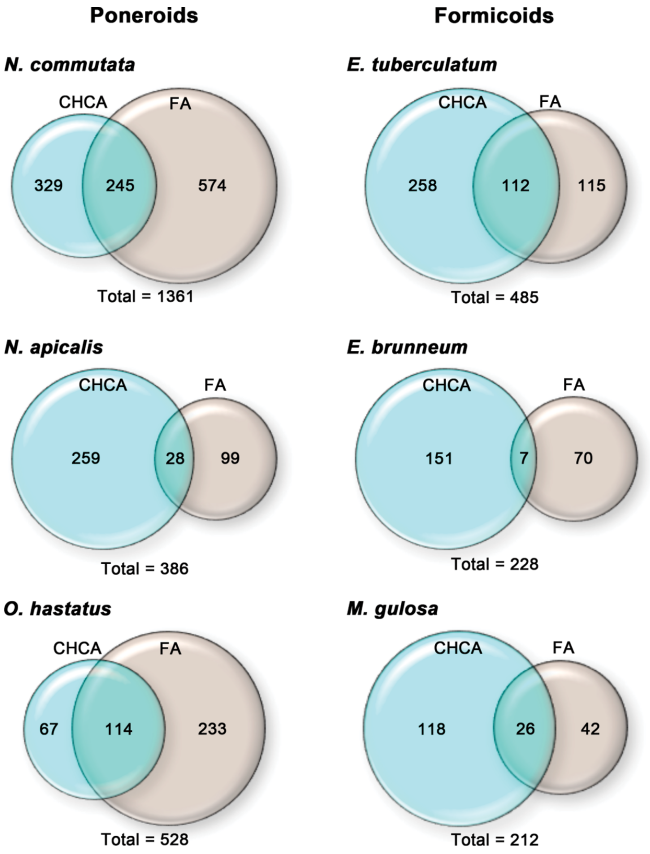


Figure 7. Area-proportional Euler plots depicting numbers of shared and distinct peptides. Shared peptides were defined as being within ±1 Da. Spectra were acquired with CHCA (cyan) and FA (beige) MALDI matrices. The numbers within the circles indicate the number of unique peptides detected employing that particular matrix. The total number of peptides detected by each matrix is indicated in the box on the bottom of the Euler plots. Area-proportional Euler plots depicting overlapping peptide masses from multiple venoms were constructed using Venn diagram plotter.

protein sequences, especially between bees, wasps, and ants.[43,44] For example, the known insect venom allergen lysosomal acid phosphatase, also found in honeybees, was identified in both ant venoms.[45−47] An interesting protein that was identified in the *N. commutata* gel was the antimicrobial, insecticidal, and hemolytic peptide ponericin. This peptide was first isolated from the venom of the closely related ant *N. goeldii* from the same genus[11] and has been found in other ants.[15] This protein was identified in several spots throughout the gel, which suggests the presence of a precursor protein that is cleaved into ponericin.

A small number of the spots were identified as contaminants or artifacts of the venom extraction procedure, as the identified proteins are not believed to represent venom peptide toxins. For example, actin present in the *N. commutata* venom is most likely a contaminant of the exoskeleton[48] or reservoir gland,[49] while endochitinase present in *E. tuberculatum* venom hydrolyses chitin, an important protein for insect growth and morphogenesis.[50] The enzyme trehalase identified in *N. commutata* venom is also believed to be another contaminant, as it is a hemolymph protein involved in the rapid supply of energy for insects by converting trehalose to glucose.[51]

Importantly, however, a large number of the peptides had no similarity to any known sequences, which highlights the uniqueness of ant venom and also the low number of available sequences in protein sequence databases. This problem has been encountered by other studies looking at hymenopteran venoms.[23,39,52] Moreover, there was a lack of separation of the basic, low mass peptides, especially those present in *N. commutata* venom; hence, the 2D-PAGE technique was unable to resolve the identity of these peptides resulting in the need to employ an alternative method for peptide characterization.

Ant Venom Peptidomes

The venoms of ants and other hymenoptera are known sources of biologically active, low-molecular-weight peptides.[3,53,54] Unlike snakes, spiders, scorpions, and marine cone snails,[37,55−57] most ant venom peptidomes remain barely explored. Accordingly, the present study represents the first in-depth study of a range of ant venom peptides using highly sensitive MS technologies.

MALDI Matrix Optimisation. Numerous matrix candidates have been evaluated, but CHCA has become the favored matrix for the analysis of peptides.[58−61] Our results confirmed recent reports that different matrices yield different spectra in MALDI-TOF MS.[62,63] We found that CHCA and FA matrices gave the

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

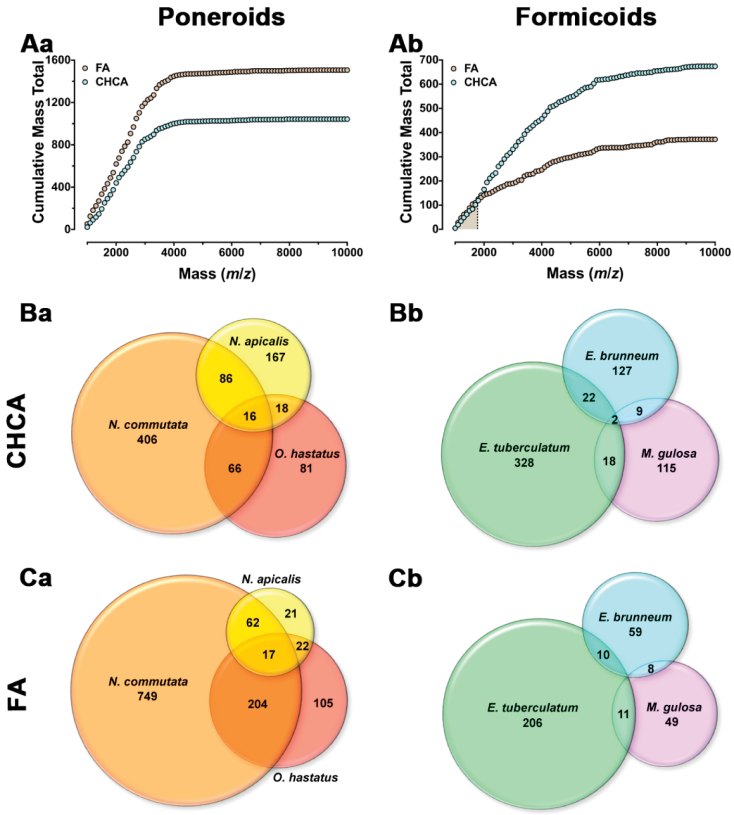


Figure 8. (A) Cumulative total of peptides detected using CHCA (cyan circles) or FA (beige circles) from the poneroid ants *N. commutata*, *P. apicalis*, and *O. hastatus* (Aa) or formicoid ants *O. hastatus*, *O. brunneum*, and *M. gulosa* (Ab). Panel Ab gray shaded area, FA was more efficient at detecting lower molecular weight peptides below 1800 Da in the formicoid venoms. (B,C) Common peptides between ant clades. Area-proportional Euler plot depicting species variation in numbers of shared and distinct peptides (±1 Da) determined using MALDI-TOF MS. Left-hand panels labeled "a" contain data from poneroid ants; right-hand panels labeled "b" show a comparison of the formicoid ant data. Panels B refer to the venom peptides detected using CHCA matrix, while Panels C represent peptides detected using FA matrix.

highest peptide counts with the broadest mass distribution, which is in accord with previous studies.[38] Moreover, the range of peptides detected by the two matrices revealed that FA was most suited for poneroid ant venoms and was slightly more effective at detecting low mass peptides <1800 Da in formicoid venoms, while CHCA was better at detecting peptides in formicoid venoms with a wider distribution of masses, for example, *E. tuberculatum* and *M. gulosa* (Figures 7 and 8). It was also determined that FA resulted in the formation of fewer adducts compared with CHCA (Figure 3B), as has been previously reported.[64] This is most likely due to its increased tolerance to salts and other impurities.[65] Nevertheless, a consistent difference in signal intensity was noted that has been reported to be higher with CHCA.[65]

Ant Venom Peptide Numbers and Distribution. The presence of higher abundance peptides and proteins in the sample results in undetected or unresolved signals from quasi-isobaric peptides and lower abundance peptide components. Accordingly, ant venoms were subjected to RP-HPLC separation prior to MALDI-TOF MS, revealing a complex mixture of peptides due to the separation and thus increased detection of trace levels of peptides. In addition, this overcame the ion suppression phenomenon by reducing the number of co-ionised peptides in the mass spectrometer source.[37,66]

The different ant venoms did not follow the same mass distribution, reflecting their peptide diversity. This is not unique to ants, as both bimodal[37] and unimodal[67] distributions have been previously reported with other venoms. Most of the ant venoms analyzed have the majority of their peptides in the low-molecular-weight range (1−4 kDa). This suggests that ant venoms have peptides that are overall smaller than the peptidic component of other venomous animals such as scorpions (4−5 kDa),[68] sea anemones (2−3.5 kDa),[69] and spiders (3−5 kDa and 6.5−8.5 kDa)[70] but not marine cone snails (0.5−3.5 kDa).[55] However, formicoid ant venoms seem to be more complex than poneroid ant venoms and possess some larger peptides with masses up to 8 kDa. These results confirm the conclusions of previous studies about ant venom peptides using lower resolution MALDI-TOF MS than the present study.[38,53] However, we found that the poneroid ant *N. apicalis* has a complex venom despite the fact that it is a highly specialized predator of only *Syntermes* termites.[71] Its venom complexity is comparable to some spider and cone snail venoms.[55,72] This suggests that predation is not the primary factor in venom

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

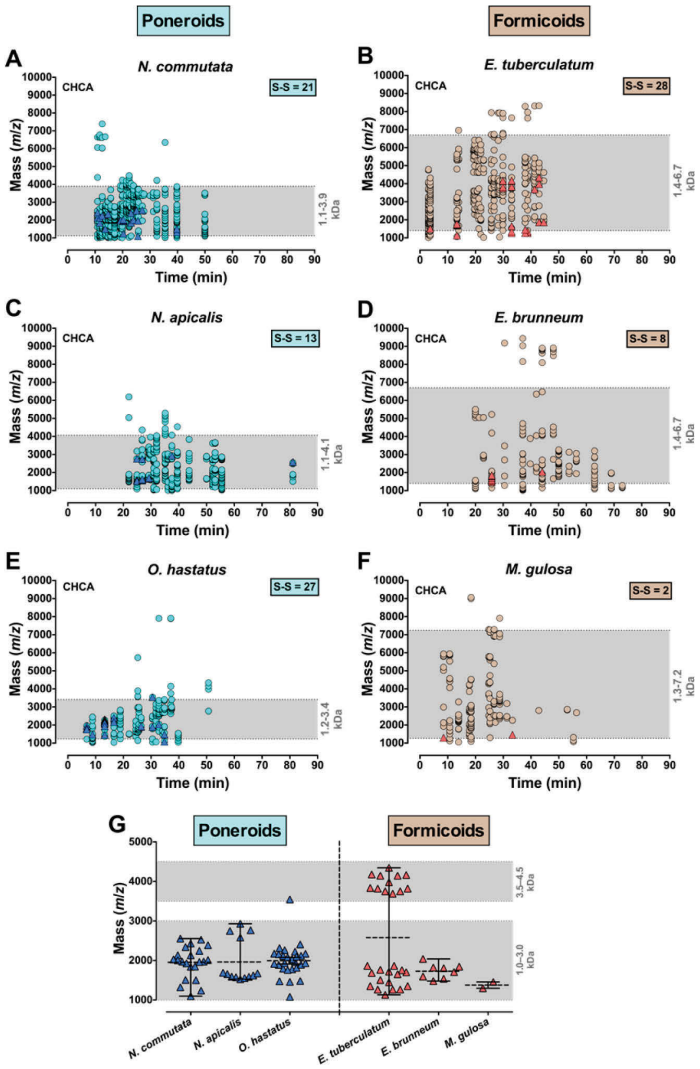


Figure 9. Determination of disulfide-bonded peptides in ant venoms. MALDI-TOF MS analysis of individual RP-HPLC fractions from five ant venoms. (A–F) 2D Landscape plots where the abscissa represents the C18 RP-HPLC retention time while the ordinate shows the masses obtained by MALDI-TOF MS. Cyan circles represent the nondisulfide-bonded peptides in the venom. Red triangles indicate peptides with a disulfide bond. Total number of disulfide-bonded peptides in each venom is indicated in the top right-hand box in each panel (S–S). (A) *N. commutata*, (B) *N.* apicalis, (C) *O. hastatus*, (D) *E. tuberculatum*, (E) *E. brunneum*, and (F) *M. gulosa*. The shaded areas represent 5–95% of the total number of peptides. (G) Scatter dot plot showing all disulfide-bonded masses from each venom; the dotted horizontal line represents the mean mass of disulfide-bonded peptides while the solid bars represent the mass range.

composition and that defense against predators and pathogens may play a larger role in the diversification of ant venom toxins.

Given the paucity of data on ant venom peptides, the present study highlights that the enormous biodiversity of ant venoms remains almost completely unexplored. If one assumes each species of ant has on average 600 unique venom peptides (average of the six ant venoms), the total biodiversity present from around 9000 species (ca. 70% of the total 13 000 extant species) of venomous ants in the world is ca. 5 400 000 distinct peptides (Aili et al., 2014). Thus, ant venoms represent an enormous pharmacological library that is significantly larger than scorpions (ca. 100 000 peptides)[57] and marine cone snail venoms (ca. 500 000 peptides)[55] but is comparable to that of spiders (ca. 10 000 000 peptides).[37] Of course, this fails to account for common peptide masses across all species (Figure 8). Hence a more conservative estimation of ant venom peptide biodiversity is around 2 000 000+ peptides. To place this potential resource in perspective, only 75 peptides from the enormous chemical library of ant venom peptidomes have been described to date, representing a mere 0.003%.

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX–XXX

Disulfide-Bonded Venom Peptides. Disulfide bonds render peptides very stable and resistant to a number of proteases and environmental extremes.[28,73] This is an advantageous feature and consequently has become a focus in the search for stable therapeutic or insecticidal peptide toxins.[28,73] Indeed, all eight venom-derived drugs on the pharmaceutical drug market contain between 1−14 disulfide bonds.[73] Furthermore, other venomous animals such as cone snails, scorpions, and spiders contain a high proportion of disulfide-rich peptides.[37,73−75] While the present study confirms previous findings that ant venom is believed to be a rich source of small linear peptides,[53] disulfide-bonded peptides are minor components of ant peptidomes; however, biological activities and functions of these reticulated peptides still remain unknown and should be investigated in future studies as potential sources of insecticides.

CONCLUDING REMARKS

This proteomic-based study has revealed not only the vast complexity of ant venom proteins and peptides but also their inter- and intrafamily variation. The mass range of ant venom peptides, particularly those from poneroid ants, is lower than that of peptides from other venoms, indicating likely novel structures and pharmacologies. The peptide mass distributions varied across ant clades and between different ant genera from the same clade, reflecting their peptide diversity. Also, given that some ant venoms were complex despite only preying upon one prey species suggests that there are likely a wide variety of bioactive molecules involved in predator defense, communication, and host defense against infection. This peptide diversity highlights the utility of ant venoms as novel chemical peptide libraries with the potential to be sources of novel therapeutic or insecticide leads.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.6b00182.

Methodology for the identification of disulfide-bonded peptides as well as annotated spectra from 2D gel spots where a protein was identified from a single peptide. (PDF)

AUTHOR INFORMATION

Corresponding Author

*Phone: (+61) 9514 2230. Fax: (+61) 9514 1656. E-mail: Graham.Nicholson@uts.edu.au.

Author Contributions

The manuscript was written through contributions of all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support for this study was provided by an Australian Postgraduate Award to Samira R. Aili, the *Programme Convergence 2007-2013, Région Guyane* from the European Community (BI-Appli, 115/SGAR-DE/2011/052274) to A.D. and an *Investissement d'Avenir* grant managed by the Agence Nationale de la Recherche (CEBA, ANR-10-LABX-25-01).

ABBREVIATIONS

ACN, acetonitrile; C7BzO, 3-(4-heptyl)phenyl-3-hydroxypropyl) dimethylammoniopropanesulfonate; 1,5-DAN, 1,5-diaminonapthalene; CHCA, α-cyano-4-hydroxycinnamic acid; FA, ferulic acid; HGD, Hymenopteran Genome Database; IEF, isoelectric focusing; ionic CHCA, *N*-*tert*-butyl-*N*-isopropyl-*N*-methylammonium α-cyano-4-hydroxycinnamate; IPG, immobilized pH gradient; LC-ESI-QTOF MS, liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry; MALDI-TOF MS, matrix-assisted laser-desorption ionization time-of-flight mass spectrometry; PLA_2, phospholipase A_2; RP-HPLC, reverse-phase high-pressure liquid chromatography; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid

REFERENCES

(1) Windley, M. J.; Herzig, V.; Dziemborowicz, S. A.; Hardy, M. C.; King, G. F.; Nicholson, G. M. Spider-venom peptides as bioinsecticides. *Toxins* **2012**, *4* (3), 191−227.

(2) Coddington, J. A.; Levi, H. W. Systematics and evolution of spiders (Araneae). *Annu. Rev. Ecol. Syst.* **1991**, *22* (1), 565−592.

(3) Aili, S. R.; Touchard, A.; Escoubas, P.; Padula, M. P.; Orivel, J.; Dejean, A.; Nicholson, G. M. Diversity of peptide toxins from stinging ant venoms. *Toxicon* **2014**, *92*, 166−78.

(4) Touchard, A.; Aili, S.; Fox, E.; Escoubas, P.; Orivel, J.; Nicholson, G.; Dejean, A. The biochemical toxin arsenal from ant venoms. *Toxins* **2016**, *8* (1), 30.

(5) Johnson, S. R.; Copello, J. A.; Evans, M. S.; Suarez, A. V. A biochemical characterization of the major peptides from the venom of the giant neotropical hunting ant *Dinoponera australis*. *Toxicon* **2010**, *55* (4), 702−10.

(6) Riffle, A.; Gavalda, S.; Téné, N.; Orivel, J.; Leprince, J.; Guilhaudis, L.; Génin, E.; Vétillard, A.; Treilhou, M. Identification and characterization of a novel antimicrobial peptide from the venom of the ant *Tetramorium bicarinatum*. *Peptides* **2012**, *38* (2), 363−370.

(7) *AntWeb*, 2014. Available online: http://www.antweb.org (accessed 27 August 2014).

(8) *Ant Ecology*; Lach, L., Parr, C. L., Abott, K. L., Eds.; Oxford University Press: Oxford, U.K., 2010.

(9) Ward, P. S. Taxonomy, Phylogenetics, and Evolution. In *Ant Ecology*; Lach, L., Parr, C. L., Abott, K. L., Eds.; Oxford University Press: Oxford, U.K., 2010; Vol. *1*.

(10) Cerdá, X.; Dejean, A. Predation by Ants on Arthropods and Other Animals. In *Predation in the Hymenoptera: An Evolutionary Perspective*; Polidori, C., Ed.; Transworld Research Network: Kerala - India, 2011; pp 39−78.

(11) Orivel, J.; Redeker, V.; Le Caer, J. P.; Krier, F.; Revol-Junelles, A. M.; Longeon, A.; Chaffotte, A.; Dejean, A.; Rossier, J. Ponericins, new antibacterial and insecticidal peptides from the venom of the ant *Pachycondyla goeldii*. *J. Biol. Chem.* **2001**, *276* (21), 17823−9.

(12) Wilson, E. O. *Success and Dominance in Ecosystems: The Case of the Social Insects*; Ecology Institute: Oldendorf/Luhe, Germany, 1990; Vol. *2*.

(13) Ebaid, H.; Al-Khalifa, M.; Isa, A. M.; Gadoa, S. Bioactivity of Samsum ant (*Pachycondyla sennaarensis*) venom against lipopolysaccharides through antioxidant and upregulation of Akt1 signaling in rats. *Lipids Health Dis.* **2012**, *11* (1), 93.

(14) Inagaki, H.; Akagi, M.; Imai, H. T.; Taylor, R. W.; Wiese, M. D.; Davies, N. W.; Kubo, T. Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (Jack Jumper Ant). *Arch. Biochem. Biophys.* **2008**, *477* (2), 411−416.

(15) dos Santos Pinto, J. R.; Fox, E. G.; Saidemberg, D. M.; Santos, L. D.; da Silva Menegasso, A. R.; Costa-Manso, E.; Machado, E. A.; Bueno, O. C.; Palma, M. S. Proteomic view of the venom from the fire ant Solenopsis invicta Buren. *J. Proteome Res.* **2012**, *11* (9), 4643−4653.

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

(16) Wiese, M. D.; Brown, S. G.; Chataway, T. K.; Davies, N. W.; Milne, R. W.; Aulfrey, S. J.; Heddle, R. J. *Myrmecia pilosula* (Jack Jumper) ant venom: identification of allergens and revised nomenclature. *Allergy* **2007**, *62* (4), 437−43.
(17) Hink, W. F.; Pappas, P. W.; Jaworski, D. C. Partial biochemical characterization of venom from the ant, *Pseudomyrmex triplarinus*. *Toxicon* **1994**, *32* (7), 763−72.
(18) Pan, J.; Hink, W. F. Isolation and characterization of myrmexins, six isoforms of venom proteins with anti-inflammatory activity from the tropical ant, *Pseudomyrmex triplarinus*. *Toxicon* **2000**, *38* (10), 1403−13.
(19) Haddad Junior, V.; Cardoso, J. L. C.; Moraes, R. H. P. Description of an injury in a human caused by a false tocandira (*Dinoponera gigantea*, Perty, 1833) with a revision on folkloric, pharmacological and clinical aspects of the giant ants of the genera Paraponera and Dinoponera (sub-family Ponerinae). *Rev. Inst. Med. Trop. Sao Paulo* **2005**, *47* (4), 235−238.
(20) Piek, T.; Duval, A.; Hue, B.; Karst, H.; Lapied, B.; Mantel, P.; Nakajima, T.; Pelhate, M.; Schmidt, J. O. Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata*. *Comp. Biochem. Physiol., C: Comp. Pharmacol.* **1991**, *99* (3), 487−495.
(21) Johnson, S. R.; Rikli, H. G.; Schmidt, J. O.; Evans, M. S. A reexamination of poneratoxin from the venom of the bullet ant Paraponera clavata. *Peptides* **2016**, DOI: 10.1016/j.peptides.2016.05.012.
(22) Pluzhnikov, K.; Nosyreva, E.; Shevchenko, L.; Kokoz, Y.; Schmalz, D.; Hucho, F.; Grishin, E. Analysis of ectatomin action on cell membranes. *Eur. J. Biochem.* **1999**, *262* (2), 501−6.
(23) Torres, A. F. C.; Huang, C.; Chong, C.-M.; Leung, S. W.; Prieto-da-Silva, A. R. B.; Havt, A.; Quinet, Y. P.; Martins, A. M. C.; Lee, S. M. Y.; Rádis-Baptista, G. Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: insights into the polypeptide toxin arsenal of Hymenopterans. *PLoS One* **2014**, *9* (1), e87556.
(24) Gonçalves Paterson Fox, E.; Russ Solis, D.; Delazari Dos Santos, L.; Aparecido Dos Santos Pinto, J. R.; Ribeiro da Silva Menegasso, A.; Cardoso Maciel Costa Silva, R.; Sergio Palma, M.; Correa Bueno, O.; de Alcantara Machado, E. A simple, rapid method for the extraction of whole fire ant venom (Insecta: Formicidae: *Solenopsis*). *Toxicon* **2013**, *65*, 5−8.
(25) Escoubas, P.; King, G. F. Venomics as a drug discovery platform. *Expert Rev. Proteomics* **2009**, *6*, 221.
(26) Vetter, I.; Davis, J. L.; Rash, L. D.; Anangi, R.; Mobli, M.; Alewood, P. F.; Lewis, R. J.; King, G. F. Venomics: a new paradigm for natural products-based drug discovery. *Amino Acids* **2011**, *40* (1), 15−28.
(27) Nolde, D. E.; Sobol, A. G.; Pluzhnikov, K. A.; Grishin, E. V.; Arseniev, A. S. Three-dimensional structure of ectatomin from *Ectatomma tuberculatum* ant venom. *J. Biomol. NMR* **1995**, *5* (1), 1−13.
(28) Saez, N. J.; Senff, S.; Jensen, J. E.; Er, S. Y.; Herzig, V.; Rash, L. D.; King, G. F. Spider-venom peptides as therapeutics. *Toxins* **2010**, *2* (12), 2851−2871.
(29) King, G. F.; Hardy, M. C. Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. *Annu. Rev. Entomol.* **2013**, *58* (1), 475−96.
(30) Deutscher, A. T.; Jenkins, C.; Minion, F. C.; Seymour, L. M.; Padula, M. P.; Dixon, N. E.; Walker, M. J.; Djordjevic, S. P. Repeat regions R1 and R2 in the P97 paralogue Mhp271 of *Mycoplasma hyopneumoniae* bind heparin, fibronectin and porcine cilia. *Mol. Microbiol.* **2010**, *78* (2), 444−458.
(31) Jobbins, S. E.; Hill, C. J.; D'Souza-Basseal, J. M.; Padula, M. P.; Herbert, B. R.; Krockenberger, M. B. Immunoproteomic approach to elucidating the pathogenesis of cryptococcosis caused by *Cryptococcus gattii*. *J. Proteome Res.* **2010**, *9* (8), 3832−3841.
(32) Raymond, B. B. A.; Tacchi, J. L.; Jarocki, V. M.; Minion, F. C.; Padula, M. P.; Djordjevic, S. P. P159 from *Mycoplasma hyopneumoniae* binds porcine cilia and heparin and Is cleaved in a manner akin to ectodomain shedding. *J. Proteome Res.* **2013**, *12* (12), 5891−5903.
(33) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551−3567.
(34) Jaskolla, T. W.; Karas, M.; Roth, U.; Steinert, K.; Menzel, C.; Reihs, K. Comparison between vacuum sublimed matrices and conventional dried droplet preparation in MALDI-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1104−1114.
(35) Escoubas, P.; Sollod, B.; King, G. F. Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon* **2006**, *47* (6), 650−663.
(36) Fukuyama, Y.; Iwamoto, S.; Tanaka, K. Rapid sequencing and disulfide mapping of peptides containing disulfide bonds by using 1, 5-diaminonaphthalene as a reductive matrix. *J. Mass Spectrom.* **2006**, *41* (2), 191−201.
(37) Palagi, A.; Koh, J. M.; Leblanc, M.; Wilson, D.; Dutertre, S.; King, G. F.; Nicholson, G. M.; Escoubas, P. Unravelling the complex venom landscapes of lethal Australian funnel-web spiders (*Hexathelidae: Atracinae*) using LC-MALDI-TOF mass spectrometry. *J. Proteomics* **2013**, *80*, 292−310.
(38) Touchard, A.; Dauvois, M.; Arguel, M.-J.; Petitclerc, F.; Leblanc, M.; Dejean, A.; Orivel, J.; Nicholson, G. M.; Escoubas, P. Elucidation of the unexplored biodiversity of ant venom peptidomes via MALDI-TOF mass spectrometry and its application for chemotaxonomy. *J. Proteomics* **2014**, *105*, 217−231.
(39) Bouzid, W.; Klopp, C.; Verdenaud, M.; Ducancel, F.; Vetillard, A. Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (Hymenoptera: Formicidae): the first transcriptome analysis of an ant species. *Toxicon* **2013**, *70*, 70−81.
(40) Leluk, J.; Schmidt, J.; Jones, D. Comparative studies on the protein composition of hymenopteran venom reservoirs. *Toxicon* **1989**, *27* (1), 105−114.
(41) Peiren, N.; de Graaf, D. C.; Vanrobaeys, F.; Danneels, E. L.; Devreese, B.; Van Beeumen, J.; Jacobs, F. J. Proteomic analysis of the honey bee worker venom gland focusing on the mechanisms of protection against tissue damage. *Toxicon* **2008**, *52* (1), 72−83.
(42) King, T. P.; Spangfort, M. D. Structure and biology of stinging insect venom allergens. *Int. Arch. Allergy Immunol.* **2000**, *123* (2), 99−106.
(43) Wanstall, J. C.; De la Lande, I. Fractionation of bulldog ant venom. *Toxicon* **1974**, *12* (6), 649−IN16.
(44) Bonasio, R.; Zhang, G.; Ye, C.; Mutti, N. S.; Fang, X.; Qin, N.; Donahue, G.; Yang, P.; Li, Q.; Li, C.; Zhang, P.; Huang, Z.; Berger, S. L.; Reinberg, D.; Wang, J.; Liebig, J. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **2010**, *329* (5995), 1068−71.
(45) Resende, V. M. F.; Vasilj, A.; Santos, K. S.; Palma, M. S.; Shevchenko, A. Proteome and phosphoproteome of Africanized and European honeybee venoms. *Proteomics* **2013**, *13* (17), 2638−2648.
(46) Hoffman, D. R. Allergens in bee venom: III. Identification of allergen B of bee venom as an acid phosphatase. *J. Allergy Clin. Immunol.* **1977**, *59* (5), 364−366.
(47) Li, R.; Zhang, L.; Fang, Y.; Han, B.; Lu, X.; Zhou, T.; Feng, M.; Li, J. Proteome and phosphoproteome analysis of honeybee (*Apis mellifera*) venom collected from electrical stimulation and manual extraction of the venom gland. *BMC Genomics* **2013**, *14*, 766.
(48) Kheyri, H.; Cribb, B. W.; Merritt, D. J. Comparing the secretory pathway in honeybee venom and hypopharyngeal glands. *Arthropod Struct. Dev.* **2013**, *42* (2), 107−114.
(49) do Amaral, J. B.; Machado-Santelli, G. M. Salivary system in leaf-cutting ants (*Atta sexdens rubropilosa* Forel, 1908) castes: A confocal study. *Micron* **2008**, *39* (8), 1222−1227.
(50) Merzendorfer, H.; Zimoch, L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* **2003**, *206* (Pt 24), 4393−412.
(51) Mori, H.; Lee, J.-H.; Okuyama, M.; Nishimoto, M.; Ohguchi, M.; Kim, D.; Kimura, A.; Chiba, S. Catalytic reaction mechanism based on alpha-secondary deuterium isotope effects in hydrolysis of trehalose

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

by European honeybee trehalase. *Biosci., Biotechnol., Biochem.* **2009**, *73* (11), 2466−2473.
(52) Sookrung, N.; Wong-din-Dam, S.; Tungtrongchitr, A.; Reamtong, O.; Indrawattana, N.; Sakolvaree, Y.; Visitsunthorn, N.; Manuyakorn, W.; Chaicumpa, W. Proteome and allergenome of Asian wasp, Vespa affinis, venom and IgE reactivity of the venom components. *J. Proteome Res.* **2014**, *13* (3), 1336−44.
(53) Touchard, A.; Koh, J.; Aili, S. R.; Dejean, A.; Nicholson, G. M.; Orivel, J.; Escoubas, P. The complexity and structural diversity of ant venom peptidomes is revealed by mass spectrometry profiling. *Rapid Commun. Mass Spectrom.* **2015**, *29* (5), 385−396.
(54) Wiese, M. D.; Chataway, T. K.; Davies, N. W.; Milne, R. W.; Brown, S. G.; Gai, W. P.; Heddle, R. J. Proteomic analysis of *Myrmecia pilosula* (jack jumper) ant venom. *Toxicon* **2006**, *47* (2), 208−17.
(55) Davis, J.; Jones, A.; Lewis, R. J. Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides* **2009**, *30* (7), 1222−1227.
(56) Escoubas, P.; Quinton, L.; Nicholson, G. M. Venomics: unravelling the complexity of animal venoms with mass spectrometry. *J. Mass Spectrom.* **2008**, *43* (3), 279−295.
(57) Possani, L. D.; Becerril, B.; Delepierre, M.; Tytgat, J. Scorpion toxins specific for Na^+ channels. *Eur. J. Biochem.* **1999**, *264* (2), 287−300.
(58) Řehulková, H.; Chalupová, J.; Šebela, M.; Řehulka, P. A convenient purification and preconcentration of peptides with α-cyano-4-hydroxycinnamic acid matrix crystals in a pipette tip for matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **2009**, *45* (1), 104−111.
(59) Kouvonen, P.; Rainio, E. M.; Suni, V.; Koskinen, P.; Corthals, G. L. Data combination from multiple matrix-assisted laser desorption/ionization (MALDI) matrices: opportunities and limitations for MALDI analysis. *Rapid Commun. Mass Spectrom.* **2010**, *24* (23), 3493−3495.
(60) Ahn, S. H.; Park, K. M.; Bae, Y. J.; Kim, M. S. Quantitative reproducibility of mass spectra in matrix-assisted laser desorption ionization and unraveling of the mechanism for gas-phase peptide ion formation. *J. Mass Spectrom.* **2013**, *48* (3), 299−305.
(61) Rappsilber, J.; Moniatte, M.; Nielsen, M. L.; Podtelejnikov, A. V.; Mann, M. Experiences and perspectives of MALDI MS and MS/MS in proteomic research. *Int. J. Mass Spectrom.* **2003**, *226* (1), 223−237.
(62) Matysiak, J.; Schmelzer, C. E.; Neubert, R. H.; Kokot, Z. J. Characterization of honeybee venom by MALDI-TOF and nanoESI-QqTOF mass spectrometry. *J. Pharm. Biomed. Anal.* **2011**, *54* (2), 273−278.
(63) Lai, Y.-H.; Wang, C.-C.; Chen, C. W.; Liu, B.-H.; Lin, S. H.; Lee, Y. T.; Wang, Y.-S. Analysis of initial reactions of MALDI based on chemical properties of matrixes and excitation condition. *J. Phys. Chem. B* **2012**, *116* (32), 9635−9643.
(64) Yang, H.; Wan, D.; Song, F.; Liu, Z.; Liu, S. α-Cyano-4-hydroxycinnamic acid, sinapinic acid, and ferulic acid as matrices and alkylating agents for matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis of cysteine-containing peptides. *Rapid Commun. Mass Spectrom.* **2013**, *27* (12), 1410−1412.
(65) Beavis, R.; Chaudhary, T.; Chait, B. α-Cyano-4-hydroxycinnamic acid as a matrix for matrixassisted laser desorption mass spectromtry. *Org. Mass Spectrom.* **1992**, *27* (2), 156−158.
(66) Trimpin, S.; Wang, B.; Lietz, C. B.; Marshall, D. D.; Richards, A. L.; Inutan, E. D. New ionization processes and applications for use in mass spectrometry. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48* (5), 409−429.
(67) Abdel-Rahman, M. A.; Quintero-Hernandez, V.; Possani, L. D. Venom proteomic and venomous glands transcriptomic analysis of the Egyptian scorpion *Scorpio maurus palmatus* (Arachnida: Scorpionidae). *Toxicon* **2013**, *74*, 193−207.
(68) Martin-Eauclaire, M.-F.; Céard, B.; Belghazi, M.; Lebrun, R.; Bougis, P. E. Characterization of the first K^+ channel blockers from the venom of the Moroccan scorpion *Buthus occitanus* Paris. *Toxicon* **2013**, *75*, 168−176.
(69) Rodriguez, A. A.; Staendker, L.; Zaharenko, A. J.; Garateix, A. G.; Forssmann, W.-G.; Beress, L.; Valdes, O.; Hernandez, Y.; Laguna, A. Combining multidimensional liquid chromatography and MALDI−TOF-MS for the fingerprint analysis of secreted peptides from the unexplored sea anemone species *Phymanthus crucifer*. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2012**, *903*, 30−39.
(70) Nicholson, G. M., Spider Venom Peptides. In *Handbook of Biologically Active Peptides*, 2nd ed.; Kastin, A. J., Ed.; Academic Press: San Diego, 2013.
(71) Mill, A. E. Predation by the ponerine ant *Pachycondyla commutata* on termites of the genus Syntermes in Amazonian rain forest. *J. Nat. History* **1984**, *18* (3), 405−410.
(72) Escoubas, P. Molecular diversification in spider venoms: a web of combinatorial peptide libraries. *Mol. Diversity* **2006**, *10* (4), 545−554.
(73) King, G. F. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther.* **2011**, *11* (11), 1469−1484.
(74) Palma, M. S. Peptides as toxins/defensins. *Amino Acids* **2011**, *40* (1), 1−4.
(75) Quinton, L.; Demeure, K.; Dobson, R.; Gilles, N.; Gabelica, V.; De Pauw, E. New method for characterizing highly disulfide-bridged peptides in complex mixtures: application to toxin identification from crude venoms. *J. Proteome Res.* **2007**, *6* (8), 3216−3223.
(76) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31* (13), 3784−3788.

DOI: 10.1021/acs.jproteome.6b00182
J. Proteome Res. XXXX, XXX, XXX−XXX

Supporting Information for Publication

Title of Paper

Comparisons of Protein and Peptide Complexity in Poneroid and Formicoid Ant Venoms

Authors

Samira R. Aili[†], Axel Touchard[‡], Jennifer M. S. Koh[†], Alain Dejean[‡,§], Jérôme Orivel[‡], Matthew P. Padula[¶], Pierre Escoubas[#], Graham M. Nicholson[†,*]

Affiliations

[†]Neurotoxin Research Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW 2007, Australia

[‡]CNRS, UMR Ecologie des Forêts de Guyane (EcoFoG), Campus Agronomique, BP 316, 97379 Kourou Cedex, France

[§]Laboratoire Écologie Fonctionnelle et Environnement, CNRS, UMR 5245, Ecolab, 118 route de Narbonne, 31062 Toulouse, France

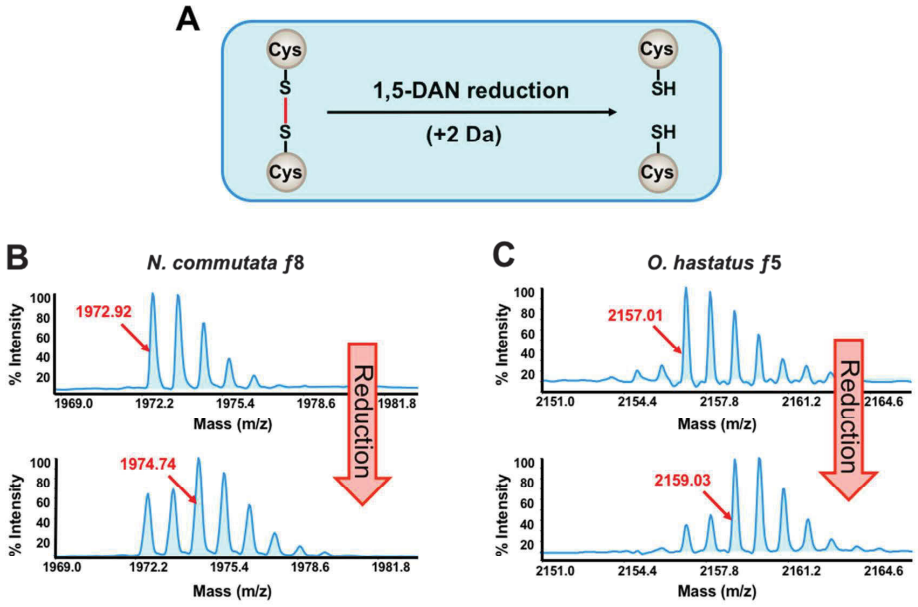
[¶]Proteomics Core Facility, Faculty of Science, University of Technology Sydney, NSW 2007, Australia

[#]VenomeTech, 473 Route des Dolines — Villa 3, Valbonne 06560, France

Corresponding Author

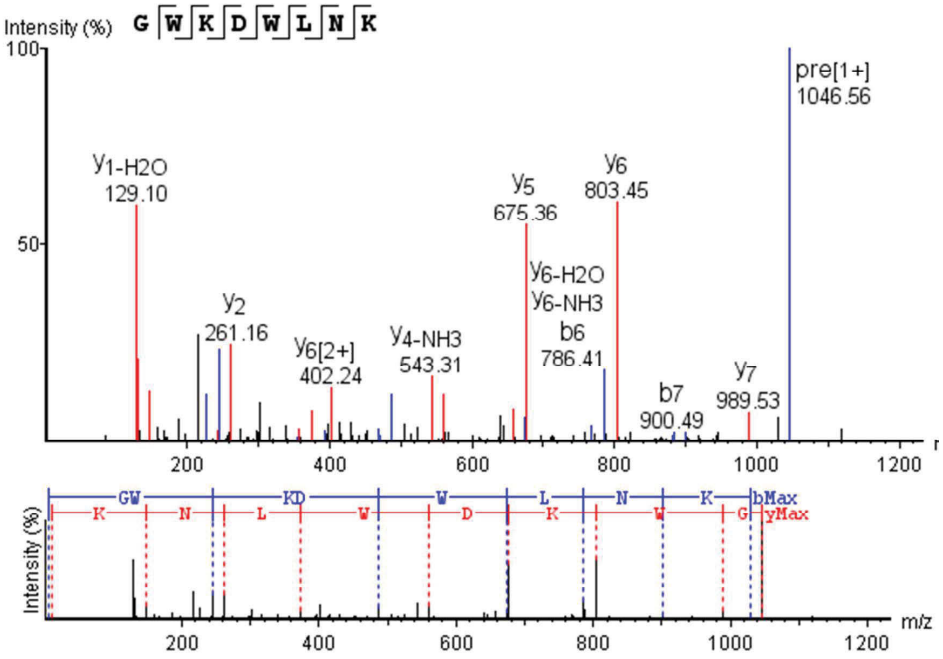
Prof. Graham M. Nicholson, Neurotoxin Research Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW 2067, Australia. Phone: (+61) 9514 2230. Fax: (+61) 9514 1656. E-mail: Graham.Nicholson@uts.edu.au

Table of Contents



Supplementary Figure S1. Identification of disulfide-bonded peptides with the use of 1,5-DAN matrix. (A) A 2 Da shift in mass results from the gain of two hydrogen atoms following reduction of a cystine bond. (B-C) Typical MALDI-TOF spectra recorded in reflector mode showing the 2 Da increase in mass following use of the reducing matrix 1,5-DAN consistent with the presence of a single disulfide bond. Upper panels show native peptide isotopic clusters obtained using non-reducing CHCA matrix, while lower panels show the same peptide in reduced form obtained in reductive 1,5-DAN matrix. (B) Peptide from *N. commutata* C18 RP-HPLC *f*8, (C) peptide from *O. hastatus* C18 RP-HPLC *f*5.

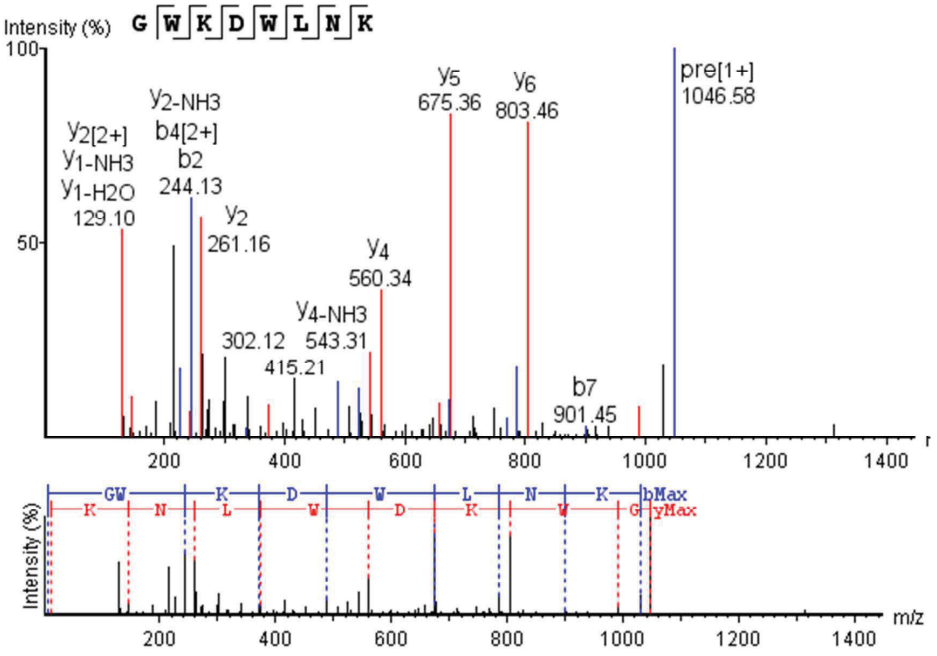
Spot 2:



Precursor *m/z*: 523.7957
Charge state: 2
Confidence score/expect value: 26.66

Supplementary Figure S2. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 2 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 2, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

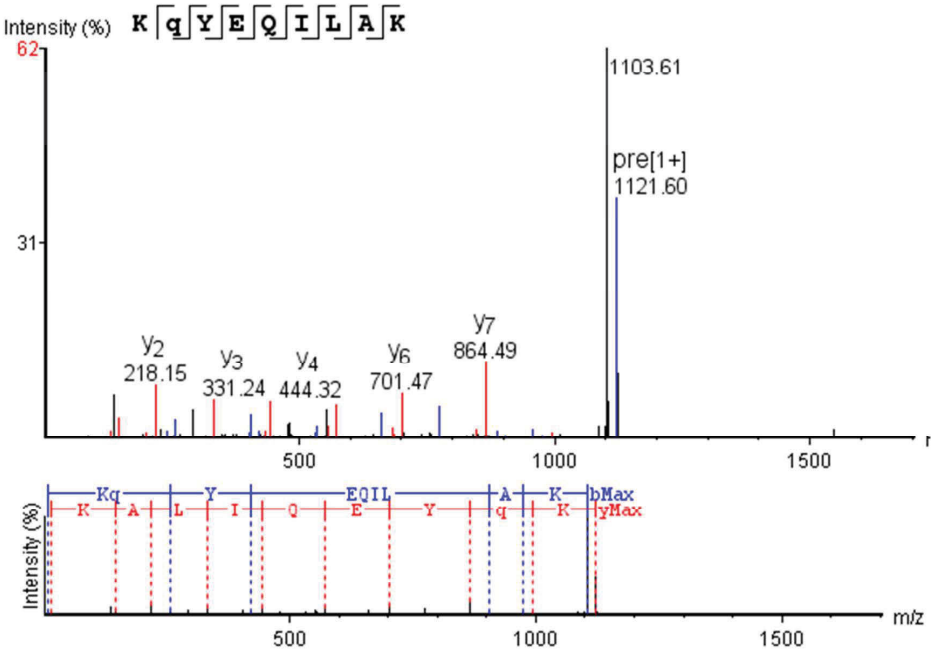
Spot 5:



Precursor m/z: 523.7988
Charge state: 2
Confidence score/expect value: 25.75

Supplementary Figure S3. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 5 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 5, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

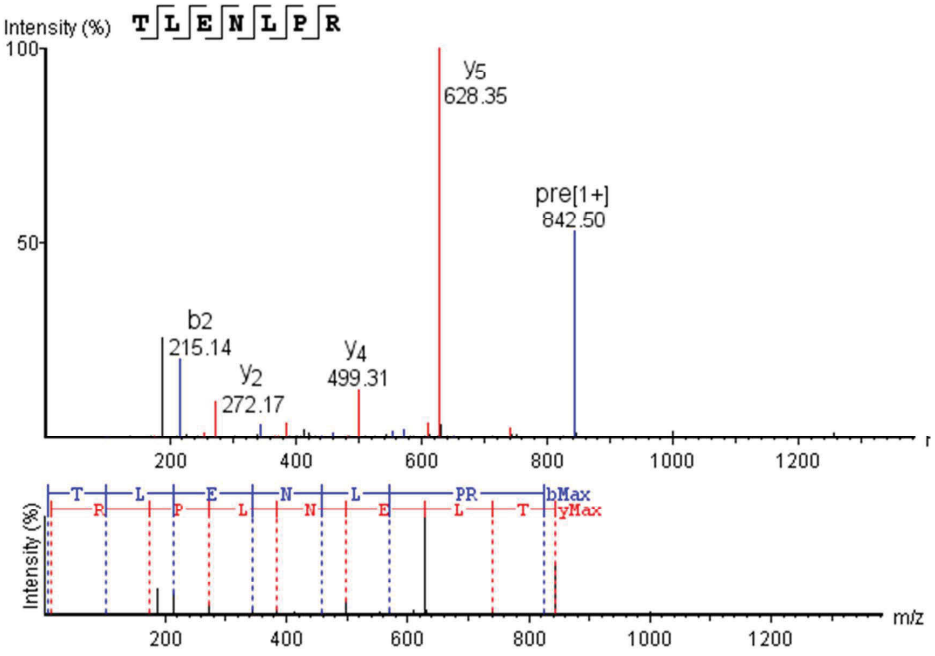
Spot 8:



Precursor m/z: 561.3240
Charge state: 2
Confidence score/expect value: 27.45

Supplementary Figure S4. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 8 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 8, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

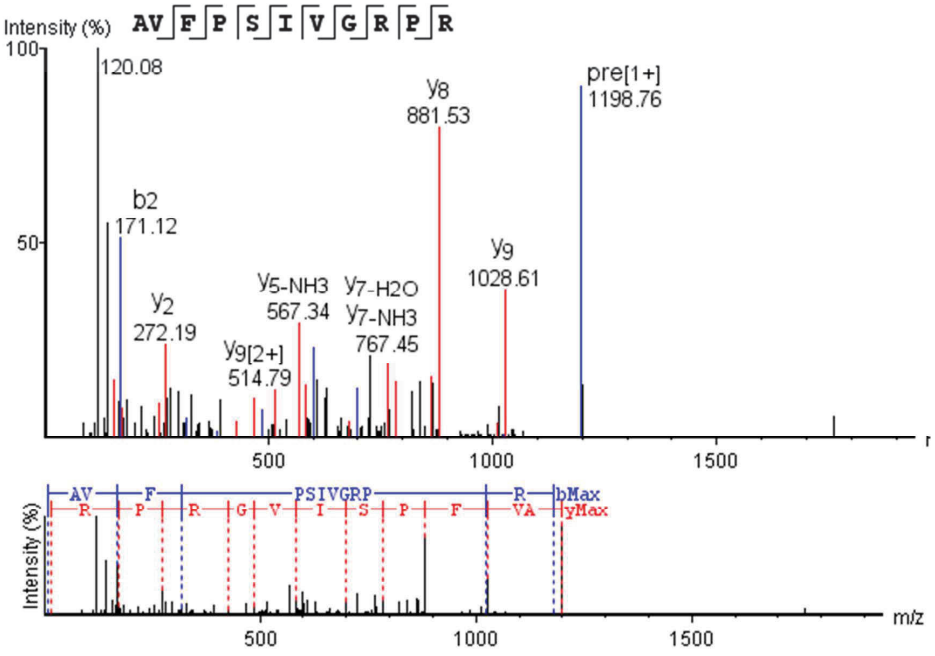
Spot 10:



Precursor m/z: 421.7591
Charge state: 2
Confidence score/expect value: 23.20

Supplementary Figure S5. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 10 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 10, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

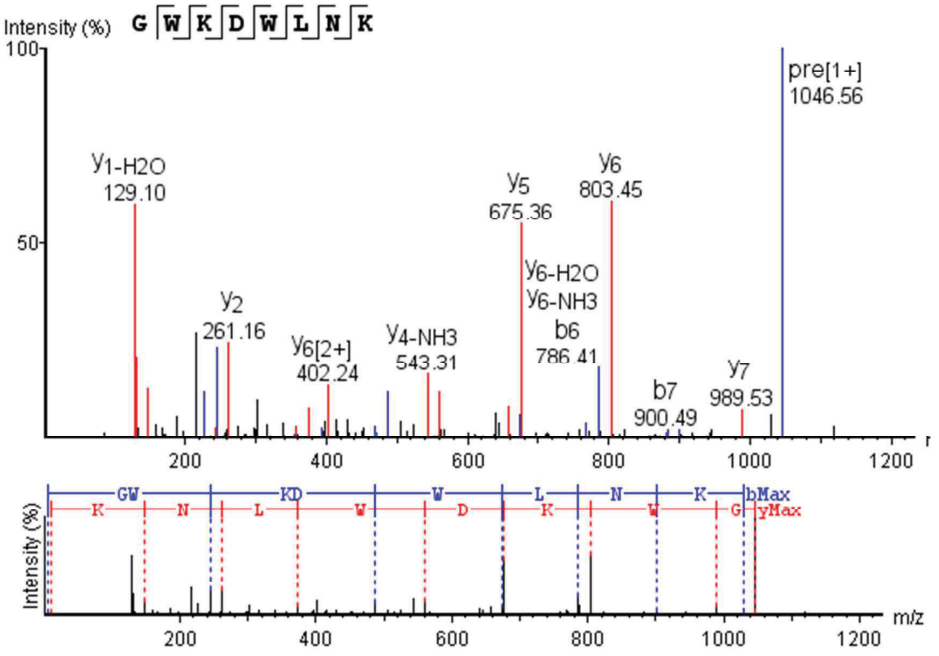
Spot 15



Precursor m/z: 599.8855
Charge state: 2
Confidence score/expect value: 43.87

Supplementary Figure S6. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 15 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 15, as well as the peptide's precursor *m/z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

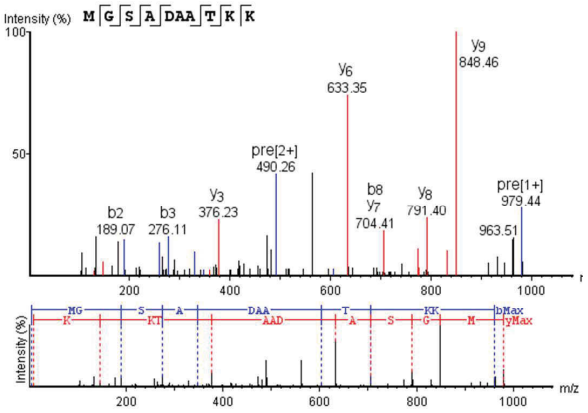
Spot 17



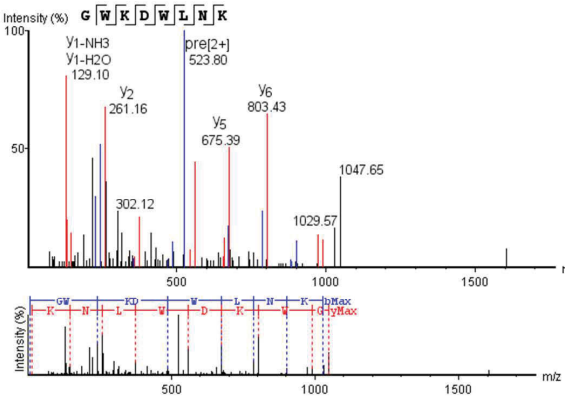
Precursor m/z: 523.7957
Charge state: 2
Confidence score/expect value: 26.66

Supplementary Figure S7. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 17 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 17, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

Spot 18:



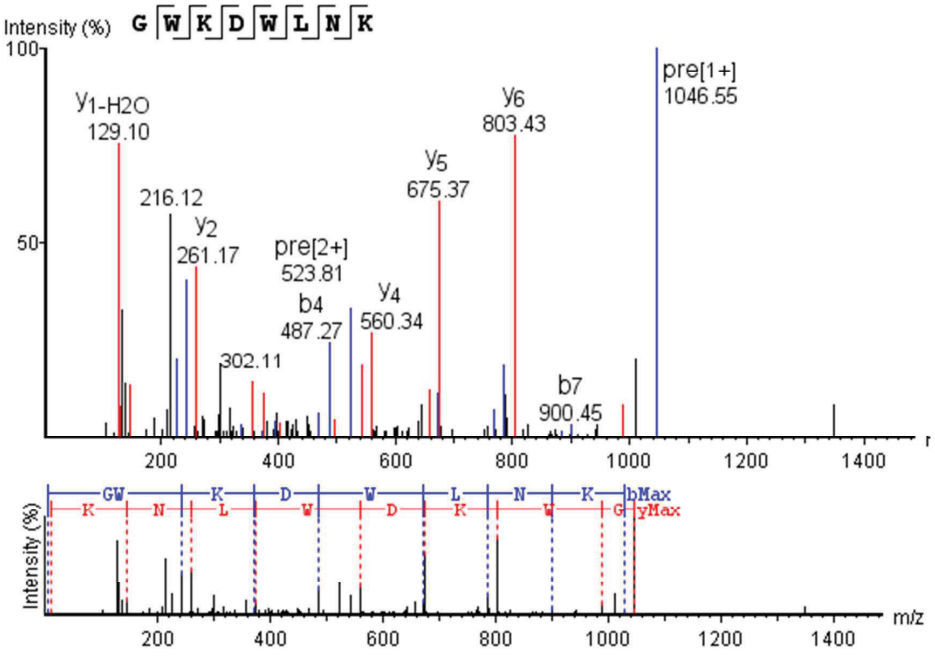
Precursor m/z: 490.2568
Charge state: 2
Confidence score/expect value: 21.29



Precursor m/z: 523.7982
Charge state: 2
Confidence score/expect value: 29.27

Supplementary Figure S8. Annotated MS spectra of the two peptides from *N. commutata* 2D gel spot 18 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of peptide matches to spot 18, as well as each peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

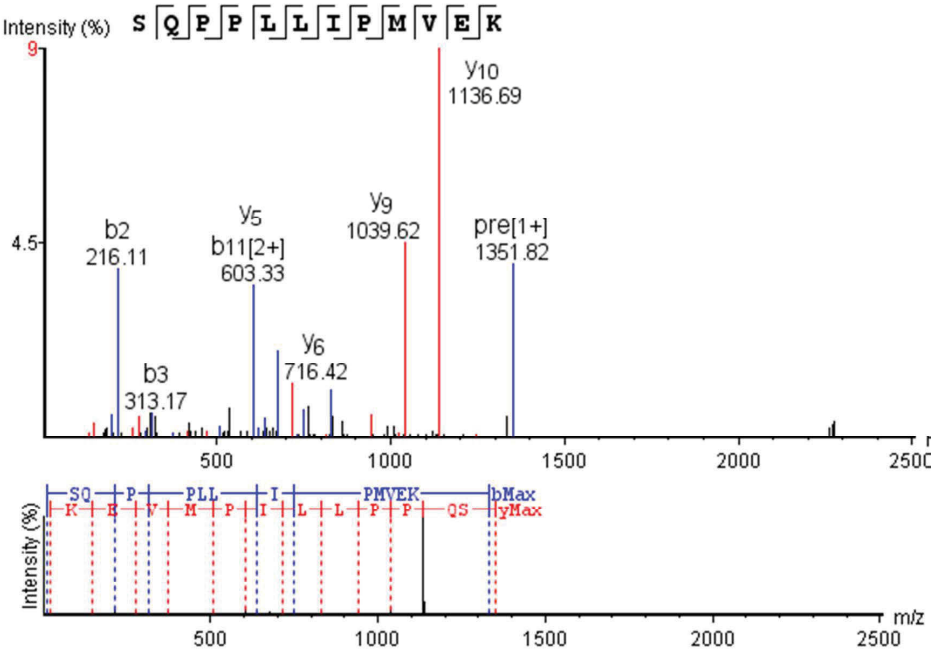
Spot 19



Precursor m/z: 523.792
Charge state: 2
Confidence score/expect value: 32.88

Supplementary Figure S9. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 19 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 19, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

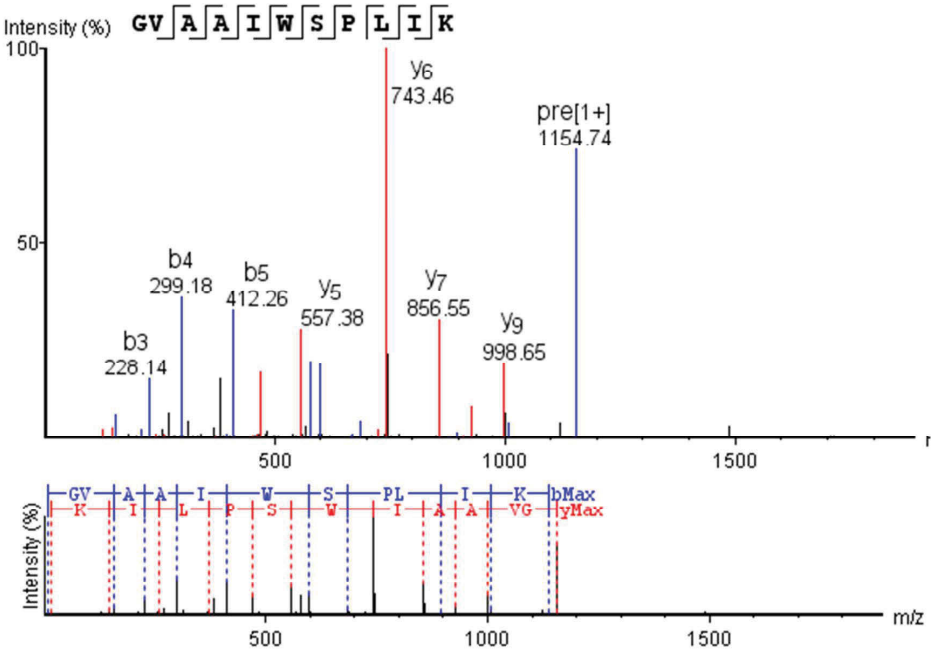
Spot 23:



Precursor m/z: 676.4149
Charge state: 2
Confidence score/expect value: 41.84

Supplementary Figure S10. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 23 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 23, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

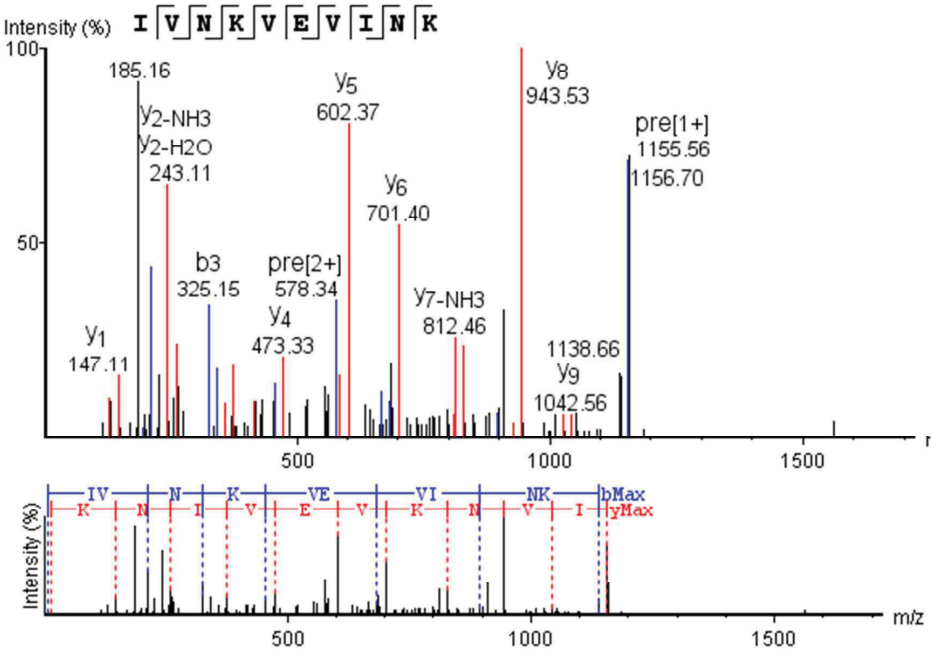
Spot 24:



Precursor m/z: 577.8694
Charge state: 2
Confidence score/expect value: 48.61s

Supplementary Figure S11. Annotated MS spectra of a peptide from *N. commutata* 2D gel spot 24 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 24, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

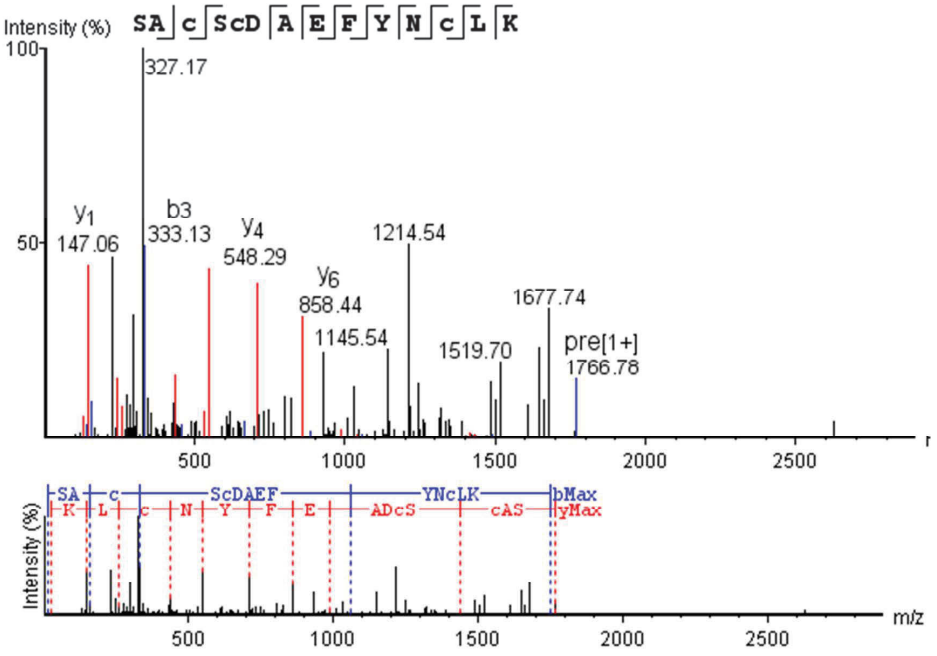
Spot 1



Precursor m/z: 578.3434
Charge state: 2
Confidence score/expect value: 24.44

Supplementary Figure S12. Annotated MS spectra of a peptide from *E. tuberculatum* 2D gel spot 1 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 1, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

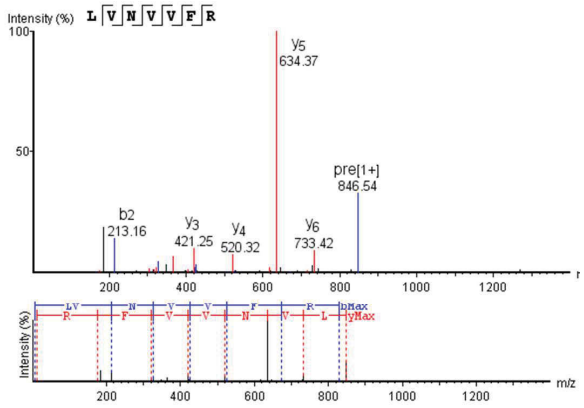
Spot 2:



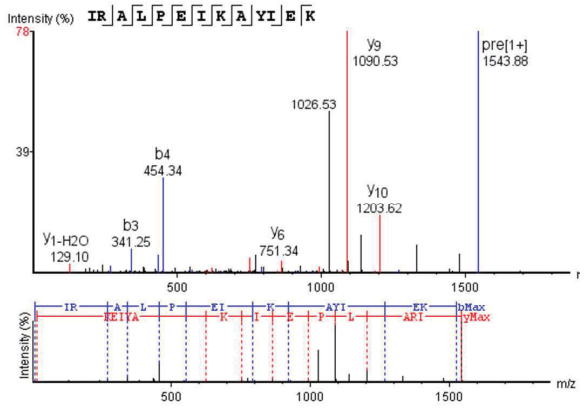
Precursor m/z: 883.9064
Charge state: 2
Confidence score/expect value: 26.03

Supplementary Figure S13. Annotated MS spectra of a peptide from *E. tuberculatum* 2D gel spot 2 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 2, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

Spot 13:



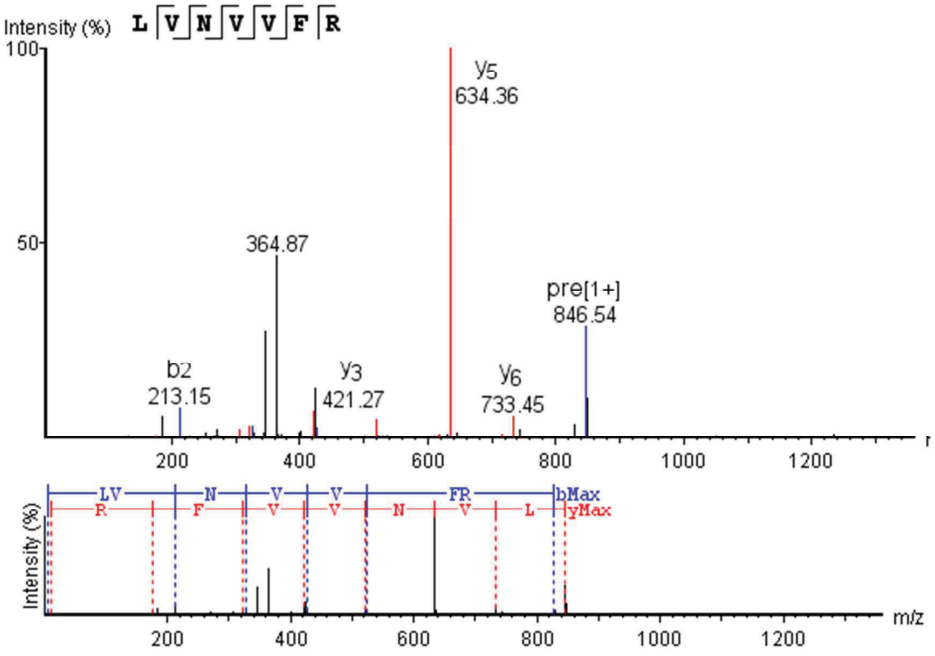
Precursor m/z: 423.77
Charge state: 2
Confidence score/expect value: 26.30



Precursor m/z: 772.43
Charge state: 2
Confidence score/expect value: 22.08

Supplementary Figure S14. Annotated MS spectra of the two peptides from *E. tuberculatum* 2D gel spot 13 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 13, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

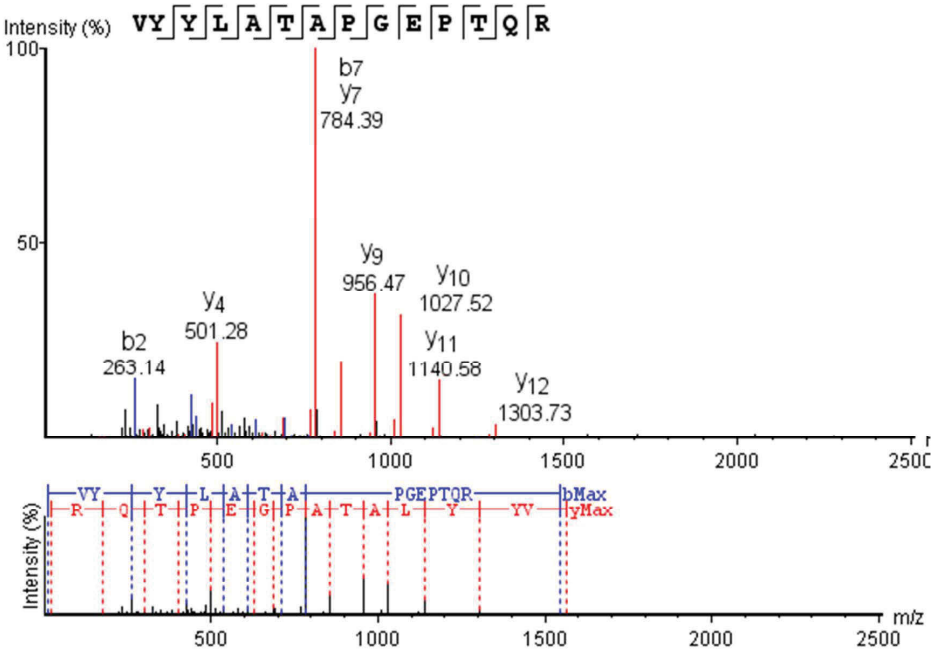
Spot 15:



Precursor m/z: 423.77
Charge state: 2
Confidence score/expect value: 33.22

Supplementary Figure S15. Annotated MS spectra of a peptide from *E. tuberculatum* 2D gel spot 15 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 15, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

Spot 19



Precursor m/z: 783.4087
Charge state: 2
Confidence score/expect value: 56.61

Supplementary Figure S16. Annotated MS spectra of a peptide from *E. tuberculatum* 2D gel spot 19 used to match against Hymenopteran proteins. The panel shows the annotated spectra and b- and y-ion alignments of the peptide match from spot 19, as well as the peptide's precursor *m*/*z*, charge state and confidence score obtained from Peaks 6 software (BSI, Waterloo, ON, Canada).

Chapter Seven

A holistic investigation of *Paraponera clavata* venom gland transcriptome and venom proteome

Chapter Seven: A Holistic Investigation of *Paraponera clavata V*enom Gland Transcriptome and Venom Proteome

Compound Abstract

A constant hurdle in ant venom proteomic investigations is the lack of a database to identify the full sequence and function of the proteins and peptides within the venom. This was exemplified by the large number of 2D-PAGE spots from the previous chapters with no matches to known proteins. In order to resolve this issue for *P. clavata* venom, a transcriptomic investigation was performed as it is able to rapidly generate sequence information that can then be used to assign the tandem MS fragmentation spectra of a peptide to a specific transcript. This was performed alongside a shotgun LC-MS/MS analysis of the venom to confirm that transcripts were being further expressed as proteins.

The combined transcriptomic and proteomic investigation of *P. clavata* was able to identify four times the number of proteins initially identified using the 2D-PAGE approach alone in the previous chapter. It was also able to identify several novel peptide sequences for future pharmacological investigations, including some with a cysteine framework complying to that of conotoxin framework I or framework VI/VII which is known to have an inhibitor cysteine knot framework. These types of peptides have previously been developed as successful pharmaceutical or bioinsecticidal peptides or are currently in clinical trials.

7.1. Introduction

A significant proportion of pharmaceutical drugs are derived from natural products. Animal venoms are becoming an increasingly popular source of drug leads as they have been shown to be rich in bioactive peptides. These peptides have evolved over millions of years of evolution and target a wide variety of enzymes, ion channels, receptors and other biological processes [59, 60]. Currently, there are six venom-derived peptides on the therapeutic drug market and one insecticide [60]. However, there is potential for many more, as only a small fraction of venom peptides have been investigated. New developments in sensitivity and accuracy in mass spectrometry and transcriptomics sequencing will accelerate this process. One of the hurdles of venom proteomics investigations has been the lack of available databases to assign peptide sequence information using bottom up mass spectrometry. However, the one of the biggest problems with proteomics is the difficulty in producing sequence data, as it needs a large amount of starting material in order to purify the protein or peptide of interest to homogeneity. This is then followed by a need to *de novo* sequence the MS fragmentation spectra of a peptide that might not have fragmented properly or ionised sufficiently to be detected by the mass spectrometer and accurately sequenced. This task is further complicated when the sequence of interest is more than 5 kDa, as it would need enzymatic digestion to produce smaller peptides which are more likely to be detected by the mass spectrometer. In addition, smaller peptides tend to fragment with better efficiency, providing richer spectra for sequence determination. Therefore, the use of mRNA to deduce protein sequence information is becoming increasingly common, particularly in conjunction with proteomics studies because it overcomes some of the problems of predicting ORFs and *de novo* sequencing.

Transcriptomics is particularly useful because venom gland mRNA profiles can be generated *de novo,* using only a small amount of tissue in a time efficient manner. Furthermore, a transcriptomics approach is relatively unbiased as it captures almost all the diversity present in the venom-gland at the time of tissue harvesting, as opposed to proteomics where ion-suppression can significantly confound analyte ionisation and detection when using certain types of instrumentation. However, there are disadvantages as well such as the fact that post-translational modifications (PTMs) such as proteolytic cleavage, amidation of the C-terminus, pyroglutamate formation, etc., are not revealed by transcriptomics, whereas they are seen in proteomics. This is where a holistic study, combining proteomics and transcriptomics, is advantageous as it allows for identification of synthesised proteins and PTMs, and confirmation of the N-terminus of the peptide/protein. On the other hand, the transcriptome will allow quantification of the expression of genes of interest. This combination is a powerful

tool when used with software packages such as PEAKS Studio v8.5 (BSI, Canada) that can assign peptide sequences to a cDNA sequence and hence a gene.

Despite these technological advances, ant venom studies have been slow in adopting transcriptomics approaches. To date, there are only three published ant venom gland transcriptomes, of which only one (*Odontomachus monticola*) is compared to the peptidome [61-63]. This is despite the fact that ants are a diverse group, with ca. 16,000 extant species (www.AntWeb.org). They are also one of the most successful organisms on this planet, occupying a diverse range of habitats, often being the species that dominate [64]. The main reason for this shortage comes down to the size of the small size of ants which makes venom and venom gland collection difficult, especially when compared to other venomous organisms such as snakes and spiders. Another reason may be the lack of available genomes (only 20 ant genomes to date)[65] which means *de novo* transcriptome assemblies are required [61]. The reason this is a problem is the fact that bioinformatics tools that assemble transcriptomes give better support when assemblies are generated using a reference genome.

Approximately 9,000 of the 16,000 ant species of ants have evolved a venom that is rich in peptides and proteins, which they use for predation, defence and communication [17]. Despite this rich source of bioactive molecules, there are few studies investigating ant venom composition. This is in part due to the initial misconception that ant venoms were mainly composed of formic acid, and the low yield of venom from these small hymenopterans [61]. The early studies of Cavill (1964) showed the presence of proteinaceous components with hyaluronidase activity in the bull ant, *Myrmecia gulosa* [66], however, the first ant venom peptide toxin, poneratoxin from the ant *Paraponera clavata*, was not isolated until 1991 [18, 22, 67, 68].

Paraponera clavata is one ant species that has received significant attention at a proteomic level, with seven publications to date – one of which was the work presented in chapter 4 [20-22, 67-71]. Moreover, all of these previously published studies focus on poneratoxin, a 25-residue peptide, from the most active fraction of the venom that elicits neurotoxic activity by blocking synaptic transmission in cockroaches and also affects vertebrate voltage-activated sodium channels ($Na_v1.7$) [67]. Nevertheless, as seen in chapter 4, there are several other components in this venom that remain uncharacterised as shown by the high number of unmatched peptides.

The overarching aim of the present study was to identify a novel peptide(s) with neurotoxic activity for use as a bioinsecticide in the future. This is an area of research with great

importance, as there is a high demand for novel ways to eliminate insect pests. The harmful side-effects associated with synthetic agrochemical insecticides have led to the de-registration of many insecticides, particularly after the introduction of new food safety regulations in the European Union [9]. Finding novel ways of combating these insect pests is critical, given that provision of food is a major problem facing mankind. Indeed, insect pests destroy almost 20% of crops or stored grains, particularly in sub-Saharan Africa where losses are much higher [1]. Therefore, insect-selective bioinsecticidal peptides are being investigated due to their lower risk to human and animal health, as well as the environment [1, 9]. Venom peptides are a particularly attractive source of these bioinsecticides due to their small size and relative ease of synthesis. Their novel modes of action also mean that no known resistance mechanisms have yet been developed by the insects. Therefore, this investigation aimed to identify novel sequences in *P. clavata* with potential bioactive properties for future pharmacological investigations into the insecticidal activity. Furthermore, this study aimed to highlight the advantages of using a combined proteomics and transcriptomics approach to provide a holistic overview of the complexity of ant venom arsenal using *P. clavata* as a representative of the undiscovered ant venom components.

7.2. Methods

7.2.1. RNA isolation and Illumina sequencing

Paraponera clavata worker ants were collected in the locality of "la Montagne de Kaw" in French Guiana. Both venom glands and venom sacs of 52 worker ants were dissected in distilled water and immediately placed into 1 mL of RNAlater [72]. The samples were stored at –80 °C prior to RNA extraction. The RNAlater was removed and the glands disrupted with a TissueLyser II (Qiagen, USA) in RLT buffer containing 10% (v/v) of 2-mercaptoethanol (RNeasy Mini Kit, Qiagen, USA). RNA was first isolated with a phenol-chloroform (5:1) solution followed by washing with a solution of chloroform-isoamyl alcohol (25:1) to remove any traces of phenol. The RNA was then bound to a Qiagen column and washed as per the manufacturer's instructions. DNAse I (Roche Diagnostic, Germany) was added in order to remove any remaining fragments of DNA. The RNA was eluted in sterile water and total RNA was determined using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, USA) with the RNA HS assay kit (Life Technologies Corp., USA). 260/280 and 260/230 nm ratios were determined using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific, USA). Finally, RNAstable (Biomatrica, USA) was added to the purified RNA and the sample dried using a Speed Vac (RC1010, Jouan, Saint Herblain, France.).

The purified RNA (1.8 ng) was sent to the Department of Biological Sciences, National University of Singapore (NUS) for RNA-seq library construction and RNAseq. Sample quality was reassessed using Agilent 2100 Bioanalyzer and cDNA library was constructed using the NEBNext Ultra Directional Library Prep Kit as per manufacturer's protocol. The 250bp paired-end library was sequenced on an Illumina HiSeq2500 sequencer on 1/14th of a lane per sample.

A summary of the combined proteomic and transcriptomic investigation of the bullet ant *P. clavata* is shown in Figure 7.1

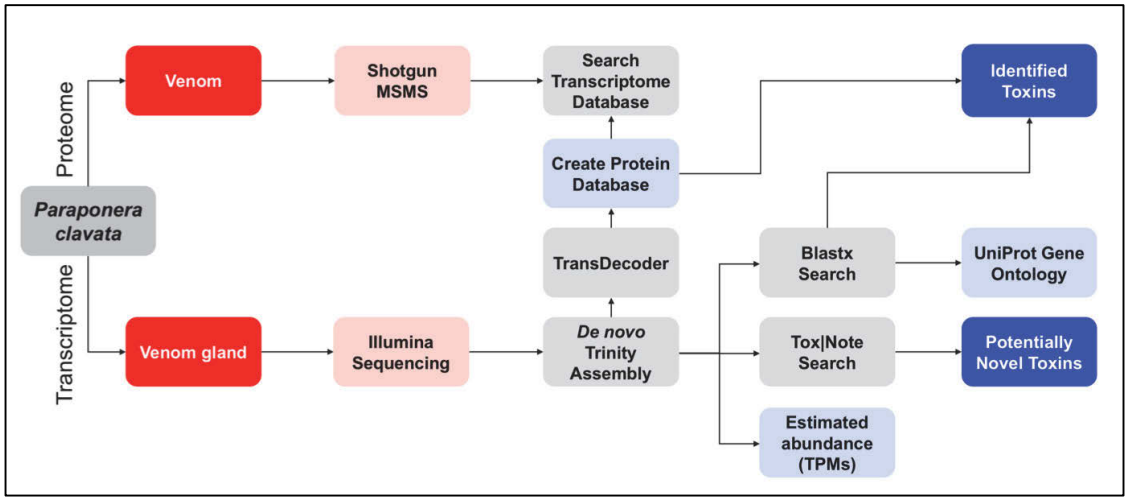


Figure 7.1: Summary of the *P. clavata* combined proteome/ transcriptome methodology.
Flow chart outlines the main steps undertaken in the present investigation using an integrated proteomic and transcriptomic methodology.

7.2.2. Quality control and de-novo assembly

After de-multiplexing, a filtering step was performed on the raw Fastq files for quality control using FastQC version 0.115 [73], and fastq-screen (version 0.9.3) (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/) with various pre-indexed genomes included (see supplementary file 2 for graphs). Primer and adaptor sequences and low quality bases were trimmed using Trimmomatic (version 0.36) with default settings [74]. The resulting paired-end reads underwent *de-novo* assembly using Trinity (version 2.2.0) with default parameters that included *in-silico* normalisation [75], generating 54,242 contigs. Various assembly statistics and further quality checks were performed using in-house scripts. Bowtie was used to assess the assembly quality by mapping the reads to the contigs [76]. Expression levels were computed in Transcripts Per Kilobase Million (herein TPM) using the RSEM package [77]. Commands used for Trinity, RSEM and bowtie are included as supplementary material.

7.2.3. Contig functional annotation

BLASTx

In order to annotate the *P. clavata* transcriptome, all assembled contigs were queried against the NCBI non-redundant online database (National Center for Biotechnology Information, August 2017) using a BLASTx algorithm (see supplementary file 1 for utilised script). All sequences with hits to the database with an e-value below e^{-4} were considered for further analysis. Translated protein sequences were assigned Gene Ontology (GO) terms using UniProtKB [78, 79]. Toxins were identified by a manual search of ca. 50 keywords, e.g. toxin, phospholipase, metalloproteinase, acid phosphatase and dipeptidyl peptidase (See Table 7.1 for all terms). The top 20 expressed toxins, based on TPM value, were annotated based on their matched proteins and re-named according to the nomenclature system below.

Table 7.1: Toxin keyword search list.
Table shows a list of common toxin protein names previously used to describe venom toxins

Common Toxin Protein Names				
Acetylcholinesterase	Bicarinalin	Ectatomin	Mastoparan	Ponericin
ACh	Calsyntenin	Esterase	Metalloendopeptidase	Serine prot
Acid Oxidase	Carboxypeptidase	Haemorrhagic	Metalloprotease/ peptidase	Serine Protease/ Proteinase
Acid phosphatase	Cholinesterase	Hyaluronidase	Metalloproteinase	Serpin (serine protease inhibitor)
ADAM	Coagulation	Icarapin	Neprilysin	SNTX (Stonustoxin)
Alkaline phosphatase	Conotoxin	Kinin	Phospholipase	Sphingomyelin D
Allergen	CRISP/ cysteine rich secretory protein	LAAO amino oxidase	Pilosulin	STX (Shiga Toxin)
Analgesic	Defensin	Lectin	PLA	Toxin/ Tx
Antigen	Dipeptidyl peptidase	Leucine-rich domain	PLB	VEGF, vascular
Arginine kinase	Disintegrin	Lipase	PLD	Venom

TransDecoder

The assembled and annotated venom gland transcriptome was translated using TransDecoder (version 3.0.1) [75, 80] using the parameters for 6-frame translation with a minimum open reading frame (ORF) of 50 amino acids. The resulting ORFs were uploaded onto PEAKS 8.5 as a database (Bioinformatics Solutions, ON, Canada) [81]. This transcriptome database was used to query the LC-MS/MS data obtained from the shotgun mass spectrometry experiment (see section 2.4). The peptides from LC-MS/MS which matched to contigs from the translated transcriptome were also searched on BLASTp to confirm the BLASTx match assigned to the contig [82].

Signal peptide prediction

Venom toxins are usually transcribed as pre-propeptide precursors that contain a signal peptide, potentially a propeptide region, and a toxin segment. These transcripts undergo processing, where signal peptidases and pro-protein convertases cleave the signal and propeptides in order to produce the mature toxin. The signal peptide cleavage site was predicted using the SignalP server (http://www.cbs.dtu.dk/services/SignalP/)[83]. Propeptides could not be predicted, as there is no available predictor for ant venom peptides and hence were predicted if proteomic data was available for that transcript.

Tox|Note

To further explore and identify novel toxin-like peptides, the assembled transcriptome fasta file was uploaded onto the Arachnoserver's Tox|Note pipeline [84] since a similar pipeline for ants, or hymenopterans in general, is unavailable. The pipeline annotates transcripts by homology or *de novo* using BLASTx or the Tox_Seek| tool respectively. The Tox|Note output also predicts the transcript cleavage sites for the signal peptide and a propeptide using a combination of the SignalP and Spider|ProHMM tools respectively [84]. A toxin was considered novel if it had no BLASTx or Tox|Note match.

Toxin identification

Toxin matches from BLASTx and Tox|Note were compiled and sorted based on their respective expression TPMs. Some of these proteins were re-searched on the web NCBI BLASTx interface to find other homologous proteins for alignment and domain annotation. A ClustalW multiple alignment was then performed using Mega7 [85].

Nomenclature

Throughout this chapter, peptide toxins (< 10 kDa) have been named according to the proposed rational nomenclature system for peptide toxins from animal venoms of King *et al.,* (2008) [86]. This same nomenclature was used to re-name known ant venom toxins in Touchard *et al.* (2016) [87]. A brief summary of the naming system can be seen in Figure 7.2.

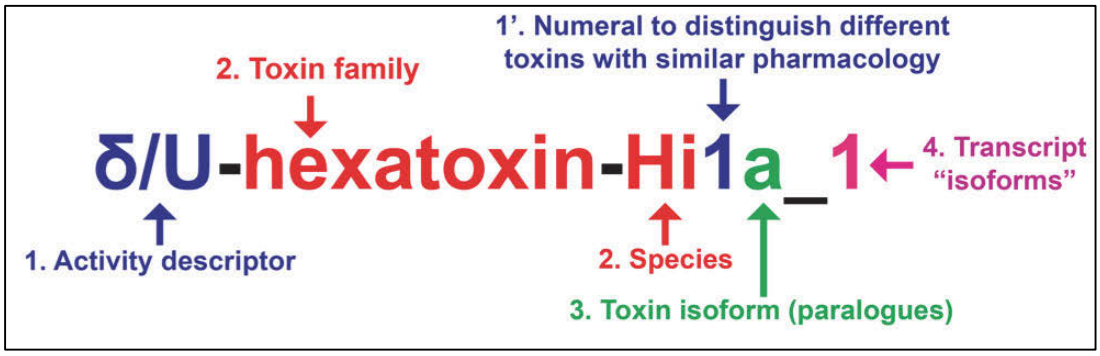


Figure 7.2: Peptide toxin nomenclature system using a spider venom peptide example.
The toxin name is divided into three parts that describe the toxin's activity (blue), biological source (red), and relationship to other toxins (green/purple). The subtype descriptor should be based on IUPHAR-recommended nomenclature for channels and receptors. Figure adapted from King et al. (2008) [86].

However, as this rational nomenclature system did not consider proteins, therefore, we adapted it to include putative protein toxins. The proposed protein toxin nomenclature method begins with a generic protein name (if known), which is based on a previously assigned name based on homology to other proteins of the same sequence and activity, for example hyaluronidase or phospholipase. If the protein has not been characterised, then it would begin with 'Uncharacterised'. This generic toxin name is followed by a number that designates a particular family of paralogous toxins with the same activity assigned to the generic protein name. This designator was introduced because often there is more than one group of proteins from the same species that act on the same molecular target. This designator is simply incremented as new groups of toxins are discovered. The toxin-family designator is followed by a lowercase letter that is used to distinguish isoforms as many venomous animals diversify their toxin repertoire through post-translational modifications that change the original toxin by a few amino acids. However, if the isoform from a transcript is identical at the protein level with a different nucleotide sequence a numeral is added after the isoform letter preceded by an underscore '_'. The last part of the protein name is an uppercase letter that identifies the genus of origin and the lowercase species name. Both genus and full species name are

required to avoid confusion because of the large number of ant species. An example of a protein isolated from this investigation is shown in Figure 7.3.

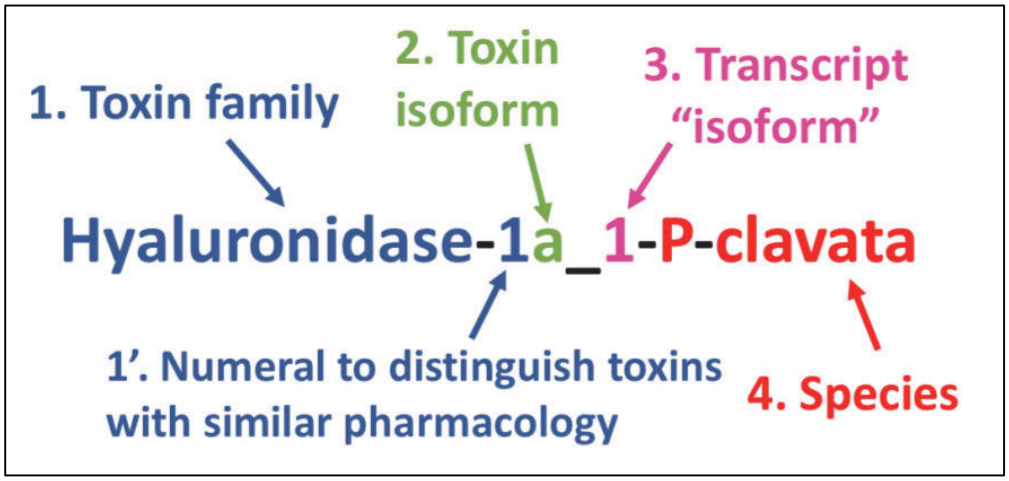


Figure 7.3: Proposed protein toxin nomenclature.
The toxin name is divided into The toxin name is divided into three parts that describe the toxin's family (blue), biological source (red), and relationship to other toxins (green/purple).

Identification of potential pharmacological targets

In order to assess the extent of homology of contigs to the matched proteins and to assess whether the critical residues were present, several potential toxin contigs were manually searched using BLASTx. These were then aligned using MEGA's ClustalW alignment feature [85].

7.2.4. Venom collection for proteomics

After collection, ants were stored at −20 °C prior to manual dissection of the venom glands. After dissection, the glands were pooled in 10% (v/v) acetonitrile (ACN)/water. Samples were then centrifuged for 5 min at 14,400 rpm (12,000 g_{av}), the supernatant was collected and lyophilised prior to storage at −20 °C.

Shotgun proteomics

To identify proteins present in the venom gland, we used a bottom-up proteomics approach. Crude venom was re-suspended in MilliQ water and digested using trypsin (Promega, Madison, USA). Peptides were then analysed using an Eksigent 415 autosampler connected to a 415 nanoLC system (Eksigent, USA), 5 µL of the sample was loaded at 300nl/min with MS buffer A (2% Acetonitrile + 0.2% Formic Acid) by direct injection onto a PicoFrit column (75 µm x 150

mM; New Objective, USA) packed with C18AQ resin (1.9 μM, 200 Å, Dr Maisch, Germany). Peptides were eluted from the column and into the source of a 6600 TripleTOF hybrid quadrupole-time-of-flight mass spectrometer (Sciex, USA) using the following program: 2–35% MS buffer B (80% Acetonitrile + 0.2% Formic Acid) over 90 min, 35-95% MS buffer B over 9 min, 95% MS buffer B for 9 min, 80–2% for 2 min. The eluting peptides were ionised at 2300 V. An Intelligent Data Acquisition (IDA) experiment was performed, with a mass range of 350-1500 Da continuously scanned for peptides of charge state 2^+–5^+ with an intensity of more than 400 counts/s. Up to 50 candidate peptide ions per cycle were selected and fragmented and the product ion fragment masses measured over a mass range of 100–2000 Da. The mass of the precursor peptide was then excluded for 15 seconds.

Protein identification

Proteins present in the venom proteome were identified by mapping the mass spectra (section 7.2.5) to the translated transcriptome assembly (section 7.2.4) combined with a contaminants database using the software PEAKS Studio v8.5 (BSI, Canada). Because of the known variability of toxins post-translationally, the following parameters were employed in order to maximise the identified proteins: semi-tryptic peptide, biological modification: deamidation and oxidation, parent mass error tolerance: 50 ppm, fragment mass error tolerance: 0.1 Da, Enzyme: trypsin, max mixed cleavages: 3. The results of the search were then filtered to include peptides with a –log10P score that was determined by the False Discovery Rate (FDR) of <2%, the score being that where decoy database search matches were <2% of the total matches to call a positive hit. Sequences with less than 95% confidence were excluded.

Transcriptome data upload

All annotated transcripts and sequence data will be deposited into the EMBL-EBI European Nucleotide Archive (ENA).

7.3. Results

7.3.1. Paraponera clavata venom gland transcriptome profile

In order to investigate the diversity and expression of venom toxins and toxin-like sequences in *P. clavata*, transcriptome sequencing was performed on the venom glands and sacs. The transcriptome assembly also provided a database for exploring the proteome of *P. clavata* venom.

7.3.1.1 Illumina sequencing

Illumina Hiseq 2500 sequencing of the cDNA library from *Paraponera clavata* venom glands yielded over 17 million reads with an average length of 547 bp after quality control (QC). In brief, *de novo* transcriptome assembly from Trinity yielded 54,242 contigs, with an N50 of 1667 bp and a maximum contig size of 23,534 (Table 7.2). Trinity was also used to determine expression levels of contigs using the transcripts per million (TPM) estimates. Full assembly statistics are provided in Table 7.2.

Table 7.2: Assembly statistics and downstream metrics from the *P. clavata* venom gland transcriptome analysis.

After QC		
	Total number of reads (paired-end)	17,026,962
Assembly summary		
	Total number of genes/ transcripts	46,282/ 54,242
	Contigs (N50)	1,667
	Contig length (Median/Average)	594/1,007
	Longest contig length	23,534
	GC content	37.3%
	Raw reads mapped to contigs	99.3%
Annotation		
	Total BLASTx hits	37,140
	Toxins	356
	Percent of assembled contigs with BLASTx hit	59%
	Contigs with Gene Ontology Annotation	6,732
Translation		
	At least one ORF	31,586

7.3.1.2 Annotation and classification of P. clavata protein functions

To identify the types of proteins and toxins in the venom gland similar to known proteins, the contig library was used to interrogate the NCBInr database using the BLASTx algorithm [82] with an e-value cut-off of e^{-4}. The search generated 37,140 matching protein hits (68% of the transcriptome). Twenty of the top 35 hits were to ants, and the remainder were to other hymenopterans (bees, wasps), with the highest matched species being to the ant *Harpegnathos saltator* (Figure 7.4), reflecting the close phylogenetic relationship between these species.

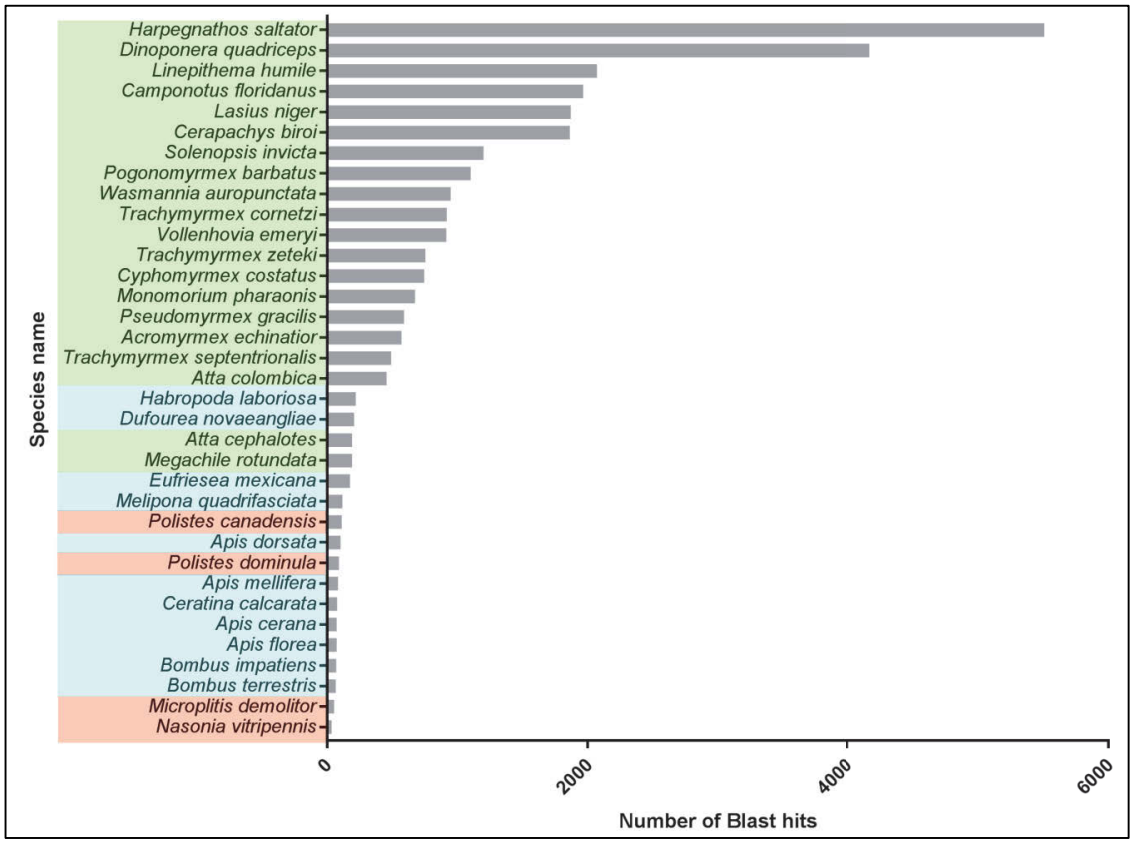


Figure 7.4: Distribution of protein hits to different hymenopteran species.
Graph shows the number of protein hits in the top 35 species matched by BLASTx, highlighted regions indicate matches to ants (green), bees (cyan) and wasps (pink). The most prevalent protein/ peptide hits were to the ant *Harpegnathos saltator.*

It was found that the majority of the contigs were not related to ant venom proteins or toxins. This was expected as ant-specific venom protein datasets are still limited in the NCBInr. Moreover, most of the transcripts from venom gland mRNA encode for proteins from the venom gland cells such as actin and myosin or are part of the machinery which helps in the processing of toxins in the venom-gland such as chaperones and signal peptidases. For functional analysis, a Gene Ontology (GO) search was conducted on UniProtKB's Retrieve/ID mapping function using the accession numbers obtained from the BLASTx results. A total of 6732 contigs (18%) were assigned GO terms belonging to the categories molecular function, biological process, and cellular components. Several of these contigs had multiple GO terms, therefore terms were initially sorted based on category. The majority of the proteins were identified to have molecular functions followed by those that were cellular components and those involved in biological processes (Figure 7.5). Most of the proteins were identified as having a binding or catalytic function or were membrane proteins. Of the proteins forming cellular components, most were associated with membranes and organelles. The proteins with biological processes functions were more evenly distributed in number, with the majority involved in cell proliferation or metabolism. This is consistent with the large number of general GO terms that have been assigned for basic metabolic processes in organisms.

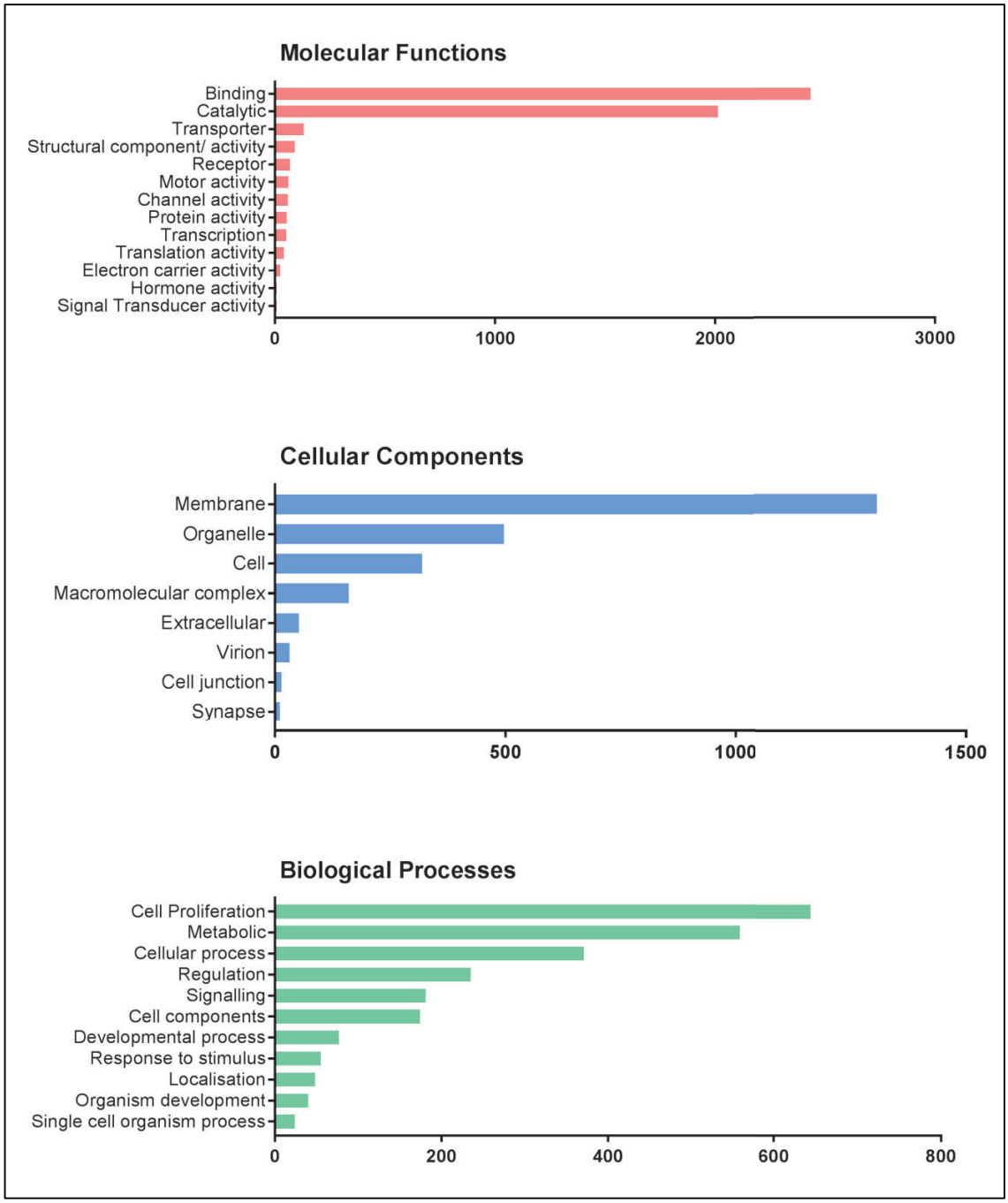


Figure 7.5: Gene Ontology classification of contigs with BLASTx hits.
Figure shows the distribution of hits to each of the categories – molecular function, biological processes and cellular processes. Results were obtained by searching accessions obtained by BLASTx in the UniProtKB Retrieve/ID mapping extension (http://www.uniprot.org/).

7.3.1.3 Identification and classification of P. clavata toxins

A large number of proteins/ peptides (32%) of the total contigs had no significant BLASTx hit. This indicates that some of these components represent potentially unique toxins in *P. clavata* venom. From this dataset, we identified 354 putative toxin sequences that were classified into 17 different toxin families (Figure 7.6). These toxins were identified through a manual search of the BLASTx output using over 50 different keywords denoting different toxin families such

as phospholipase, neurotoxin, hyaluronidase, allergen and toxin (see Table 7.1 in methods for all search terms). This toxin list also includes hits from the Tox|Note search that were not observed in the BLASTx output. Tox|Note queries a combined database that includes the manually curated UniProtKB database of animal venom toxins and ArachoServer. Analysis of the relative abundance of these toxins in the venom gland based on their TPM value revealed that the majority of the toxins were proteases (e.g. serine proteases and cathepsins) and metalloproteinases (Figure 7.6). However, neurotoxins were the most abundant contigs, based on TPM values. These were subsequently identified as poneratoxins through proteomic analyses. Contigs in the 'toxin-like' category include peptides such as ant omega-conotoxin-like proteins which, despite having an assigned mode of action to inhibit calcium channels, have not been proven to have any toxin activity. The full list of toxins identified through BLASTx and Tox|Note can be found in supplementary file 3.

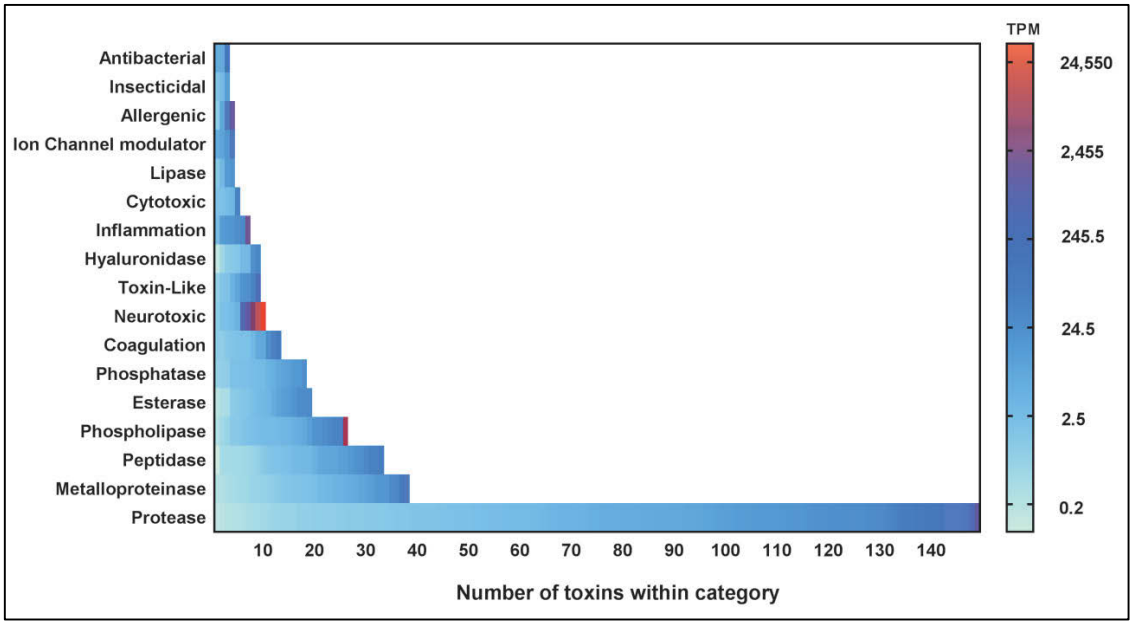


Figure 7.6: Abundance and expression levels of *P. clavata* toxins.
Heat map showing toxin contigs classified into 17 categories. Expression TPM levels are depicted by a colour scale between cyan and red, with cyan being least abundant (less than 0.2 transcripts per million) and red signifying over 1000 transcripts per million (see right-hand scale). The x-axis indicates the number of transcripts within each category.

Among all proteins, a poneratoxin contig was the second most abundantly expressed at TPM 39,657, with the structural protein actin being the most highly expressed protein (TPM 55,551). Within proteins identified as toxins however, the same poneratoxin contig was the most abundantly expressed toxin at several orders of magnitude higher than the next most abundant toxin, a phospholipase-like protein (TPM 6328). The most highly expressed toxins, identified through BLASTx, are summarized in Table 7.3. Other highly expressed toxins included phospholipases, serine proteases, arginine kinases, metalloproteinases and allergens. Other toxins had relatively lower TPMs such as defensin-2 (TPM 89) and cathepsin (TPM 77.2). An unusual annotation that was found was a translationally controlled tumor protein homolog from the ant *D. quadriceps*, which was identified by Tox|Note as a toxin involved in inflammation [88]. All *P. clavata* venom contigs shown in Table 7.3 were subsequently named using the nomenclature proposed in the methods section.

Although the majority of hits from Tox|Note were to toxins with proven toxicity, several unusual hits were observed to very large toxins such as α-latrotoxin. Alpha-Latrotoxin is known to be composed of several ankyrin repeats and only the full length protein, presently only detected in the venom of theridiid spiders, is neurotoxic [59, 89]. The transcripts from *P. clavata* only matched to the ankyrin domains of the 130 kDa toxin, but not other regions important for activity [59, 89]. Calglandulin was also identified by Tox|Note as one of the highly expressed 'toxin' transcripts in *P. clavata* (TPM: 196), but it is involved in the processing and release of venom proteins, rather than being a toxin per se [90]. Neither of these hits were included in the toxins list generated by Tox|Note.

Table 7.3: Top 20 most expressed toxins (based on TPM) from *P. clavata* venom gland transcriptome BLASTx and Tox|Note searches.
A complete list of all the BLASTx toxins can be found in supplementary file 3.

Toxin Name	BLASTx/ Tox|Note Protein name	NCBInr Accession	Species	Trinity ID	TPM	Toxin Category
δ-Paraponeritoxin-Pc1e_1 **δ-Paraponeritoxin-Pc1e_2** **δ-Paraponeritoxin-Pc1e_3** **δ-Paraponeritoxin-Pc1e_4**	Poneratoxin	P41736.1	*Paraponera clavata*	DN10254_c0_g26_i11 DN9710_c0_g11_i2 DN10254_c0_g26_i12 DN10254_c0_g26_i5	39,657 9279 2414 797	Neurotoxic
Phospholipase-A2-1a_1-P-clavata	PREDICTED: phospholipase A2	XP_011262712	*Camponotus floridanus*	DN9894_c0_g8_i1	6328	Phospholipase
Translationally-controlled-tumor-protein-1a-P-clavata	PREDICTED: translationally-controlled tumor protein homolog	XP_014489127	*Dinoponera quadriceps*	DN9555_c0_g10_i3	1229	Inflammation
Icarapin-1a-P-clavata	PREDICTED: icarapin-like	XP_011863857	*Vollenhovia emeryi*	DN9969_c1_g1_i4	969	Allergenic
Serine-Protease-1a-P-clavata **Serine-Protease-2a-P-clavata**	Serine proteinase stubble	EZA53191.1	*Cerapachys biroi*	DN10218_c2_g11_i1 DN10218_c2_g7_i1	835 364	Protease
Arginine-kinase-1a_1-P-clavata	PREDICTED: arginine kinase isoform X1	XP_014467324	*Dinoponera quadriceps*	DN9942_c0_g1_i6	382	Neurotoxic
U_1-Paraponeritoxin-Pc1a	PREDICTED: omega-conotoxin-like protein 1	XP_011257508	*Camponotus floridanus*	DN5808_c0_g1_i1	253	Unknown
Serine-Protease-3a-P-clavata	PREDICTED: serine protease easter-like isoform X2	XP_014488937	*Dinoponera quadriceps*	DN10218_c1_g1_i2	161	Protease
Serine-Protease-4a-P-clavata	Serine protease snake	KMR05322.1	*Lasius niger*	DN11547_c0_g1_i1	138	Protease
Venom-Allergen3-1a-P-clavata	PREDICTED: venom allergen 3-like	XP_012534360	*Monomorium pharaonis*	DN10202_c0_g4_i2	127	Allergenic
Cysteine-rich protein-1a-P-clavata	Cysteine-rich protein 1	KYN02327.1	*Cyphomyrmex costatus*	DN7413_c0_g4_i1	123	Ion channel modulator
Serine-Protease-1b-P-clavata **Serine-Protease-5a-P-clavata**	PREDICTED: serine protease 42-like isoform X3	XP_014488938	*Dinoponera quadriceps*	DN10218_c2_g8_i1 DN10218_c2_g6_i7	113 103	Protease

Disintegrin-and-metalloproteinase-domain-containing-protein-1a-P-clavata	Disintegrin and metalloproteinase domain-containing protein	XP_012242558	*Bombus impatiens*	DN21256_c0_g1_i1	113	Metalloprotein-ase
Disintegrin-and-metalloproteinase-domain-containing-protein-2a-P-clavata		EZA52689.1	*Cerapachys biroi*	DN20154_c0_g1_i1	75	
Lysosomal-aspartic-protease-1a-P-clavata	PREDICTED: lysosomal aspartic protease	XP_014468660	*Dinoponera quadriceps*	DN7128_c0_g1_i1	111	Protease
Defensin-2-like-1a-P-clavata	PREDICTED: defensin-2-like	XP_012537462	*Monomorium pharaonis*	DN5917_c0_g1_i1	89	Antibacterial
Serine-Protease-inhibitor-1a-P-clavata	PREDICTED: serine protease inhibitor 3/4-like isoform X2	XP_012539061	*Monomorium pharaonis*	DN9965_c0_g10_i1	86	Protease
Serine-Protease-6a-P-clavata	PREDICTED: venom serine protease-like	XP_014477136	*Dinoponera quadriceps*	DN9366_c0_g3_i1	85	Protease
Serine-Protease-7a-P-clavata	PREDICTED: serine protease 52-like	XP_011641802	*Pogonomyrmex barbatus*	DN10218_c2_g11_i2	84	Protease
Serine-Protease-8a-P-clavata	PREDICTED: serine protease snake-like	XP_011142823	*Harpegnathos saltator*	DN9987_c0_g2_i9	83	Protease
Cathepsin-L-1a-P-clavata	Cathepsin L	EFN75465.1	*Harpegnathos saltator*	DN9254_c0_g3_i1	77	Protease

7.3.2. Paraponera clavata venom proteome

Proteomics is an important tool used to confirm expressed proteins and peptides within venoms. However, a major problem in many ant venom proteomics studies is the lack of a comprehensive database that can be used to match the identified peptides. Therefore, the assembled contigs were initially translated in 6-frames using TransDecoder [75], generating 34,586 open reading frames (ORFs). This database and was then uploaded onto PEAKS Studio v8.5 as a database to search the shotgun proteomic LC-MS/MS data.

The mass spectra search against the transcriptome assembly was performed using PEAKS Studio v8.5 and generated peptide matches to 438 contigs with a minimum –10lgP score of 15. The contigs were manually classified into 11 functional categories based on their BLASTx ID and an e-value of e^{-4} (Figure 7.7). If a contig had no BLASTx hit, the matched peptides were searched on BLASTp to confirm that the protein was indeed uncharacterised. This approach was undertaken because a highly expressed contig (TPM: 9279) that was initially uncharacterised was later confirmed as poneratoxin using BLASTp. The majority of venom proteins were uncharacterised (Figure 7.7), followed by proteins with structural/ motor or cellular function which is comparable to the GO results for the transcriptome. Among other protein categories were toxin-like proteins, and proteins involved in metabolic processes, and regulation as well as chaperones, oxidoreductases, kinases and miscellaneous proteins. Positive matches between proteins identified by proteomics and transcriptomics allowed for the confirmation of proteomics data as well as complete predicted full-length protein sequences. For a full list of the proteins present in Figure 7.7, see supplementary file 4.

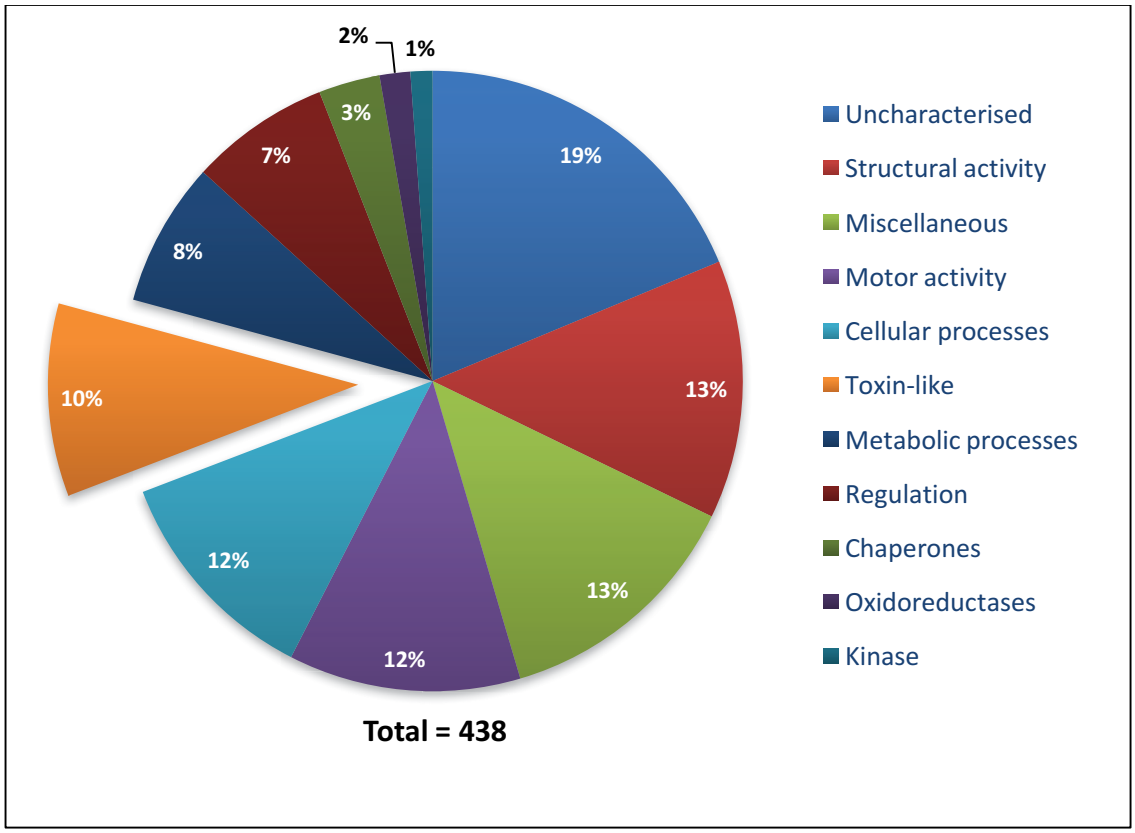


Figure 7.7: Protein categories in *P. clavata* venom.
Figure shows *P. clavata* contigs identified by mapping the proteome to the transcriptome. Annotation from BLASTx was used to place each contig into one of the above categories. The orange portion shows the proportion of contigs that encode toxins (10%).

There were a total of 44 contigs in the toxin-like category, which are shown in Table 7.4. The best match, based on a PEAKS score of –10lgP, was to hyaluronidase followed by dipeptidyl peptidase, poneratoxin and phospholipase A_2. All contigs in Table 7.4 were also re-named based on BLASTx or BLASTp hits according to the proposed nomenclature systems and queried for signal peptides using SignalP v4.1 [83]. Interestingly, many contigs had no predicted signal peptide despite the fact that they were seen in the proteome.

Table 7.4: *Paraponera clavata* toxin transcripts identified by proteomics analysis and their BLASTx hit.
Table shows transcripts which had at least two peptides from the shotgun MS/MS analysis identified using PEAKS. Also included are the TPM values, and BLASTx match (or BlastP if no match by BLASTx) for each transcript. Additional information regarding transcripts, accessions and full protein names can be found in supplementary file 4.

Toxin name	BLASTx hit	Species	Trinity ID	Unique peptide	% Cover	-10 lgP	TPM	SigP (y/n)
Hyalyronidase-1a_1-P-clavata	PREDICTED: hyaluronidase-like	*D. quadriceps*	DN10062_c0_g13_i9	70	65	333	38	Y
Hyalyronidase-1a_2-P-clavata			DN10062_c0_g13_i12	70	65	333	23	Y
Hyalyronidase-1a_3-P-clavata			DN10062_c0_g13_i7	70	65	333	3	Y
Hyalyronidase-1a_4-P-clavata			DN10062_c0_g13_i16	70	65	333	3	Y
Hyalyronidase-1a_5-P-clavata			DN10062_c0_g13_i15	70	65	333	1	Y
Hyalyronidase-1a_6-P-clavata			DN10062_c0_g13_i3	70	65	333	1	Y
Hyalyronidase-1a_7-P-clavata			DN10062_c0_g13_i10	70	65	333	1	Y
Hyalyronidase-1a_8-P-clavata			DN10062_c0_g13_i5	70	65	333	1	Y
Venom-dipeptidyl-peptidase-1a_1-P-clavata	PREDICTED: venom dipeptidyl peptidase 4	*L. humile*	DN8858_c0_g1_i1	26	29	281	60	N
Venom-dipeptidyl-peptidase-1a_2-P-clavata	PREDICTED: venom dipeptidyl peptidase 4		DN8858_c0_g1_i2	26	29	281	21	N
δ-Paraponeritoxin-Pc1e_2	Poneratoxin	*P. clavata*	DN9710_c0_g11_i2	20	45	260	9279	Y
δ- Paraponeritoxin-Pc1e_5			DN10254_c0_g26_i8	20	45	260	2	Y
δ- Paraponeritoxin-Pc1e_6			DN10254_c0_g26_i3	20	45	260	0	Y
Phospholipase-A2-1a_1-P-clavata	PREDICTED: phospholipase A2	*C. floridanus*	DN9894_c0_g8_i1	24	48	235	6328	Y
Phospholipase-A2-1a_2-P-clavata			DN9894_c0_g2_i2	24	48	235	7	Y
Venom-Allergen3-1a-P-clavata	PREDICTED: venom allergen 3-like	*M. pharaonis*	DN10202_c0_g4_i2	14	44	198	127	Y
Venom-Allergen3-1b-P-clavata			DN10202_c0_g4_i1	14	53	198	1	Y
Carboxypeptidase-Q-1a_1-P-clavata	PREDICTED: carboxypeptidase Q	*C. biroi*	DN8288_c0_g1_i2	12	22	180	3	N
Carboxypeptidase-Q-1a_2-P-clavata			DN8288_c0_g1_i1	12	23	180	2	N

Arginine-kinase-1a_1-P-clavata	PREDICTED: arginine kinase isoform X1	*D. quadriceps*	DN9942_c0_g1_i6	10	32	169	382	Y
Arginine-kinase-1a_2-P-clavata			DN9942_c0_g1_i1	10	32	169	4	N
Pancreatic-lipase-1a-P-clavata	Pancreatic lipase-related protein 2	*C. floridanus*	DN10238_c1_g31_i1	8	23	169	15	Y
Pancreatic-lipase-1b-P-clavata			DN10238_c1_g10_i1	8	25	169	3	Y
Serine-Protease-6a-P-clavata	PREDICTED: venom serine protease-like	*D. quadriceps*	DN9366_c0_g3_i1	6	10	135	85	Y
Matrix-metalloproteinase-14-like-1a-P-clavata	Matrix metalloproteinase-14	*C. biroi*	DN18586_c0_g1_i1	3	23	104	5	N
Serine-Protease-inhibitor-1a-P-clavata	PREDICTED: serine protease inhibitor 3/4-like isoform X2	*M. pharaonis*	DN9965_c0_g10_i1	4	11	97	86	Y
Alaserpin-X5-like-1a-P-clavata	PREDICTED: alaserpin-like isoform X5	*P. barbatus*	DN9965_c0_g10_i2	4	12	97	18	Y
Pancreatic-triacylglycerol-lipase-X1-like-1a-P-clavata	PREDICTED: pancreatic triacylglycerol lipase-like isoform X1	*D. quadriceps*	DN8997_c0_g2_i2	3	6	93	5	Y
Pancreatic-triacylglycerol-lipase-X1-like-2a-P-clavata			DN8997_c0_g2_i3	1	3	37	2	Y
Pancreatic-triacylglycerol-lipase-X1-like-3a-P-clavata			DN8997_c0_g2_i1	1	8	37	1	Y
Venom-acid-phosphatase-1a-P-clavata	PREDICTED: venom acid phosphatase Acph-1	*C. biroi*	DN5022_c0_g1_i1	3	7	63	5	N
Iron-zinc-purple-acid phosphatase-like-1a_1-P-clavata	Iron zinc purple acid phosphatase-like protein	*L. niger*	DN8672_c0_g1_i2	2	4	55	17	N
Iron-zinc-purple-acid phosphatase-like-1a_2-P-clavata			DN8672_c0_g1_i1	2	4	55	9	N
Alkaline-Phosphatase-1a_1-P-clavata	PREDICTED: alkaline phosphatase 4-like	*L. humile*	DN9670_c0_g1_i1	1	2	51	24	N
Alkaline-Phosphatase-1a_2-P-clavata	Alkaline phosphatase 4	*H. saltator*	DN9670_c0_g1_i2	1	2	51	2.3	N

Neprilysin-11-isoform-X2-1a_1-P-clavata	PREDICTED: neprilysin-11 isoform X2	*D. quadriceps*	DN10125_c0_g9_i1	2	4	43	4.5	Y
Neprilysin-11--isoform-X2-1a_2-P-clavata	PREDICTED: neprilysin-11 isoform X2		DN10125_c0_g18_i1	2	4	43	1.7	Y
Matrix-metalloproteinase-14-like-2a-P-clavata	PREDICTED: matrix metalloproteinase-14 isoform	*D. quadriceps*	DN21699_c0_g1_i1	2	6	41	1.9	Y
Protein-5NUC-1a-P-clavata	PREDICTED: protein 5NUC-like	*Wasmannia auropunctata*	DN10097_c1_g11_i1	1	2	37	25.2	N
Icarapin-1a-P-clavata	PREDICTED: icarapin-like	*D. quadriceps*	DN9969_c1_g1_i4	1	7	32	969	N
Cytosolic-dipeptidase-1a-P-clavata	PREDICTED: cytosolic non-specific dipeptidase	*D. quadriceps*	DN9768_c0_g1_i2	1	2	30	1.8	Y
Phenoloxidase-2-like-1a-P-clavata	PREDICTED: phenoloxidase 2	*S. invicta*	DN15311_c0_g1_i1	1	4	24	3.2	Y

7.3.3. Paraponera clavata toxins

Many of the toxin sequences identified in this study have not been previously reported in *P. clavata*. Therefore, a comparative study with homologous sequences was undertaken to determine the potential pharmacological targets. This was achieved through comparison of critical residues for activity between sets of homologous sequences. Proteins chosen for alignment are detailed below.

7.3.3.1 Poneratoxin

We report for the first time the protein precursor of poneratoxin, which includes the signal peptide and propeptide. The most abundant toxin identified in *P. clavata's* transcriptome is poneratoxin, as seen in Table 7.3, which is consistent with previous investigations [18, 22, 68]. To determine sequence homology with the previously identified poneratoxins, the contigs of δ-paraponeritoxin-Pc1e (Pc1e_1-6), from French Guiana, were aligned with previously published poneratoxin sequences from Brazil, Peru, Panama and Costa Rica (Pc1a-d) using Mega 7 [85]. The δ-paraponeritoxin-Pc1e isoforms have identical amino acid residues, however the nucleotide sequences were different (hence all were Pc1e). The δ-paraponeritoxin-Pc1e transcripts included a signal peptide (residues 1–24) and a propeptide (25–38) preceding the first amino acid of the published mature ORF (residues 39–65; with a calculated average mass of 2968.7 Da) (Figure 7.8).

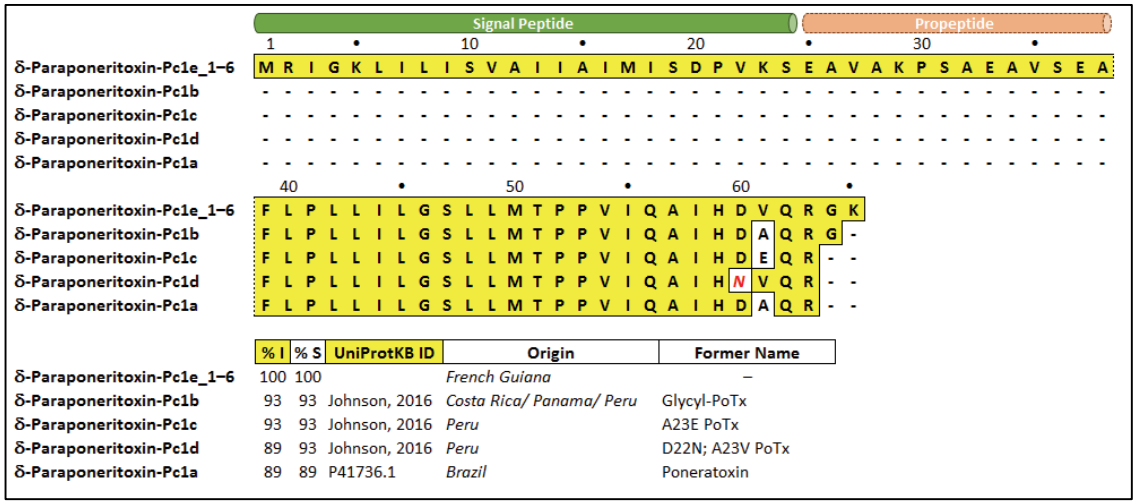


Figure 7.8: Amino acid sequence alignment of δ-paraponeritoxin-Pc1e isoforms.
Figure shows the alignment of six of the *P. clavata* transcripts matching to poneratoxin (δ-paraponeritoxin-Pc1e_1 to Pc1e_6) aligned to published δ-paraponeritoxin sequences. Predicted signal peptide (www.cbs.dtu.dk/services/SignalP/) is indicated by a green bar above the sequences and the predicted propeptide is indicated by an orange bar. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Percentage identity (%I) is relative to δ-paraponeritoxin-Pc1e while percentage similarity (%S) includes conservatively substituted residues. 'UniProtKB ID' indicates the UniProt Knowledge base ID code or source reference. 'Origin' refers to the country where the venom was collected. 'Former name' refers to the previously published name assigned to that sequence. Previously published sequences (δ-Paraponeritoxin-Pc1b, -Pc1c and -Pc1d) represent isoforms of -Pc1a (formerly poneratoxin)[20, 22].

The mature toxin present in this transcriptome also has an alanine to valine substitution at position 61 (A23V in the mature toxin), and additional GK residues at the C-terminus, compared to the original Brazilian δ-paraponeritoxin-Pc1a [22]. It seems to be most similar to the recently published "glycl-PoTx" [20], re-named here as δ-paraponeritoxin-Pc1b, as it also contains the additional C-terminal glycine but not the lysine residues. Nevertheless the shotgun proteomics data found a mature δ-paraponeritoxin-Pc1e mass of 2783 Da, consistent with the terminal GK residues undergoing cleavage and a loss of 185.1 Da (Figure 7.9). If these two C-terminal residues are removed from the alignment, δ-paraponeritoxin-Pc1e has 96% identity and 100% similarity to the Peruvian isoform Pc1d (Figure 7.8).

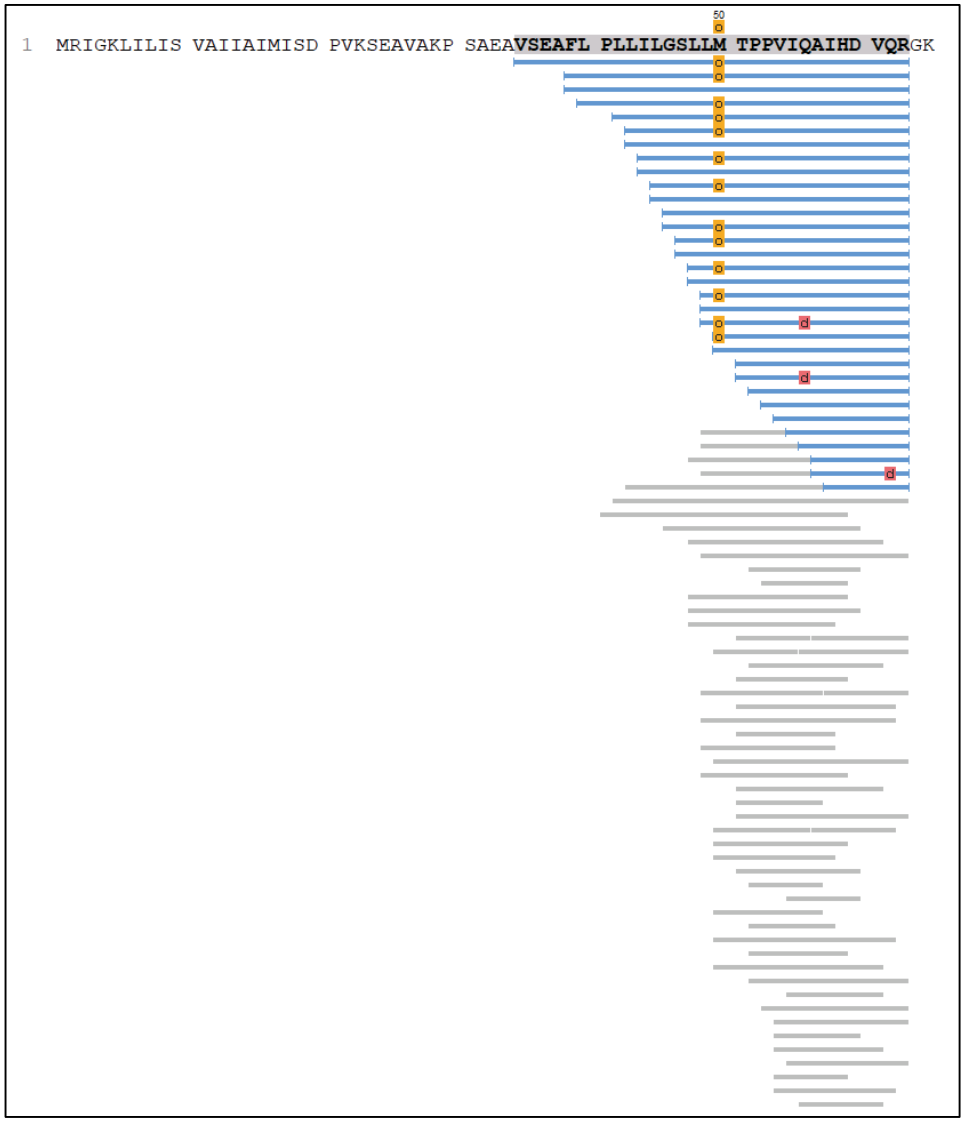


Figure 7.9: δ-Paraponeritoxin-Pc1e LC-MS/MS coverage.
Figure shows translated δ-paraponeritoxin-Pc1e_1 transcript coverage obtained from the shotgun mass spectrometer search result from PEAKS. Blue lines indicate an identified peptide sequence and grey bars indicate a *de novo* only tag match. Yellow 'o' indicates oxidation of methionine and red 'd' indicates deamidation.

7.3.3.2 Conotoxin-like-protein

Peptides highly homologous to ant omega-conotoxin-like protein were also found in the *P. clavata* venom transcriptome (Figure 7.10). The sequence of the *P. clavata* omega-conotoxin-like protein was 90% identical and 96% homologous to that of *Camponatus floridanus* and 90% homologous to the ant *Trachymyrmex cornetzi.* Interestingly, identity notably decreased when compared to the same peptide from the bee *Apis florea* (63%), however the homology was still relatively high (81%). Manual BLASTx results did not produce significant hits to any marine cone snail peptides. We therefore re-searched with the organism filter restricted to "Conus" in

order to confirm the similarity of these ant "conotoxin-like-proteins" to that of cone snails. However, as seen in Figure 7.10, the highest match which was to *Conus lividus,* which had a very different sequence to all the ant sequences, with an identity of only 15% and homology of 31% to that of *P. clavata*. The alignment also highlighted that the main conserved amino acids between the sequences were the cysteine residues, which are normally conserved due to their role in the structural framework. Therefore, we did not name the *P. clavata* contig using the same nomenclature as the ant omega-conotoxin-like proteins, instead, we named it U_1-Paraponeritoxin-Pc1a because the main activity associated with omega-conotoxins, an action to inhibit voltage-gated calcium channels, is very unlikely given the low sequence homology.

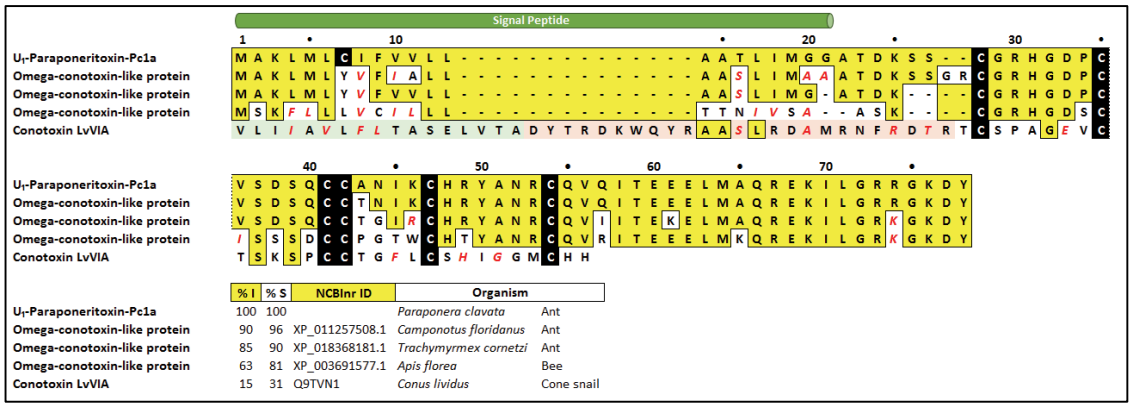


Figure 7.10: Amino acid sequence alignment of omega-conotoxin-like contigs.
The figure shows alignment of *P. clavata* contigs and proteins from the ant species *C. floridanus* and *T. cornetzi,* the bee *Apis florea* and the marine cone snail *Conus lividus.* The predicted signal peptide for all sequences (except conotoxin LvVIA) is indicated by a green bar above the sequences. The signal peptide for conotoxin LvVIA is indicated by light green shading, while orange shading indicates the propeptide. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to U_1-paraponeritoxin-Pc1a, while percentage similarity (%S) includes conservatively substituted residues.

7.3.3.3 Phospholipase A_2 (PLA_2)

The second most highly expressed transcript in *P. clavata* venom after poneratoxin (TPM 6328) was a 202 residue phospholipase A_2 (PLA_2) (Table 7.3). These proteins have previously shown to be highly abundant in *P. clavata* [71]. PLA_2 proteins are not unique to ant venoms, but are found in almost all venomous animals [91-94]. The size of the *P. clavata* PLA_2 and the presence of a His48/Asp49 pair, however, suggests that this is a group III PLA_2 also present in hymenopteran venoms including bees and ants [61, 95]. The presence of PLA_2 in the venom proteome (Table 7.4) also confirms its secretion in high enough abundance for detection by mass spectrometry. The combined transcriptomic and proteomic evidence increases the reliability of the deduced *P. clavata* PLA_2 amino acid sequence.

BLASTx analysis identified a number of homologous insect PLA_2 sequences as shown in Figure 7.11. For example, PLA_2 from *C. floridanus* shares a 68% homology while other ant species show a 60% homology. Interestingly, identity was not very high to any of the ant species, with the highest being 52% and the next highest identity to the beetle *Tribolium castaneum* with an identity of 36%. Although the identities were not high, it should be noted that all critical residues are conserved in the *P. clavata* contig. For example, almost all cysteine residues were also conserved and the previously described calcium binding loop and active site is relatively conserved across the sequences, particularly the critical histidine 48, aspartate 49 and glycine 32 residues.

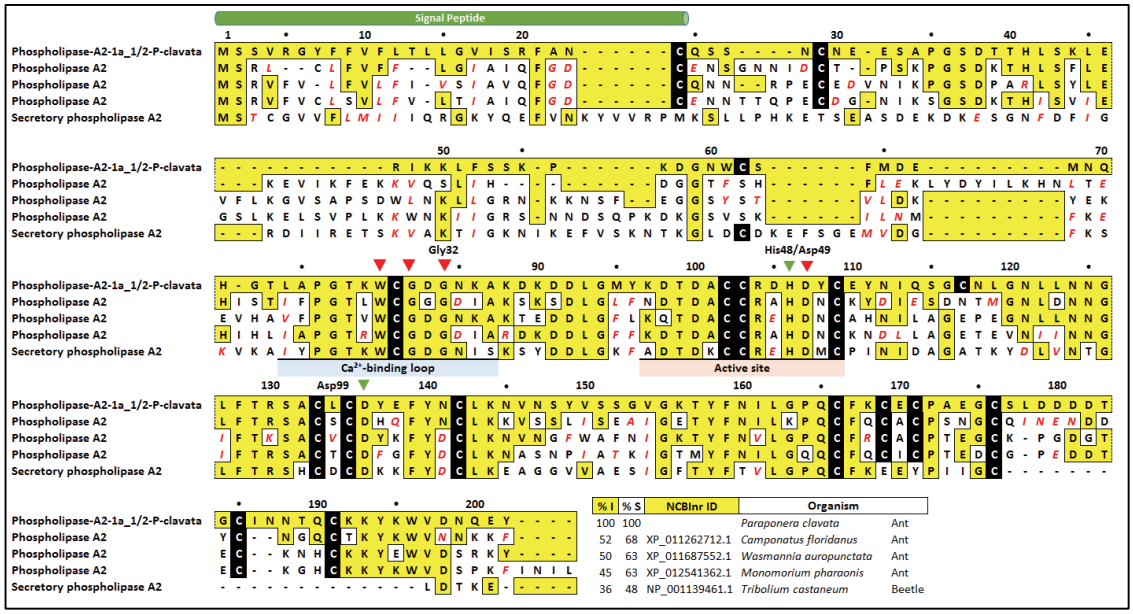


Figure 7.11: Amino acid alignment of *P. clavata* phospholipase A2 isoforms.
The figure shows the alignment of *P. clavata*'s transcripts with Phospholipase A2-like proteins from the ant species *Camponatus floridanus, Wasmannia auropunctata, Monomorium pharaonis* and the red flour beetle *Tribolium castaneum.* Predicted signal peptide (green bar) is indicated above the sequences. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimize the alignments. The blue triangle indicates the predicted N-terminus [92]. The underlined coloured regions indicate the Ca^{2+} binding loop (blue), active site (orange) and conserved region (green) [92, 95, 96]. The red triangles indicate residues critical for calcium binding and the green triangles indicate the active site residues [92, 95, 97, 98]. Percentage identity (%I) is relative to the first peptide of each family, while percentage similarity (%S) includes conservatively substituted residues.

7.3.3.4 Hyaluronidase

A transcript composed of 357 residues with sequence homology to hyaluronidase was identified by a BLASTx search of the NCBInr. It's presence and expression in the venom was confirmed through the shotgun proteomics experiment as the protein with highest –10lgP score of 333. There were a total of eight contigs (1a_1 to 1a_8) for hyaluronidase-1a-P-clavata (Figure 7.12). The translated protein sequence had the highest homology to a hyaluronidase-like sequence from the ant *Pogonomyrmex barbatus* (85%; Figure 7.12). All cysteines were conserved across matched proteins with the majority of the differences found near the N-terminus of the sequences. All ant hyaluronidases contain the conserved active site residues previously described in bee venom hyaluronidases (Asp111, Phe112 and Glu113, Glu247; numbered according to the bee venom hyaluronidase) [99].

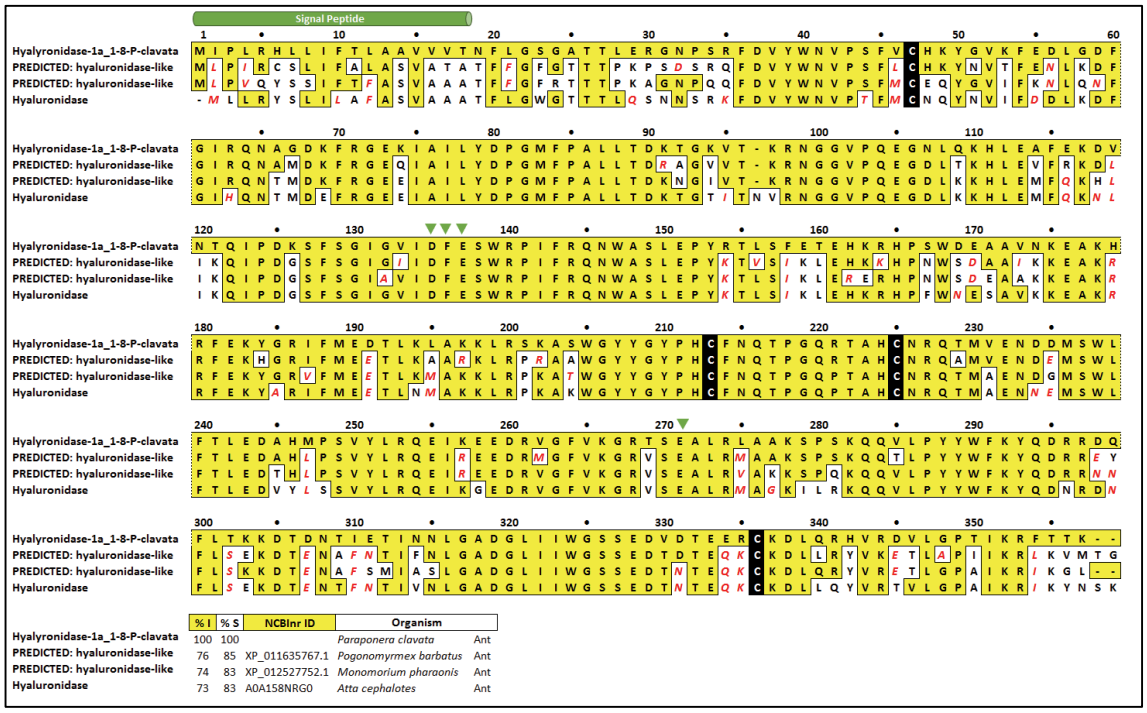


Figure 7.12: Alignment of hyaluronidase-like proteins from *P. clavata* and other ants.
The Figure shows the alignment of *P. clavata* transcripts (hyaluronidase-1a_1-P-clavata through to '-1a_8' identical ORFs) and hyaluronidase proteins from the ants *Pogonomyrmex barbatus*, *Monomorium pharaonis* and *Atta cephalotes*. A signal peptide was predicted for all sequences except for that of *M. pharonis* (indicated by a green bar above the sequences). All signal peptides were from amino acid 1–18, except that of *A. cephalotes* which was from 1-17. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to hyaluronidase-1a_1-P-clavata, while percentage similarity (%S) includes conservatively substituted residues. Green arrowheads denote the active-site residues Asp, Phe, Glu and Glu of bee venom hyaluronidase [99].

7.3.3.5 Icarapin

The allergenic protein icarapin, is another protein that were detected in both the proteome and the transcriptome. The transcript icarapin-1-P-clavata is a 163-residue protein that had relatively low homology to known proteins, with the highest match being to icarapin from the ant *Linepithema humile* (58% identity; Figure 7.13). Interestingly, from residues 73 to 118 (numbering from icarapin-1-P-clavata) all sequences share very high homology, which is where the consensus icarapin sequence is located [100]. However, the *P. clavata* transcript was not 100% identical to the consensus, as there were a few substitutions (positions 87, 100, 108 and 111) and a threonine insertion at position 92.

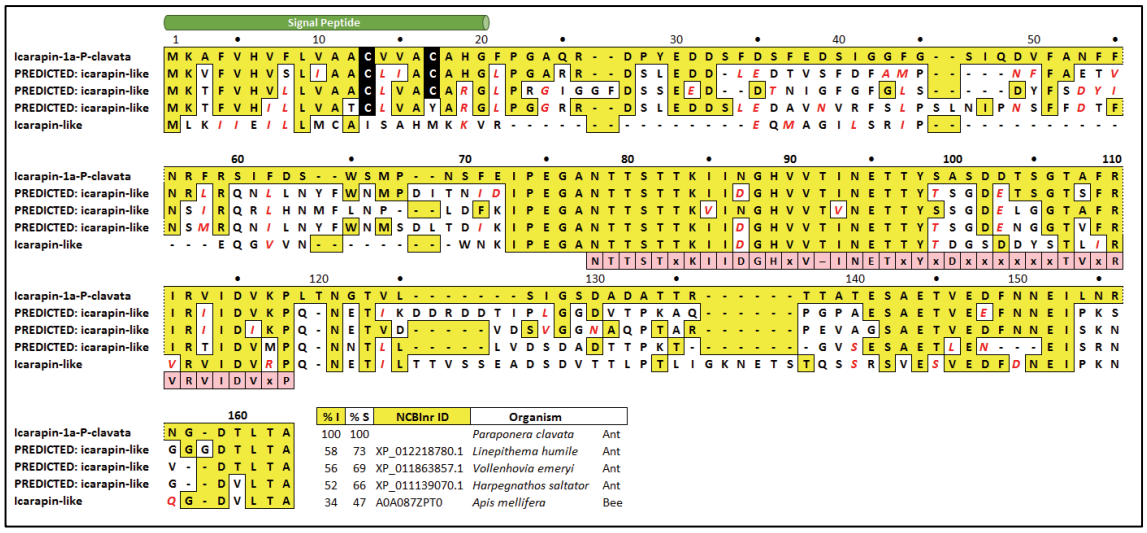


Figure 7.13: Alignment of icarapin-like proteins.
The figure shows the alignment of the *P. clavata* transcript with homologous icarapin-like proteins from the ant species *Linepithema humile*, *Vollenhovia emeryi*, *Harpegnathos saltator* and the bee *Apis mellifera.* Predicted signal peptide (green bar) is indicated above the sequences, it was the same size for all sequences except *A. mellifera's* which ends at amino acid 17. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimise the alignments. Boxed residues in pink indicate the conserved icarapin residues described by Pieren (2006) [100]. Percentage identity (%I) is relative to icarapin-1a-P-clavata, while percentage similarity (%S) includes conservatively substituted residues.

7.3.3.6 Arginine Kinase

Of the all the proteins described here, the most conserved of all matches was that of the 355 residue arginine kinase, with the lowest percentage homology of 97% (Figure 7.14). This protein was detected initially by BLASTx homology searching; it was then confirmed through proteomic analysis of the venom. Importantly, all residues critical for arginine binding, ATP binding and catalysis are conserved amongst homologous sequences [101]. Additionally, the two specificity loops implicated in the arginine kinase activity and the phosphagen kinase site were also conserved. No signal peptide was predicted by SignalP v4.1 for the *P. clavata* transcript or the other matched proteins.

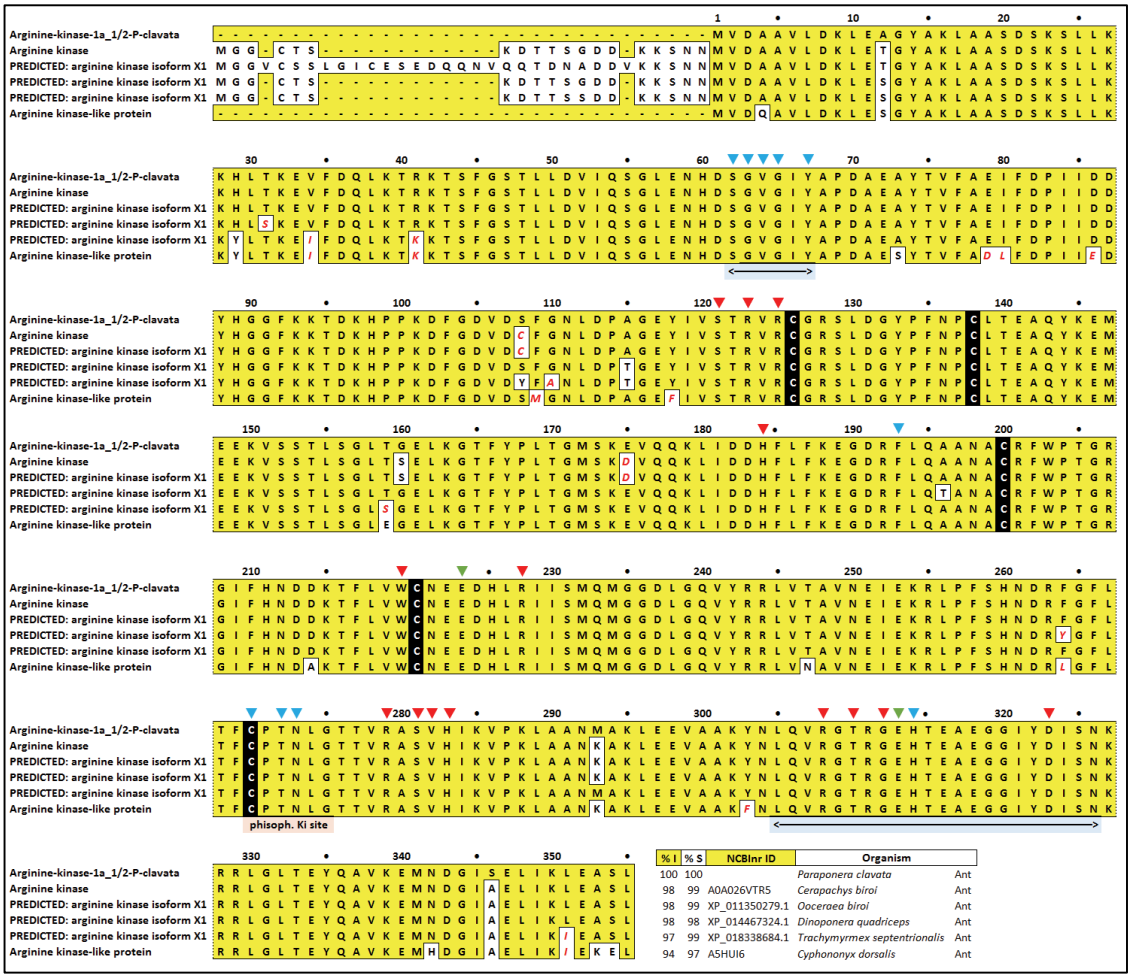


Figure 7.14: Amino acid alignment of arginine kinase transcripts.
Alignment of arginine-kinase-1a-P-clavata to proteins from the ant species *Cerapachys biroi*, *Ooceraea biroi*, *Dinoponera quadriceps*, *Trachymyrmex septentrionalis* and *Cyphononyx dorsalis*. Identical residues in the peptide sequences are boxed in yellow while conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimize the alignments. Percentage identity (%I) is relative to arginine-kinase-1a-P-clavata, while percentage similarity (%S) includes conservatively substituted residues. Blue triangles indicate amino acid residues implicated in arginine binding, red triangles show residues important for ATP binding and green triangles indicate residues important for catalytic binding [102]. Blue boxed regions with '<->' indicates the specificity loops [101]. Finally, the orange boxed region labelled "phosphagen kinase" represents the phosphagen kinase site [102].

7.3.3.7 Serine Proteases

Serine proteases were the most common toxins detected in the BLASTx search (Figure 7.6). The longest transcript recovered from our transcriptome, translated to 256 amino acid residues and was named serine-protease-5a-P-clavata. Interestingly, we also found transcripts that appear to be truncated versions of isoform 6a, namely: isoforms 5a and 8a. These isoforms were not only truncated but also had a few non-conserved and conserved substitutions. The BLASTx match with the highest homology was to a serine proteinase stubble protein from the ant *Ooceraea biroi* (81% identity, 89% homology). However, it must be noted that the *O. biroi* protein has 173 amino acids preceding the ones shown in Figure 7.15, which suggests that the *P. clavata* proteins are all truncated versions of this full-length proteinase. An important feature of serine proteinases is the catalytic triad (His37, Asp87 and Ser190, naming according to the *P. clavata* serine proteinase) which was conserved across all the ant serine proteinases shown in Figure 7.15 apart from the serine-protease-8a-P-clavata, which had a glycine instead of a serine at position 190. The residues important for the specificity pocket were also present in all of the *P. clavata* isoforms except serine-protease-5a-P-clavata which had a conservatively substituted alanine instead of a glycine.

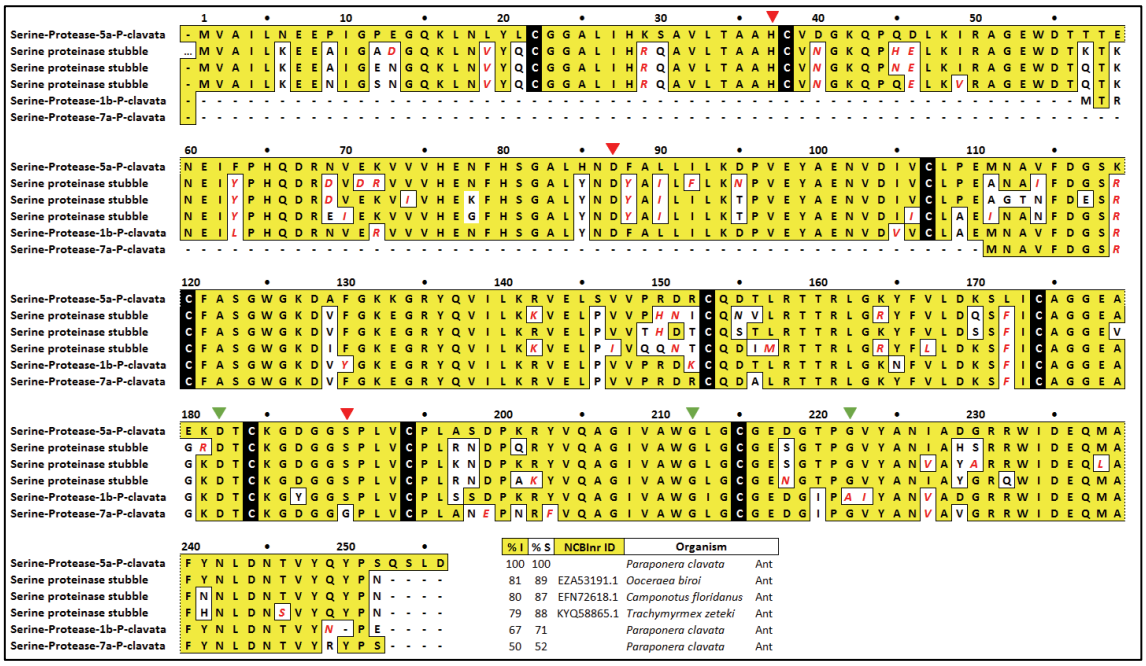


Figure 7.15: Alignment of serine proteases from *P. clavata* and other insect species.

The Figure shows alignment of three *P. clavata* transcripts (serine-protease-5a-P-clavata, serine-protease-1b-P-clavata and serine-protease-7a-P-clavata) that have homology to other serine protease-like proteins from the ant species *Ooceraea biroi*, *Camponotus floridanus* and *Trachymyrmex zeteki.* The only sequence with a predicted signal peptide was that of *O. biroi* (not shown, but started at amino acid 174). Identical residues in the peptide sequences are boxed in yellow whilst conservative substitutions are shown in red italic text. Cysteines are boxed in black. Gaps were introduced to optimize the alignments. Solid green triangles indicate the important residues for the specificity pocket [103, 104]. Solid red triangles indicate the catalytic triad (His, Asp, Ser) [105, 106]. Percentage identity (%I) is relative to serine-protease-5a-P-clavata, while percentage similarity (%S) includes conservatively substituted residues.

7.3.4. Novel P. clavata toxins

In order to identify novel peptide toxins in the venom of *P. clavata*, the transcriptome was analysed with Tox|Note, a pipeline developed mainly for the identification of novel spider-venom toxins. This platform was chosen due to the lack of an ant- or hymenopteran-venom specific platforms. The output contains the predicted toxin/toxin-like sequences and information regarding the predicted prepropeptides and mature toxin. Predicted peptide toxins were then sorted manually to identify those with a 'toxin-like' cysteine framework, defined as peptides with an even number of greater than 4 cysteines, and then double checked for novelty by BLASTp. As seen in Figure 7.16, there were a total of 190 peptides conforming to these conditions with the majority (72%) having 4 cysteines and the highest number of cysteines being 12 which would indicate six disulfide bonds. The majority of the peptides with cysteine patterns consistent with an ICK structural framework (containing 6 or more cysteines conforming to "C-C-CC-C-C") showed very low expression, with the highest being a TPM of 44, while conotoxin cysteine framework 1 toxins (containing 4 cysteines conforming to "CC-C-C") had higher transcript levels overall, with the highest TPM value of 771 for one of the contigs.

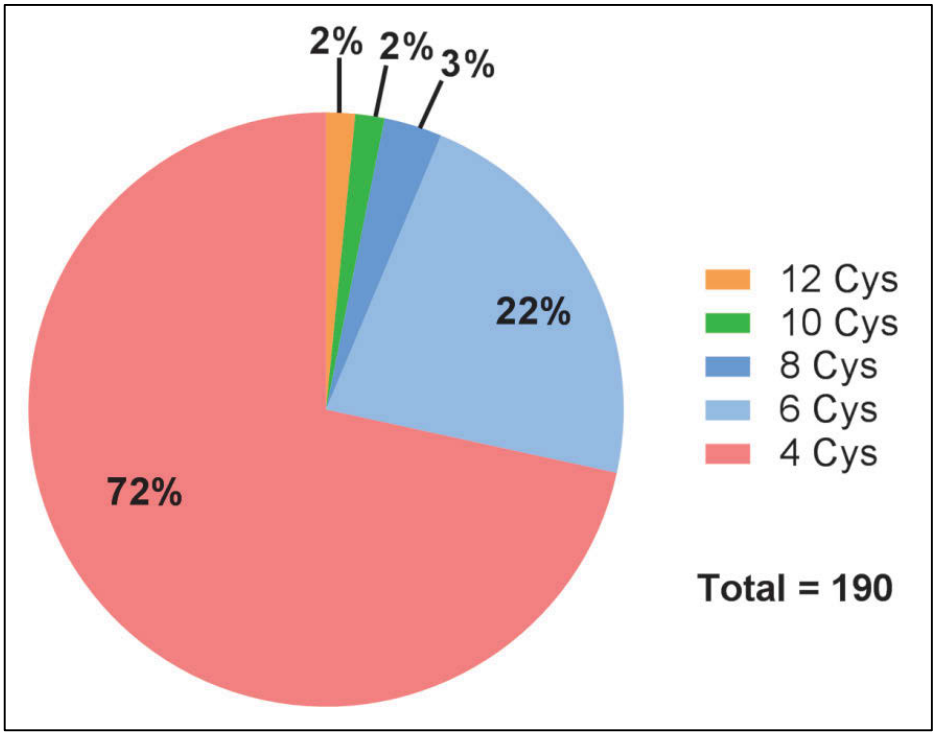


Figure 7.16: Distribution of novel *P. clavata* toxin-like peptides with four or more cysteines. Predicted peptide toxins were identified by Tox|Note.

Upon closer inspection of the sequences that contained six cysteines, we found seven peptides that adhered to a canonical inhibitor cysteine knot (ICK) framework similar to the conotoxin VI/VII framework as shown in Figure 7.17 panel A. There were also other peptides conforming to the ICK framework, however, they had 8 and 10 cysteines. These are shown in Figure 7.17 panel B, however the disulfide bonding pattern of the cysteine residues was not predicted as there are many combinations that could occur.

Figure 7.17: Novel *P. clavata* toxins with a predicted inhibitor cysteine knot (ICK) structural framework.
Predicted peptide toxins were identified using Tox|Note. Green residues indicate signal peptides. Cysteine residues are indicated in bold red text. Panel A: predicted novel ICK peptides with six cysteine residues and the predicted disulphide bonding pattern, according to conotoxin framework VI/VII, is represented by dotted blue lines. Panel B: predicted novel ICK peptides with eight and ten cysteine residues.

We also found nine sequences conforming to the conotoxin framework I, which are shown in Figure 7.18. We predicted the disulphide-bond pairs according to the conotoxin frameworks (www.conoserver.org). Of all the peptides conforming to framework I or VI/VII framework, only one had a predicted pro-peptide (Figure 7.18).

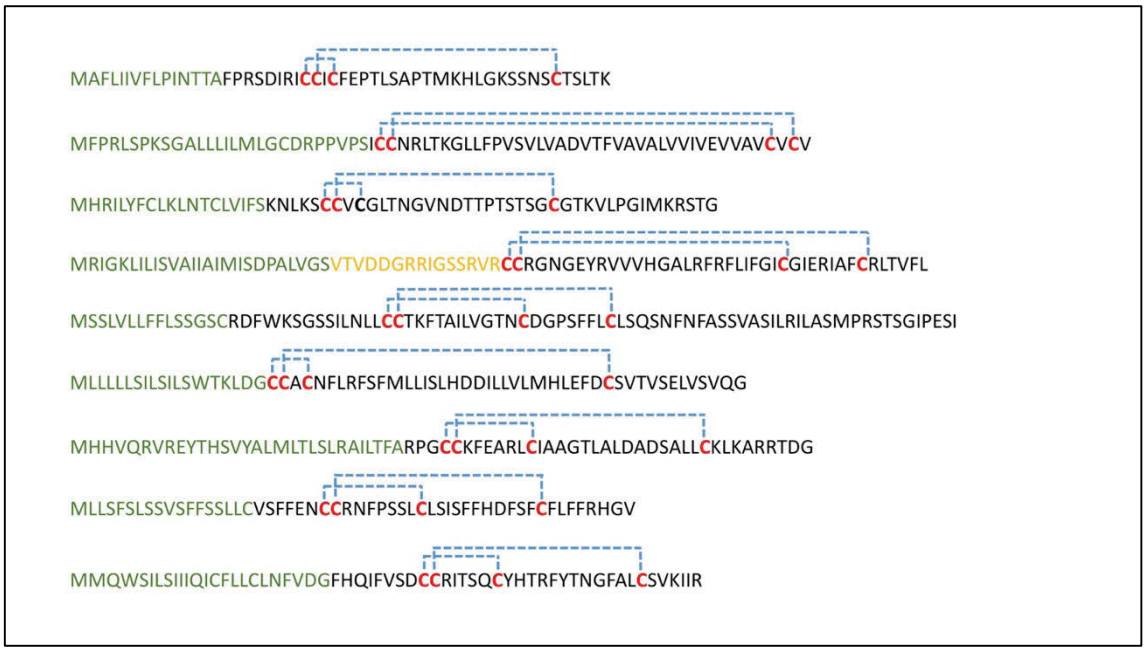


Figure 7.18: Novel *P. clavata* toxins with predicted conotoxin framework I cysteine structural framework.

Predicted peptide toxins were identified using Tox|Note. Green residues indicate signal peptides whilst residues in yellow indicate propeptides. Cysteine residues are indicated in bold red text and the predicted disulphide bonding pattern, according to conotoxin framework I, is represented by dotted blue lines.

7.4. Discussion

7.4.1. Paraponera clavata transcriptome

Ants are one of the most successful social insects with almost 16,000 currently described species (https://www.antweb.org/), however, their venom remains extremely understudied, with only 3 transcriptomes published to date [61-63]. Here, we present the transcriptome and proteome of the bullet ant *P. clavata*, which is known to inflict the most painful insect sting [107, 108]. Understanding this venom is critical for understanding the basis for envenomation but more importantly, as a source of potential bioactive compounds for development of insecticides or therapeutics [20, 62]. In the past, such an approach was not possible due to the lack of sequencing data and the large amounts of sample required for analysis. However, the advancement in the sensitivity of separation and mass spectrometry techniques, coupled with multi-"omic" technologies and their increasingly lower cost, has facilitated the identification of bioactive compounds from venom glands.

Hymenopteran venoms, including ants, are known to be rich in peptides and proteins such as enzymes, putative toxins and regulatory proteins [61]. The present transcriptomic analysis yielded a total of 54,242 assembled contigs (31,586 ORFs), which is the largest transcriptomic dataset reported to date, exceeding that of the previously published ants *T. bicarinatum* (36,042 contigs), *D. quadriceps* (18,546 contigs) and *O. monticola* (49,639 contigs) [61-63]. In order to annotate the transcriptome of *P. clavata*, a BLASTx search of the NCBInr database was performed and generated over 37,000 significant hits (68%) which were queried for GO terms to determine the types of functions of the peptides and proteins. The number of significant matches in the NCBInr database highlights the continuing progress to annotate genes, as the previous ant venom transcriptomes had a lower number of hits, while the scorpion *Urodacus yaschenkoi* which is from a far more studied group of venomous animals had only 51% annotation [109]. It was found that overall, the molecular functions predominated, which was also seen in the ant venom gland transcriptomes [61] and jellyfish transcriptome [110]. The most common molecular functions of *P. clavata* venom peptides and proteins were binding and catalytic activity and the most common biological processes related to metabolic and cellular processes, which was also the case with the ant *T. bicarinatum* [61]. These proteins are likely involved in the folding of peptides, their cleavage or their secretion to the venom gland.

The majority of contigs with BLASTx hits matched to components in the venom of the ants *H. saltator* and *D. quadriceps* which was not surprising, given the availability of a *D. quadriceps* transcriptome and an *H. saltator* genome [111]. Both of these ants also belong to the

subfamily Ponerinae, one of the phylogenetically closest subfamilies to *Paraponerinae* [17]. This is further proven by the lack of matches to *T. bicarinatum,* an ant belonging to the subfamily *Myrmicinae*, a phylogenetically distant subfamily to *Paraponerinae* (see Chapter 2, Figure 1 for a phylogenetic tree of ants). The lack of available data on ant venom proteins was also highlighted during functional annotation of the *P. clavata* transcriptome using GO terms, where the majority of protein hits from BLASTx did not have a GO term. However, this is attributable to the fact that most of the BLASTx hits from the NCBInr database were to predicted proteins, which are not within the curated UniProtKB database used for GO annotation.

Previous ant venom transcriptomes showed that ants contain several toxin-like transcripts [61-63]. However, the lack of a pipeline for putative toxin identification meant that manual searches of the BLASTx output were used to identify toxins. Our analysis using the recently developed Tox|Note pipeline found 400 toxins that were classified into 17 categories based on protein name or function. However, this number could be higher as there might be several unmatched proteins to toxins that were not identified as there were no homologous proteins in the databases searched. Whilst homologs of some of the identified *P. clavata* proteins had already been described in the venom of other ants, others which matched to other venomous animals were identified and described for the first time in this investigation.

The most abundant group of toxins, in terms of contig numbers, were proteases with 149 different isoforms in *P. clavata* venom. This was not expected as other ant transcriptomes reported allergens and phospholipases as the abundant proteins in ant venoms [61, 62]. Additionally, previous 2D-PAGE proteomic studies of *P. clavata* venom suggested that δ-paraponeritoxin-Pc1e (formerly poneratoxin) and various phospholipases were more prevalent [71]. Importantly however, the expression levels (i.e. TPM values) of the various proteases were relatively (all < 834 TPM) while the number of reads of both poneratoxin and PLA_2 confirmed the high levels in the venom.

This study also identified many novel toxin families not seen in previous proteomic analyses of *P. clavata* [71] including pro-inflammatory proteins, allergens and ICK peptides. The following section summarises the most important characteristics of a selected group of eight highly expressed protein families.

7.4.1.1 Neurotoxins

Paraponera clavata is one of only a few ant species to have been studied previously and from which a peptide toxin has been sequenced and tested for activity [18, 20, 22, 68]. The majority of the pain inflicted by the sting of *P. clavata* has been attributed to the 25-residue peptide poneratoxin (now renamed δ-paraponeritoxin-Pc1a), first identified by Piek in 1991 [18, 22, 68]. Surprisingly, our initial BLASTx search of contigs against the NCBInr database did not identify δ-paraponeritoxin-Pc1a as a hit. However, a BLASTp search of peptides identified by LC-MS/MS analysis identified the toxin and matched this sequence to the contigs δ-paraponeritoxin-Pc1e in the translated database generated in this work. One of these contigs (δ-paraponeritoxin-Pc1e) showed very high expression levels (TPM 39,657) and revealed that poneratoxin was indeed the most abundant peptide in the venom of *P. clavata*. This is a prime example of the advantages of combining proteomics and transcriptomics to identify venom components. Importantly, we report, for the first time, a signal and propeptide for a δ-paraponeritoxin. It was safe to assume that the propeptide is the 14-residue segment (EAVAKPSAEAVSEA) following the 24-residue signal peptide (MRIGKLILISVAIIAIMISDPVKS), as all proteomics experiments and previous studies [18, 20, 22, 67, 68] indicate that phenylalanine is the N-terminus of the mature toxin. It was unexpected that the Tox|Note BLASTx search did not identify a hit to δ-paraponeritoxin-Pc1a either, which might be due to the algorithm's generation of incomplete alignments in some cases that results in an incorrect annotation [84]. Curiously, Tox|Note then classified the majority of the δ-paraponeritoxin-Pc1e contig as a propeptide, perhaps because the Spider|ProHMM predictive tool used by Tox|Note to predict propeptide cleaving sites was designed to detect spider venom toxins. Most notably, the propeptide in δ-paraponeritoxin-Pc1e seems to be unusual and completely different to the propeptides reported in spiders and cone snails, where a conserved arginine is usually located at position -1 from the N-terminus of the mature peptide [112, 113]. Only once there is sufficient transcriptomic and proteomic data for ant venom peptides would it be possible to develop a similar script, capable of predicting propeptides in ants and perhaps other hymenopterans.

As seen in Figure 7.8, the alignment of *P. clavata* transcripts with other homologous δ-paraponeritoxins shows that the isoforms detected as part of this study were not identical to any of the previously described sequences, differing by one or two amino acids. This subtle difference between isoforms is usually responsible for changes in the activity and specificity of peptides. For example, natural mutations in specific regions of δ-paraponeritoxins, showed

that a decrease in net charge and hydrophobicity at the C-terminus decreased the activity of the toxin while an increase in net charge and hydrophobicity increased the activity [20]. These results highlight the fact that even a subtle change in sequence can cause a significant enhancement or reduction of the activity, and therefore correct assignment of function can only be determined by activity testing. The additional Gly and Lys residue at the end of the δ-paraponeritoxin-Pc1e transcript is likely an amidation signal that is cleaved during the maturation of the peptide [114]. However, comparison of the measured mass via MALDI-TOF MS from chapter 5, also collected in French Guiana (2783.4 Da), was consistent with the calculated mass (2783.6 Da) of the *P. clavata* δ-paraponeritoxin-Pc1e transcript, without the Gly and Lys residues at the C-terminus (2783.6 Da), indicating C-terminal cleavage of these residues and no amidation.

Although the algesic activity seen following *P. clavata* envenomation has been attributed to δ-paraponeritoxin, the peptidomic/ proteomic investigations described in Chapter 4 indicates that there are other proteins in this venom that could contribute to the pain-producing activity during envenomation. The holistic proteomic/ transcriptomic approach has increased the pool of proteins and identified proteins such as arginine kinase, which has not been previously identified and characterised in ant venoms. Arginine kinases are major components of wasp venoms where they have been associated with the induction of pain and have been shown to be paralytic to spiders in some cases [101, 115]. Similar to wasp venoms, *P. clavata* venom seems to have a significant amount of this protein as it was found both proteomically and with high expression levels in the transcriptome (TPM: 382). Importantly, there was a 97% percentage similarity of arginine-kinase-1a-P-clavata to the arginine kinase from the venom of the *Cyphononyx dorsalis,* the first venom in which arginine kinase was identified and found to be paralytic [116]. Additionally, all residues critical for arginine binding, catalysis, and ATP binding [101] were conserved in arginine-kinase-1a-P-clavata. It should also be noted that although this protein was detected proteomically as well and is found in the venom, no signal peptide was identified, suggesting the ORF is not complete despite the detection of a start (ATG) codon. This highlights that perhaps there is a start codon further upstream.

7.4.1.2 Proteases

Proteases are major components of venoms involved in digestion [106], hemostasis and thrombosis [103, 117] and have also been found to be allergenic as they bind to IgE [61, 118]. They are commonly found in the venoms of wasps [119], bees [106], ants [61] and snakes [117] where they have also been identified to act as factors that allow the spreading of other venom components [61]. In the present study, the major proteases identified were serine proteases which have been implicated in many biological processes in arthropods such as activation of the Toll signaling cascade and regulation of the polyphenoloxidase activation that affects immune defense mechanisms and inhibition of melanisation, [119] [120].

Serine proteases found in snake venoms have been implicated in thrombin-like activity, formation of fibrin clots, kininogenase activity and activation of coagulation factor V, platelets or plasminogen [117]. The alignment of the *P. clavata*'s serine protease sequence with other homologous proteins (see Figure 7.15) showed that all residues critical for catalytic activity (His37, Asp87, Ser190) corresponding to the catalytic triad [105] [106] are conserved, except in the case of isoform 8 indicating that the majority likely retain their catalytic activity. The presence of a glycine instead of a serine at position 190 in isoform 8 indicates that this serine protease may belong to the PPAF-II family of serine proteases that lack catalytic activity but still have the three disulfide bonds, referred to as a 'clip' domain [104] which have been found in solitary wasps and more recently in the *D. quadriceps* transcriptome [62, 119]. Interestingly, we also recovered two truncated versions of this same group of proteins that may, or may not, be functional. These truncations may also represent an active form of zymogen which has been previously reported in snakes and bees [106].

Nevertheless, since there is a diverse range of serine proteases, those found in *P. clavata* venom need to be tested for the type of activity they elicit in order to better understand their role in the venom of *P. clavata*. They might be similar to those of bee venoms which have been shown to have different activities in mammals and insects as it kills target insects via a melanisation strategy and exhibits fibrin(ogen)olytic activity in mammals [106]. The high number of transcripts assigned as serine proteases is unusual for ant venoms, as the previous three ant venom transcriptomes indicated they were minor venom components [61-63], However, this is not unusual, as bee venoms have been found to have different levels of serine proteases in different species of bees, with bumblebees containing more than honeybees [106].

The second most abundant protease was identified as a cathepsin-L, not previously reported in ants. Cathepsins are a large superfamily of lysosomal proteinases or endopeptidases [121]. Cathepsin L proteins have been described as important digestive enzymes in insect orders and have been targeted for pest control [122]. Cathepsin L is also involved in intra- and extra cellular protein degradation in many organisms such as bacteria and mammals with the first report in Hymenoptera by Laurino (2016) [119]. Our data is the first report in ants and second in Hymenopterans.

7.4.1.3 Phospholipases

One of the highest expressed enzymes in *P. clavata* venom was phospholipase A2 (PLA_2), also commonly seen in ant venoms [61-63] and consistent with the large number of spots identified as PLA_2 proteins in the 2D-PAGE data (see Chapter 5). Phospholipases are not exclusive to ants, and are found in a wide variety of animal venoms including scorpions [109], snakes [97, 123] and bees [92]. Amongst hymenopterans, the transcriptomics data from the ants *T. bicarinatum* and *O. monticola* have been reported to have low phospholipase expression, with levels being significantly higher in wasps [61]. Whilst that of *D. quadriceps* showed high expression, which is consistent with our data [61, 62].

To date, there are at least 15 classes of secreted PLA_2s that have been described in the literature including several subtypes within each class [91]. BLASTx searches revealed strong homology of two *P. clavata* transcripts to a PLA_2 found mainly in hymenopteran venoms (as shown in Figure 7.11). PLA_2s are enzymes that catalyse the hydrolysis of phospholipids at the sn-2 bond of the glycerol backbone resulting in the release of lysophospholipids and fatty acids that are involved in signal transduction and act as precursors for pharmacological mediators such as leukotrienes [124]. They have a variety of functions in venoms including antiplatelet, myotoxic, and neurotoxic activity in snakes [97] and allergenic and hemolytic activity in hymenopterans [61]. Pharmacological tests would be necessary to elucidate the activity and structure/function relationship of the PLA_2 contig we report here. Nevertheless, it can be assumed that this isoform is not myotoxic due to the presence of an aspartate residue instead of a lysine at position 117. *P.* clavata PLA_2 is also likely to be a group III phospholipase [93] which has been reported in other ant venom transcriptomes [61-63]. In support, group III PLA_2s contain the highly conserved active site residues: histidine-48 immediately preceding the calcium binding residue aspartate-49 as well as a conserved glycine-32 [92, 94].

7.4.1.4 Hyaluronidases

Hyaluronidases are a class of highly conserved proteins found in almost all venoms and aid in the spreading of other venom components by hydrolysing hyaluronic acid (HA), a key component of the extracellular matrix of vertebrates [125]. Eight transcripts from the *P. clavata* venom transcriptome had significant homology to hyaluronidase-like proteins from other ants and their expression was confirmed by the proteomics investigation. Alignment of these sequences with ClustalW confirmed that all residues critical for structure (four cysteines) and the active site (Asp135, Phe136, Glu137) were conserved [125]. Hyaluronidases are also considered a major allergen in venoms of bees and wasps, however it has been shown that not all species of hymenopterans have high expression of hyaluronidases [99, 126]. The high conservation across hyaluronidases may play a role in the cross-reactivity seen in allergic reactions to hymenopterans where patients can be allergic to both bee and wasp venoms [125] [126]. Ants, however, seem to have less hyaluronidase expression, with the highest level reported from the harvester ant *Pogonomyrmex* from which allergic reactions have been reported [127]. Interestingly, the transcripts with homology to known hyaluronidases in *P. clavata* were highly expressed*;* therefore they may induce similar activity in their victims, however, a biochemical characterisation is necessary to understand its true role as no allergic reactions have been reported from *P. clavata* envenomations [107].

7.4.1.5 Allergenic proteins

There were also hits to proteins such as icarapin and pilosulin 4, both of which are allergenic proteins found in ant venoms. Icarapin is a protein first isolated from the honeybee, *A. mellifera,* and is involved in IgE binding [100]. Recently, the transcriptome of the ant *O. monticola* also found evidence of icarapin in their combined peptidome/ transcriptome, which is the first report of this peptide in ant venoms. The high homology between the transcripts identified in this transcriptome with that of the bee *A. mellifera* suggests that it might have a similar role in this venom. Pilosulin 4, an ant venom allergenic peptide originally isolated from the ant *Myrmecia pilosula* was also identified in the *P. clavata* transcriptome (Supplementary file 3). However, the low TPM value of this transcript (7.64) suggests that this peptide does not play a major role in the venom of *P. clavata.* Therefore, it's role should still be investigated, as this peptide has been shown to be both allergenic and antibacterial [33, 128].

7.4.1.6 Metalloproteinases

Several metalloproteinase isoforms were identified in the venom of *P. clavata.* Metalloproteinases are proteases with fibrinolytic activity and the ability to degrade the extracellular matrix [117]. They are considered to be highly toxic in venoms as they can cause severe bleeding by disrupting blood coagulation [117]. A number of metalloproteinases have been reported in venoms of different hymenopteran [61, 119], snake [117] and scorpion venoms [109]. None of the transcripts identified in this study had a significant TPM count, which is consistent with the other ant transcriptome studies, which also did not find metalloproteinases amongst their highly abundant toxins [61-63]. As other metalloproteinases are associated with tissue degradation and immune suppression, we speculate a similar role in the venom of *P. clavata* [119].

7.4.1.7 Toxin-like Proteins

Several transcripts showed homology to toxins without a known activity such as U_8-agatoxin-Ao1a-like protein from the spider *Agelena orientalis* and a number of ω-conotoxin-like peptides (misnamed proteins) from other ants. These peptides are yet to be tested for pharmacological activity. Surprisingly, the majority of transcripts in this category had homology to ω-conotoxin-like peptides. However, alignment of these peptides with those from *P. clavata* (see Figure 7.10) and ω-conotoxin peptide from the marine cone snail C*onus lividus* revealed that there was very little homology to that of the cone snail aside from the cysteine framework (15% identity). Additionally, the cone snail ω-conotoxin had a different signal peptide as well as a propeptide which we were not able to predict due to lack of software and proteomic data. Omega conotoxins block calcium channels and result in analgesic effects [129, 130]. In fact, the related ω-conotoxin MVIIA (Ziconotide) is one of the few venom-derived toxins on the therapeutic drug market (Prialt®) [60]. However, the activity of the *P. clavata* transcript will need to be tested in order to identify its target, particularly because *P. clavata* is known for its ability to induce pain (algesic) not its analgesic effects [107, 108]. Therefore, the transcripts isolated from *P. clavata* have been named U_1-paraponeritoxin-Pc1a, based on the King (2008) nomenclature systems for toxins with unknown activity [86].

7.4.1.8 Insecticidal proteins

Three transcripts with homology to sphingomyelinase phosphodiesterase proteins (highest TPM: 27) were identified in this investigation. Sphingomelinase D is a spider venom toxin that has been identified as a potent dermonectrotic protein as well as an insecticidal protein, as it has a role in immobilising prey [131]. This is a highly interesting protein as this is the first report of this protein in a venom other than that of the *Loxosceles* spider [132]. This finding suggests that the insecticidal activity seen in *P. clavata* might not be due to poneratoxin alone, as previously suggested [21], but rather a combination of several insecticidal components.

7.4.1.9 Novel P. clavata toxins

One of the most interesting results of this study is the discovery of several novel peptides which we have identified in the venom of *P. clavata* with potential as novel bioactive molecules. Tox|Note identified hundreds of peptides as novel predicted toxins. Manual curation and inspection of these peptides was performed in order to select for transcripts with the following characteristics: number of cysteines ≥4, even number of cysteines, mature peptide length < 100 amino acids. This resulted in the identification of 190 peptides, as shown in Figure 7.16. Cysteine-rich peptides were of particular interest due to their high diversity, stability and likelihood of bioactivity [133]. Further analysis of these sequences revealed several peptides conforming to the inhibitor cysteine knot (ICK) framework which is a highly stable protein fold with a characteristic pseudo-knot formed by an α-helix, β-sheet or random coil connected by two disulfide bonds and interlaced with a third disulfide bond. ICK peptides are yet to be characterised in ant venoms, as their existence in ants was only recently identified through transcriptomics in the venom of *T. bicarinatum* by Torres (2014) who identified a single peptide conforming to this framework with low expression [62]. The present study identified six peptides with this framework, nevertheless, we also saw low expression levels, with TPM values ranging from 3 to 44, and no detection in the proteome. Other peptides conforming to this framework contained 8 and 10 cysteines. However, their cysteine bonding pattern was not predicted given the potential of multiple combinations of disulfide bonding, as the ICK motif is difficult to predict unless synthesised and checked against the native toxin.

One of the highly expressed toxin-like peptides identified by Tox|Note (TPM: 771) was a peptide conforming to conotoxin cysteine framework 1 also observed with α-conotoxins [134]. We found nine transcripts (Figure 7.17) with this framework which is the first report of these peptides in ant venoms. It would be interesting to test the activity of these toxins, as α-conotoxins are known for their activity on nicotinic acetylcholine receptors [130]. Medical applications of these peptides include potential use in neuropsychiatric disorders such as Parkinson's and Alzheimer's as this receptor has been implicated in their pathophysiology [134].

It must be noted that Tox|Note is a toxin prediction pipeline specifically designed for spider toxins, therefore all data was manually curated to ensure that the correct ORF translation and propeptide assignment was selected in all transcripts. Additionally, the Tox|Note output only reports ORFs with a predicted signal peptide due to the fact that signal peptides identify peptides for secretion, and reduces the number of predictions that need to be revised manually [84]. However, there might be other peptides that are secreted into the venom gland using other secretion systems or with signal peptides that are cleaved at a different site to that predicted by the SignalP algorithm. Furthermore, not all toxins contain signal peptides [112], therefore there might be other toxins in this transcriptome that Tox|Note has not identified.

7.4.2. Holistic proteomic/ transcriptomic investigation

In this study, we present the first combined proteomic and transcriptomic investigation of an ant venom gland. The identification of individual peptides from complex mixtures such as venoms remains a challenge despite recent advances in technologies, in fact, there is no single method capable of identifying all components of a given species. As mentioned above, RNAseq alone would not have detected poneratoxin, which was identified through proteomics. Several groups are now moving to a combined proteomic (indicates what is found in the venom)/ transcriptomic (reveals what is expressed in the venom gland) methodology in an attempt to overcome this issue [63, 110]. Unsurprisingly, minimal overlap has often been obtained between the two techniques [110]. This was also the case in our investigation, as only 44 of the 453 toxin proteins identified in the transcriptome were identified in the proteome. However, this may be a taxa-dependent issue or a technical issue as ~90% overlap was obtained between the peptidome and transcriptome of the cone snail *Conus marmoreous* when using triple quadrupole mass spectrometry and high-depth pyrosequencing [135]. Therefore, in future high resolution mass spectrometry coupled to high resolution

chromatography should be employed to try and increase this overlap as it has been suggested that the low overlap might be due to: i) the inherent limitations of either technique; ii) the fact that venoms are a mixture of hundreds of components often with similar molecular size and isoelectric point, including closely related isoforms with different expression levels, which makes separation difficult, and/or (iii) the many potential post-translational modifications (PTMs) in the peptides/ proteins.

Nevertheless, there are many advantages to this holistic approach, as it allows for confirmation of expression of transcripts found within the transcriptome. It also allows for better protein identification by proteomics with a bottom-up approach. This is exemplified by the present study detecting four times the number of proteins compared to our previously published *P. clavata* proteome [71] where we found a total of 95 proteins overall, whereas here we report 495 using the shotgun approach. There are many reasons for this, such as the limits of mass spectrometry, as we did use a more sensitive mass spectrometer in this investigation compared to the previous. It is also noteworthy that according to the transcriptome, there should be over 54,000 expressed proteins with differing expression levels. However, a transcriptome is representative of transcripts being transcribed at the point of time that they were collected, and one cannot assume that all are expressed at the protein level. Nevertheless, there is no clear explanation as to why the correlation remains so imperfect, and more work needs to be done in order to improve the match between venom proteome and transcriptome.

This investigation highlighted the enormous lack of tools for ant venom transcriptome assembly as reconstructing a comprehensive transcriptome from a non-model organism with no genome, using short reads, is particularly difficult. For example, there was an issue in finding a translation tool that was free of translation errors, as there were several instances in which the programs selected the wrong ORF as the 'best' ORF which is defined as the longest ORF with a stop and start codon. This was also the case with TransDecoder, which was used in the present study, however, this was overcome by double-checking the ORF of any contig that was annotated using the online ORF finder which can be used for single proteins or peptides (https://www.ncbi.nlm.nih.gov/orffinder/). An additional error we found through orffinder was in the ORF itself, was that the start site was actually several amino acids upstream from the one selected by TransDecoder. Another issue encountered was when searching the NCBInr database using BLASTx, as the large size of the database led to the need for access to a high performance computing cluster, which still crashed. This issue was eventually resolved by

separation of the transcript file into several smaller ones and querying them separately then concatenating the results together. This is not an uncommon issue, as it is recognised as one of the limiting factors of RNAseq data analysis [136]. Although BLASTx required more time compared to a more curated database such as that of UniProtKB, it was a necessary task as we were interested in identifying novel peptides and proteins, which could have been false positives using the UniProtKB database.

7.5. Conclusion and future directions

This chapter presents the first holistic investigation of an ant venom gland using a combined proteomic and transcriptomic approach. This was performed using the non-model ant species *P. clavata* (no genome and the only ant in the subfamily) with a focus on the toxin proteins and peptides which would be responsible for the extreme pain elicited by this ant. Illumina sequencing resulted in an in-depth coverage of the venom gland components and revealed which sequences are involved in cellular processes or are putative venom compounds. We also confirmed δ-paraponeritoxin (formerly poneratoxin) as the major toxin in this venom along with the presence of high levels of PLA_2. Additionally, δ-paraponeritoxin-Pc1e is a novel homologue not previously identified in the venoms of *P. clavata* from Peru, Brazil, Panama or Costa Rica. Aside from δ-paraponeritoxin-Pc1e and the insecticidal sphingomyelinase, the other toxins seen in this venom are consistent with what is found in other ants and hymenopterans. This is supported by the cross-reactivity in allergic reactions due to highly conserved allergenic proteins amongst Hymenoptera. A small fraction of the transcriptome was composed of novel ICK-like peptides and peptides conforming to conotoxin framework I that could harbour potential therapeutic or bioinsecticidal bioactivities. Taken together, this work greatly expands the current knowledge on the molecular basis of the *P. clavata* sting and other ant venoms. In addition, we have confirmed the complexity of the venom protein and peptide pool, which could be exploited for potential therapeutic drugs or insecticides.

Future work should investigate the structure and function of the different toxin sequences identified in this study. This can be achieved by initially synthesising the toxins through either recombinant expression of proteins or solid phase synthesis of shorter peptides and subsequent pharmacological screening to determine the target and mode of action. Bioactive peptides with cysteine-stabilised structures are ideal as leads for the development of novel pharmaceutical drugs or bioinsecticides. Understanding the effects of these venom toxins will help develop strategies to neutralise the toxicity of this, and other, ant venoms. A positive step towards this is the recent initiative by the Global Ant Genomics Alliance (GAGA), which aims to sequence 200 genomes by the year 2020, one of which is that of *P. clavata* [65]. This will not only be useful in the case of *P. clavata* but will also aid in the transcriptome assembly of other ant venom glands in the future, helping us understand one of the most interesting groups of animals on the planet.

Chapter Eight

Concluding Remarks and Future Directions

Chapter Eight: Concluding Remarks and Future Directions

Ant venom is a rich source of peptide and protein toxins that have evolved over millions of years. Selection pressure over these years has resulted in toxins that have been optimised for their highly potent and selective effects when predating or protecting their colonies from aggressors. Ant venom is not usually lethal to humans, as the main prey of ants is insects; however, fatalities have been reported in people with allergies to ant venom. Interestingly, individuals with allergies to bee venom are also likely to have allergies to ant venom as there is a high degree of homology between proteins of these two venoms. These include hyaluronidases and phospholipases which have both been implicated as allergens. The similarity between these two hymenopteran venoms has been reported for several decades and was also observed throughout this thesis. However, despite this, there has been little interest in the investigation and understanding of the venoms of these insects which are considered to be one of the most successful animals in many ecosystems. The main reason for this is their small size which makes their investigation more challenging given the low venom yield, though this is now a hurdle that can be overcome due to the advancements in the sensitivity and accuracy of analytical technologies.

Another reason for the lack of studies on ant venoms is the misconception that these venoms are not rich in bioactive peptides [17]. Therefore, the aims of this investigation was to (i) confirm the insecticidal and antibacterial activity of the ant venoms used in this project, then to (ii) determine the best collection method of ant venoms as this has not been done previously. Due to lack of sufficient samples, insecticidal assays were performed on 4 of the 7 ants investigated in this project and antibacterial activity was determined for 2 of these ants. This led to the establishment that ant venoms do indeed contain insecticidal properties, as 2 out of the 4 ants did have insecticidal activity and one of the venoms which didn't have insecticidal activity had considerable antibacterial activity. The next step would be to conduct a bioassay-guided fractionation of the venoms to identify the peptide(s) responsible for these activities. It is also likely that the activity of the isolated peptide will be higher than that of the crude venom, where there are several other components masking the activity. This would involve initially fractionating with strong cation exchange chromatography, testing the fraction for antibacterial or insecticidal activity then further fractionating using an orthogonal method such as RP-HPLC.

The next aim of the project was to determine the best collection method for ant venom. This was of even more importance following the discovery of the amounts of venom required for the bioassays detailed above. The venom collection methods tested were electrical stimulation and venom gland dissection. It was found that although venom gland dissection is able to detect larger numbers of peptides, the electrical stimulation also yielded several other peptides. Surprisingly, there was very little overlap in the venom peptidomes using the two collection methods. This may suggest that ant venoms do not secrete the entirety of their venom gland components but could also be the result of ion suppression of peptides from the dissected venom by contaminant proteins from cellular structures. This would mask any peptides with lower levels of abundance peptides seen in venom collected by electrical stimulation. This led to the conclusion that, for the purposes of venom characterisation, venom gland dissection is the better option of collection as it would offer a better representation of the venom gland components. However, in cases where the desired peptide/ protein is known then electrical stimulation would be the better option as it doesn't kill the ant and it is also an easier method of collection. One such method includes the bee venom electric board (http://www.beewhisper.com/). This is able to extract large quantities of crude venom and is able to be carried to the field. This removes the need for dissecting the ants under a microscope in the lab after freezing them. It would also be interesting to perform imaging mass spectrometry on the venom gland to identify patterns of peptide distribution within the venom gland. In the case of cone snails and centipedes, the use of venom in a defensive or offensive strategy it has been shown to result in a different venom composition [137].

Another interesting finding uncovered during the *P. clavata* venom collection investigation was the intraspecies peptide diversity. It was found that specimens of *P. clavata* collected from different nests had vastly different peptidomes with hardly any overlapping peptides indicating that ant venoms are indeed highly diverse and suggestive that they are dependent on a range of biotic and abiotic conditions such as microclimate and local prey and predator species. This led to the third aim of the project, which was to perform a deep investigation of the peptidomes and proteomes of venoms from different ant clades to establish the diversity of the venom components. This resulted in highly surprising results, that ant venoms are extremely diverse with hundreds of peptides per venom and the majority of these being unique to each species. Overall, there seemed to be a pattern of low molecular weight peptides and higher peptide counts in poneroids compared to the formicoids which had lower peptide counts and a wider mass distribution. However, closer analysis showed that there was

little overlap between the peptides of species from the same clade, further highlighting that ant venoms represent a large pool of potential bioactive peptides.

All the ant venom proteomes investigated showed similar profiles when the functions of the proteins were compared. This is consistent with other hymenopteran venom proteins which seem to be similar across this insect Order. This is not surprising, as cross-reactivity has been reported in allergic reactions to hymenopteran venoms and has been associated with the high conservation of proteins between hymenopterans. One issue that was common to all the proteomics experiments performed in this study was that many of the 2D-PAGE gel spots were not matched to any known proteins. This was due to the lack of ant venom peptide/ protein sequence data in the various nucleotide and protein databases, and highlights the need for continued investigation of ant venoms to help build this library.

As indicated above, the lack of ant venom peptide/ protein amino acid sequence data in databases compromised the ability to perform bottom-up sequencing of our MS/MS data. In order to help resolve this issue, a transcriptomic investigation was performed on the ant *P. clavata* which is known to have the most painful insect stings. This was performed in conjunction with a proteomics investigation to have a complimentary set of data and to aid in the annotation of the venom components. This approach was very successful with the number of characterised *P. clavata* venom proteins increasing from our previously reported number of 95 characterised proteins to 438 proteins. The next step for this investigation, which has already commenced with our collaborators, is to characterise the pharmacological activity of these peptides by initially synthesising them using either solid phase extraction or expression in bacteria, then performing the assays to determine their activity.

Overall, this project was successful in identifying the insecticidal activity of four ant venoms, the antibacterial activity of two ant venoms and found that the collection method of venom does affect the venom profile. In addition, it has resulted in the peptidomic characterisation of seven ant species and proteomic characterisation of three ant species, all of which have not been previously reported in this level of detail. Moreover, we present the first holistic investigation of an ant venom using a combined proteomic and transcriptomic investigation. This has confirmed that ants are a viable candidate for investigation as sources of insecticidal and antibacterial peptides for development into commercial products.

Supplementary Material

Disclaimer: these files are only relate to chapter seven and due to the size the supplementary files, they are too large to be inserted into the thesis and are therefore available on the USB drive provided. A PDF version of the thesis is also included. The following materials can be found:

- **Supplementary file 1:** a word document contacting the commands used in the Trinity assembly and QC of the transcriptome.

- **Supplementary file 2:** Figure of the FastQC report

- **Supplementary file 3:** an excel workbook containing all the toxins identified by BLASTx along with the TPM values. These were used to construct Figure 7.6

- **Supplementary file 4:** an excel worksheet containing all protein hits from the proteome investigation using the transcriptome database.

- **PDF file of the Thesis**

References

1. FAO I: **WFP. 2015**. *The state of food insecurity in the world* 2015.
2. Sallam MN: **Insect damage: damage on post-harvest**. *AGSI/FAO: INPhO Available via* http://wwwfaoorg/inpho/content/compend/text/ch02-01htm 2008, **22**.
3. Brogdon WG, McAllister JC: **Insecticide resistance and vector control**. *Emerg Infect Dis* 1998, **4**(4):605-613.
4. Brouqui P, Parola P, Raoult D: **Insecticide resistance in mosquitoes and failure of malaria control**. *Expert Rev Anti Infect Ther* 2012, **10**(12):1379-1381.
5. Gerland P, Raftery AE, Ševčíková H, Li N, Gu D, Spoorenberg T, Alkema L, Fosdick BK, Chunn J, Lalic N: **World population stabilization unlikely this century**. *Science* 2014, **346**(6206):234-237.
6. Pittendrigh B, Vaughan A, Anthony N: **Why are there so few resistance–associated mutations in insecticide target genes?** *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 1998, **353**(1376):1685-1693.
7. Hemingway J, Ranson H: **Insecticide resistance in insect vectors of human disease**. *Annu Rev Entomol* 2000, **45**(1):371-391.
8. Liu N: **Insecticide resistance in mosquitoes: impact, mechanisms, and research directions**. *Annu Rev Entomol* 2015, **60**:537-559.
9. Czaja K, Góralczyk K, Struciński P, Hernik A, Korcz W, Minorczyk M, Łyczewska M, Ludwicki JK: **Biopesticides–towards increased consumer safety in the European Union**. *Pest Management Science* 2015, **71**(1):3-6.
10. Copping LG, Menn JJ: **Biopesticides: a review of their action, applications and efficacy**. *Pest Management Science* 2000, **56**:651-676.
11. Lucas J: **Advances in plant disease and pest management**. *The Journal of Agricultural Science* 2011, **149**(S1):91-114.
12. Wang C, St Leger RJ: **A scorpion neurotoxin increases the potency of a fungal insecticide**. *Nat Biotechnol* 2007, **25**(12):1455-1456.
13. Zeng X-Z, Xiao Q-B, Liang S-P: **Purification and characterization of raventoxin-I and raventoxin-III, two neurotoxic peptides from the venom of the spider *Macrothele raveni***. *Toxicon* 2003, **41**(6):651-656.
14. de Figueiredo SG, de Lima ME, Cordeiro MN, Diniz CR, Patten D, Halliwell RF, Gilroy J, Richardson M: **Purification and amino acid sequence of a highly insecticidal toxin from the venom of the Brazilian spider *Phoneutria nigriventer* which inhibits NMDA-evoked currents in rat hippocampal neurones**. *Toxicon* 2001, **39**(2):309-317.
15. Matsushita N, Miyashita M, Sakai A, Nakagawa Y, Miyagawa H: **Purification and characterization of a novel short-chain insecticidal toxin with two disulfide bridges from the venom of the scorpion *Liocheles australasiae***. *Toxicon* 2007, **50**(6):861-867.
16. McGain F, Harrison J, Winkel KD: **Wasp sting mortality in Australia**. *Med J Aust* 2000, **173**(4):198-200.
17. Aili SR, Touchard A, Escoubas P, Padula MP, Orivel J, Dejean A, Nicholson GM: **Diversity of peptide toxins from stinging ant venoms**. *Toxicon* 2014, **92**:166-178.
18. Piek T: **Neurotoxins from venoms of the Hymenoptera-twenty-five years of research in Amsterdam**. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology and Toxicology* 1990, **96**(2):223-233.
19. Tene N, Bonnafe E, Berger F, Rifflet A, Guilhaudis L, Segalas-Milazzo I, Pipy B, Coste A, Leprince J, Treilhou M: **Biochemical and biophysical combined study of bicarinalin, an ant venom antimicrobial peptide**. *Peptides* 2016, **79**:103-113.
20. Johnson SR, Rikli HG, Schmidt JO, Evans MS: **A reexamination of poneratoxin from the venom of the bullet ant *Paraponera clavata***. *Peptides* 2017, **98**:51-62.

21. Szolajska E, Poznanski J, Ferber ML, Michalik J, Gout E, Fender P, Bailly I, Dublet B, Chroboczek J: **Poneratoxin, a neurotoxin from ant venom. Structure and expression in insect cells and construction of a bio-insecticide**. *Eur J Biochem* 2004, **271**(11):2127-2136.
22. Piek T, Duval A, Hue B, Karst H, Lapied B, Mantel P, Nakajima T, Pelhate M, Schmidt JO: **Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata***. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology and Toxicology* 1991, **99**(3):487-495.
23. Pluzhnikov K, Nolde D, Tertyshnikova S, Sukhanov S, Sobol A, Torgov M, Filippov A, Arseniev A, Grishin E: **Structural and functional studies of toxic principle of *Ectatomma tuberculatum* ant venom**. *Bioorg Khim* 1994, **20**(8-9):857-871 (In Russian).
24. Pluzhnikov K, Nosyreva E, Shevchenko L, Kokoz Y, Schmalz D, Hucho F, Grishin E: **Analysis of ectatomin action on cell membranes**. *Eur J Biochem* 1999, **262**(2):501-506.
25. Nolde DE, Sobol AG, Pluzhnikov KA, Grishin EV, Arseniev AS: **Three-dimensional structure of ectatomin from *Ectatomma tuberculatum* ant venom**. *J Biomol NMR* 1995, **5**(1):1-13.
26. Arseniev A, Pluzhnikov K, Nolde D, Sobol A, Torgov MY, Sukhanov S, Grishin E: **Toxic principle of selva ant venom is a pore-forming protein transformer**. *FEBS Lett* 1994, **347**(2):112-116.
27. Davies NW, Wiese MD, Brown SG: **Characterisation of major peptides in 'jack jumper'ant venom by mass spectrometry**. *Toxicon* 2004, **43**(2):173-183.
28. Ebaid H, Al-Khalifa M, Isa AM, Gadoa S: **Bioactivity of Samsum ant (*Pachycondyla sennaarensis*) venom against lipopolysaccharides through antioxidant and upregulation of Akt1 signaling in rats**. *Lipids Health Dis* 2012, **11**(1):93.
29. Johnson SR, Copello JA, Evans MS, Suarez AV: **A biochemical characterization of the major peptides from the venom of the giant neotropical hunting ant *Dinoponera australis***. *Toxicon* 2010, **55**(4):702-710.
30. Warren WC, Hillier LW, Graves JAM, Birney E, Ponting CP, Grützner F, Belov K, Miller W, Clarke L, Chinwalla AT: **Genome analysis of the platypus reveals unique signatures of evolution**. *Nature* 2008, **453**(7192):175-183.
31. King GF, Hardy MC: **Spider-venom peptides: structure, pharmacology, and potential for control of insect pests**. *Annu Rev Entomol* 2013, **58**(1):475-496.
32. Orivel J, Redeker V, Le Caer JP, Krier F, Revol-Junelles AM, Longeon A, Chaffotte A, Dejean A, Rossier J: **Ponericins, new antibacterial and insecticidal peptides from the venom of the ant *Pachycondyla goeldii***. *The Journal of Biological Chemistry* 2001, **276**(21):17823-17829.
33. Inagaki H, Akagi M, Imai HT, Taylor RW, Kubo T: **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* 2004, **428**(2):170-178.
34. Lima D, Torres A, Mello C, Menezes RRPPB, Sampaio T, Canuto J, Silva J, Freire V, Quinet Y, Havt A: **Antimicrobial effect of *Dinoponera quadriceps* (Hymenoptera: Formicidae) venom against *Staphylococcus aureus* strains**. *J Appl Microbiol* 2014, **117**(2):390-396.
35. Pluzhnikov KA, Kozlov SA, Vassilevski AA, Vorontsova OV, Feofanov AV, Grishin EV: **Linear antimicrobial peptides from *Ectatomma quadridens* ant venom**. *Biochimie* 2014, **107 Pt B**:211-215.
36. Rifflet A, Gavalda S, Téné N, Orivel J, Leprince J, Guilhaudis L, Génin E, Vétillard A, Treilhou M: **Identification and characterization of a novel antimicrobial peptide from the venom of the ant *Tetramorium bicarinatum***. *Peptides* 2012, **38**(2):363-370.
37. Viljakainen L, Pamilo P: **Selection on an antimicrobial peptide defensin in ants**. *J Mol Evol* 2008, **67**(6):643-652.

38. Zhang Z, Zhu S: **Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species**. *Dev Comp Immunol* 2012, **38**(2):262-274.
39. Fuda C, Fisher J, Mobashery S: **β-Lactam resistance in *Staphylococcus aureus*: the adaptive resistance of a plastic genome**. *Cell Mol Life Sci* 2005, **62**(22):2617.
40. Korzeniewska E, Korzeniewska A, Harnisz M: **Antibiotic resistant *Escherichia coli* in hospital and municipal sewage and their emission to the environment**. *Ecotoxicol Environ Saf* 2013, **91**:96-102.
41. Organization WH: **Antimicrobial resistance: global report on surveillance**: World Health Organization; 2014.
42. Oguiura N, Boni-Mitake M, Affonso R, Zhang G: **In vitro antibacterial and hemolytic activities of crotamine, a small basic myotoxin from rattlesnake *Crotalus durissus***. *J Antibiot* 2011, **64**(4):327.
43. Smith PK, Krohn RI, Hermanson G, Mallia A, Gartner F, Provenzano M, Fujimoto E, Goeke N, Olson B, Klenk D: **Measurement of protein using bicinchoninic acid**. *Anal Biochem* 1985, **150**(1):76-85.
44. Xiang F, Beavis RC, Ens W: **A method to increase contaminant tolerance in protein matrix-assisted laser desorption/ionization by the fabrication of thin protein-doped polycrystalline films**. *Rapid Commun Mass Spectrom* 1994, **8**(2):199-204.
45. Escoubas P, Sollod B, King GF: **Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach**. *Toxicon* 2006, **47**(6):650-663.
46. Lipkin A, Kozlov S, Nosyreva E, Blake A, Windass J, Grishin E: **Novel insecticidal toxins from the venom of the spider *Segestria florentina***. *Toxicon* 2002, **40**(2):125-130.
47. Maggio F, King GF: **Role of the structurally disordered N- and C-terminal residues in the Janus-faced atracotoxins**. *Toxicon* 2002, **40**(9):1355-1361.
48. Wang X, Smith R, Fletcher JI, Wilson H, Wood CJ, Howden ME, King GF: **Structure-function studies of omega-atracotoxin, a potent antagonist of insect voltage-gated calcium channels**. *Eur J Biochem* 1999, **264**(2):488-494.
49. Burkhardt C, Fairchild M: **Toxicity of insecticides to house crickets and bioassay of treated soils in the laboratory**. *J Econ Entomol* 1967, **60**(6):1496-1503.
50. Manzoli-Palma M, Gobbi N, Palma M: **Insects as biological models to assay spider and scorpion venom toxicity**. *J Venom Anim Toxins Incl Trop Dis* 2003, **9**(2):174-185.
51. Balaji RA, Sasaki T, Gopalakrishnakone P, Sato K, Kini RM, Bay B-H: **Purification, structure determination and synthesis of covalitoxin-II, a short insect-specific neurotoxic peptide from the venom of the *Coremiocnemis validus* (Singapore tarantula)**. *FEBS Lett* 2000, **474**(2-3):208-212.
52. Gentz MC, Jones A, Clement H, King GF: **Comparison of the peptidome and insecticidal activity of venom from a taxonomically diverse group of theraphosid spiders**. *Toxicon* 2009, **53**(5):496-502.
53. Dejean A: **Prey capture behavior in an arboreal African ponerine ant**. *PLoS One* 2011, **6**(5):e19837.
54. Mill AE: **Predation by the ponerine ant *Pachycondyla commutata* on termites of the genus Syntermes in Amazonian rain forest**. *Journal of Natural History* 1984, **18**(3):405-410.
55. Aili SR, Touchard A, Koh JMS, Dejean A, Orivel J, Padula MP, Escoubas P, Nicholson GM: **Comparisons of protein and peptide complexity in poneroid and formicoid ant venoms**. *J Proteome Res* 2016, **15**(9):3039-3054.
56. de Melo ET, Estrela AB, Santos ECG, Machado PRL, Farias KJS, Torres TM, Carvalho E, Lima JPMS, Silva-Júnior AA, Barbosa EG *et al*: **Structural characterization of a novel peptide with antimicrobial activity from the venom gland of the scorpion *Tityus stigmurus*: Stigmurin**. *Peptides* 2015, **68**:3-10.

57. Corzo G, Escoubas P, Villegas E, Barnham KJ, He W, Norton RS, Nakajima T: **Characterization of unique amphipathic antimicrobial peptides from venom of the scorpion *Pandinus imperator***. *Biochem J* 2001, **359**(Pt 1):35-45.
58. Han SM, Kim JM, Hong IP, Woo SO, Kim SG, Jang HR, Pak SC: **Antibacterial activity and antibiotic-enhancing effects of honeybee venom against methicillin-resistant *staphylococcus aureus***. *Molecules* 2016, **21**(1):79.
59. Madio B, Undheim EA, King GF: **Revisiting venom of the sea anemone *Stichodactyla haddoni*: Omics techniques reveal the complete toxin arsenal of a well-studied sea anemone genus**. *J Proteomics* 2017, **166**:83-92.
60. Robinson SD, Undheim EA, Ueberheide B, King GF: **Venom peptides as therapeutics: advances, challenges and the future of venom-peptide discovery**. *Expert review of proteomics* 2017, **14**(10):931-939.
61. Bouzid W, Verdenaud M, Klopp C, Ducancel F, Noirot C, Vétillard A: **De Novo sequencing and transcriptome analysis for *Tetramorium bicarinatum*: a comprehensive venom gland transcriptome analysis from an ant species**. *BMC Genomics* 2014, **15**(1):987.
62. Torres AFC, Huang C, Chong C-M, Leung SW, Prieto-da-Silva ARB, Havt A, Quinet YP, Martins AMC, Lee SMY, Rádis-Baptista G: **Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: insights into the polypeptide toxin arsenal of Hymenopterans**. *PLoS One* 2014, **9**(1):e87556.
63. Kazuma K, Masuko K, Konno K, Inagaki H: **Combined Venom Gland Transcriptomic and Venom Peptidomic Analysis of the Predatory Ant *Odontomachus monticola***. *Toxins (Basel)* 2017, **9**(10):323.
64. Smith CR, Dolezal A, Eliyahu D, Holbrook CT, Gadau J: **Ants (Formicidae): models for social complexity**. *Cold Spring Harbor protocols* 2009, **2009**(7):pdb.emo125.
65. Boomsma JJ, Brady SG, Dunn RR, Gadau J, Heinze J, Keller L, Sanders NJ, Schrader L, Schultz TR, SundstrÖM L *et al*: **The Global Ant Genomics Alliance (GAGA)**. *Myrmecological News* 2017, **25**:61-66.
66. Cavill GW, Robertson PL, Whitfield FB: **Venom and venom apparatus of the bull ant, *Myrmecia gulosa* (Fabr.)**. *Science* 1964, **146**(3640):79-80.
67. Duval A, Malécot CO, Pelhate M, Piek T: **Poneratoxin, a new toxin from an ant venom, reveals an interconversion between two gating modes of the Na channels in frog skeletal muscle fibres**. *Pflügers Archiv* 1992, **420**(3-4):239-247.
68. Piek T, Hue B, Mantel P, Nakajima T, Schmidt JO: **Pharmacological characterization and chemical fractionation of the venom of the ponerine ant, *Paraponera clavata* (F.)**. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology and Toxicology* 1991, **99**(3):481-486.
69. Hendrich AB, Mozrzymas JW, Konopińska D, Scuka M: **The effect of poneratoxin on neuromuscular transmission in the rat diaphragm**. *Cell Mol Biol Lett* 2001, **7**(2):195-202.
70. Rykaczewska-Czerwinska M, Radosz A, Konopinska D, Wrobel M, Plech A: **Antinociceptive effect of poneratoxin [PoTX] in rats**. *Pestycydy* 2008(1-2):135-141.
71. Aili SR, Touchard A, Petitclerc F, Dejean A, Orivel J, Padula MP, Escoubas P, Nicholson GM: **Combined peptidomic and proteomic analysis of electrically stimulated and manually dissected venom from the south American bullet ant *Paraponera clavata***. *J Proteome Res* 2017, **16**(3):1339-1351.
72. Lader ES: **Methods and reagents for preserving rna in cell and tissue samples**. In.: Google Patents; 2012.
73. Andrews S: **FastQC: A Quality Control Tool for High Throughput Sequence Data (Babraham Bioinformatics, Cambridge, UK)**. 2011.

74. Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B: **R obi NA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics**. *Nucleic Acids Res* 2012, **40**(W1):W622-W627.
75. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M: ***De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis**. *Nat Protoc* 2013, **8**(8):1494-1512.
76. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2**. *Nature Methods* 2012, **9**(4):357-359.
77. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *BMC Bioinformatics* 2011, **12**(1):323.
78. Consortium U: **The universal protein resource (UniProt)**. *Nucleic Acids Res* 2008, **36**(suppl 1):D190-D195.
79. Huntley RP, Sawford T, Mutowo-Meullenet P, Shypitsyna A, Bonilla C, Martin MJ, O'Donovan C: **The GOA database: Gene Ontology annotation updates for 2015**. *Nucleic Acids Res* 2015, **43**(Database issue):D1057-D1063.
80. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson Da, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al*: **Full-length transcriptome assembly from RNA-Seq data without a reference genome**. *Nat Biotechnol* 2011, **29**.
81. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, Zhang W, Zhang Z, Lajoie GA, Ma B: **PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification**. *Mol Cell Proteomics* 2012, **11**(4):M111.010587.
82. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25**(17):3389-3402.
83. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nature Methods* 2011, **8**(10):785-786.
84. Pineda SS, Chaumeil P-A, Kunert A, Kaas Q, Thang MW, Li L, Nuhn M, Herzig V, Saez NJ, Cristofori-Armstrong B: **ArachnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins**. *Bioinformatics* 2017.
85. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets**. *Mol Biol Evol* 2016, **33**(7):1870-1874.
86. King GF, Gentz MC, Escoubas P, Nicholson GM: **A rational nomenclature for naming peptide toxins from spiders and other venomous animals**. *Toxicon* 2008, **52**(2):264-276.
87. Touchard A, Aili S, Fox E, Escoubas P, Orivel J, Nicholson G, Dejean A: **The biochemical toxin arsenal from ant venoms**. *Toxins (Basel)* 2016, **8**(1):30.
88. Sade YB, Bóia-Ferreira M, Gremski LH, da Silveira RB, Gremski W, Senff-Ribeiro A, Chaim OM, Veiga SS: **Molecular cloning, heterologous expression and functional characterization of a novel translationally-controlled tumor protein (TCTP) family member from *Loxosceles intermedia* (brown spider) venom**. *The International Journal of Biochemistry & Cell Biology* 2012, **44**(1):170-177.
89. Graudins A, Little MJ, Pineda SS, Hains PG, King GF, Broady KW, Nicholson GM: **Cloning and activity of a novel alpha-latrotoxin from red-back spider venom**. *Biochem Pharmacol* 2012, **83**(1):170-183.
90. de LM Junqueira-de-Azevedo I, Pertinhez T, Spisni A, Carreño FR, Farah CS, Ho PL: **Cloning and expression of calglandulin, a new EF-hand protein from the venom glands of *Bothrops insularis* snake in E. coli**. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 2003, **1648**(1):90-98.
91. Gutiérrez JM, Lomonte B: **Phospholipases A2: Unveiling the secrets of a functionally versatile group of snake venom toxins**. *Toxicon* 2013, **62**(Supplement C):27-39.

92. Scott DL, Otwinowski Z, Gelb MH, Sigler PB: **Crystal structure of bee-venom phospholipase A2 in a complex with a transition-state analogue**. *Science* 1990, **250**(4987):1563-1567.
93. Six DA, Dennis EA: **The expanding superfamily of phospholipase A 2 enzymes: classification and characterization**. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 2000, **1488**(1):1-19.
94. Valentin E, Ghomashchi F, Gelb MH, Lazdunski M, Lambeau G: **Novel human secreted phospholipase A2 with homology to the group III bee venom enzyme**. *J Biol Chem* 2000, **275**(11):7492-7496.
95. Kuchler K, Gmachl M, Sippl MJ, Kreil G: **Analysis of the cDNA for phospholipase A2 from honeybee venom glands**. *Eur J Biochem* 1989, **184**(1):249-254.
96. White SP, Scott DL, Otwinowski Z, Gelb MH, Sigler PB: **Crystal Structure of Cobra-Venom Phospholipase A (2) in a Complex with a Transition-State Analogue**. *Science* 1990, **250**(4987):1560.
97. Kini RM: **Structure–function relationships and mechanism of anticoagulant phospholipase A 2 enzymes from snake venoms**. *Toxicon* 2005, **45**(8):1147-1161.
98. Chioato L, Ward RJ: **Mapping structural determinants of biological activities in snake venom phospholipases A 2 by sequence analysis and site directed mutagenesis**. *Toxicon* 2003, **42**(8):869-883.
99. Marković-Housley Z, Miglierini G, Soldatova L, Rizkallah PJ, Müller U, Schirmer T: **Crystal structure of hyaluronidase, a major allergen of bee venom**. *Structure* 2000, **8**(10):1025-1035.
100. Peiren N, de Graaf DC, Brunain M, Bridts CH, Ebo DG, Stevens WJ, Jacobs FJ: **Molecular cloning and expression of icarapin, a novel IgE-binding bee venom protein**. *FEBS Lett* 2006, **580**(20):4895-4899.
101. Baek JH, Lee SH: **Identification and characterization of venom proteins of two solitary wasps, *Eumenes pomiformis* and *Orancistrocerus drewseni***. *Toxicon* 2010, **56**(4):554-562.
102. Strong SJ, Ellington WR: **Isolation and sequence analysis of the gene for arginine kinase from the chelicerate arthropod, *Limulus polyphemus*: insights into catalytically important residues**. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1995, **1246**(2):197-200.
103. Jiang H, Kanost MR: **The clip-domain family of serine proteinases in arthropods**. *Insect Biochem Mol Biol* 2000, **30**(2):95-105.
104. Piao S, Song YL, Kim JH, Park SY, Park JW, Lee BL, Oh BH, Ha NC: **Crystal structure of a clip-domain serine protease and functional roles of the clip domains**. *The EMBO journal* 2005, **24**(24):4404-4414.
105. Ross J, Jiang H, Kanost MR, Wang Y: **Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships**. *Gene* 2003, **304**:117-131.
106. Choo YM, Lee KS, Yoon HJ, Kim BY, Sohn MR, Roh JY, Je YH, Kim NJ, Kim I, Woo SD: **Dual function of a bee venom serine protease: prophenoloxidase-activating factor in arthropods and fibrin (ogen) olytic enzyme in mammals**. *PLoS One* 2010, **5**(5):e10393.
107. Freeman TM, Tracy JM: **Stinging Insect Allergy: A Clinician's Guide**: Springer; 2017.
108. Schmidt JO: **The sting of the wild**: JHU Press; 2016.
109. Luna-Ramírez K, Quintero-Hernández V, Juárez-González VR, Possani LD: **Whole transcriptome of the venom gland from *Urodacus yaschenkoi* scorpion**. *PLoS One* 2015, **10**(5):e0127883.
110. Ponce D, Brinkman DL, Potriquet J, Mulvenna J: **Tentacle transcriptome and venom proteome of the pacific sea nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa)**. *Toxins (Basel)* 2016, **8**(4):102.

111. Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, Hagen DE: **Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine**. *Nucleic Acids Res* 2016, **44**(D1):D793-800.
112. Kozlov S, Malyavka A, McCutchen B, Lu A, Schepers E, Herrmann R, Grishin E: **A novel strategy for the identification of toxinlike structures in spider venom**. *Proteins: Structure, Function, and Bioinformatics* 2005, **59**(1):131-140.
113. Wong ES, Hardy MC, Wood D, Bailey T, King GF: **SVM-based prediction of propeptide cleavage sites in spider toxins identifies toxin innovation in an Australian tarantula**. *PLoS One* 2013, **8**(7):e66279.
114. Mouls L, Subra G, Aubagnac JL, Martinez J, Enjalbal C: **Tandem mass spectrometry of amidated peptides**. *J Mass Spectrom* 2006, **41**(11):1470-1483.
115. Labella C, Kanawati B, Vogel H, Schmitt-Kopplin P, Laurino S, Bianco G, Falabella P: **Identification of two arginine kinase forms of endoparasitoid *Leptomastix dactylopii* venom by bottom up-sequence tag approach**. *J Mass Spectrom* 2015, **50**(5):756-765.
116. Yamamoto T, Arimoto H, Kinumi T, Oba Y, Uemura D: **Identification of proteins from venom of the paralytic spider wasp, *Cyphononyx dorsalis***. *Insect Biochem Mol Biol* 2007, **37**(3):278-286.
117. Matsui T, Fujimura Y, Titani K: **Snake venom proteases affecting hemostasis and thrombosis**. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 2000, **1477**(1):146-156.
118. Winningham KM, Fitch CD, Schmidt M, Hoffman DR: **Hymenoptera venom protease allergens**. *J Allergy Clin Immunol* 2004, **114**(4):928-933.
119. Laurino S, Grossi G, Pucci P, Flagiello A, Bufo SA, Bianco G, Salvia R, Vinson SB, Vogel H, Falabella P: **Identification of major *Toxoneuron nigriceps* venom proteins using an integrated transcriptomic/proteomic approach**. *Insect Biochem Mol Biol* 2016, **76**:49-61.
120. Asgari S, Zhang G, Zareie R, Schmidt O: **A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph**. *Insect Biochem Mol Biol* 2003, **33**(10):1017-1024.
121. Tan G-J, Peng Z-K, Lu J-P, Tang F-Q: **Cathepsins mediate tumor metastasis**. *World J Biol Chem* 2013, **4**(4):91.
122. Beton D, Guzzo CR, Ribeiro AF, Farah CS, Terra WR: **The 3D structure and function of digestive cathepsin L-like proteinases of *Tenebrio molitor* larval midgut**. *Insect Biochem Mol Biol* 2012, **42**(9):655-664.
123. Kini RM: **Excitement ahead: structure, function and mechanism of snake venom phospholipase A 2 enzymes**. *Toxicon* 2003, **42**(8):827-840.
124. Burke JE, Dennis EA: **Phospholipase A2 structure/function, mechanism, and signaling**. *J Lipid Res* 2009, **50**(Supplement):S237-S242.
125. Jacomini DLJ, Pereira FDC, dos Santos Pinto JRA, dos Santos LD, da Silva Neto AJ, Giratto DT, Palma MS, de Lima Zollner R, Braga MRB: **Hyaluronidase from the venom of the social wasp *Polybia paulista* (Hymenoptera, Vespidae): Cloning, structural modeling, purification, and immunological analysis**. *Toxicon* 2013, **64**:70-80.
126. Jin C, Focke M, Léonard R, Jarisch R, Altmann F, Hemmer W: **Reassessing the role of hyaluronidase in yellow jacket venom allergy**. *J Allergy Clin Immunol* 2010, **125**(1):184-190. e181.
127. Hoffman DR: **Hymenoptera venom allergens**. *Clin Rev Allergy Immunol* 2006, **30**(2):109-128.
128. Wiese MD, Brown SG, Chataway TK, Davies NW, Milne RW, Aulfrey SJ, Heddle RJ: ***Myrmecia pilosula* (Jack Jumper) ant venom: identification of allergens and revised nomenclature**. *Allergy* 2007, **62**(4):437-443.
129. McCleskey E, Fox A, Feldman D, Cruz L, Olivera B, Tsien R, Yoshikami D: **Omega-conotoxin: direct and persistent blockade of specific types of calcium channels in**

neurons but not muscle. *Proceedings of the National Academy of Sciences* 1987, **84**(12):4327-4331.
130. Terlau H, Olivera BM: **Conus venoms: a rich source of novel ion channel-targeted peptides**. *Physiol Rev* 2004, **84**(1):41-68.
131. Zobel-Thropp PA, Kerins AE, Binford GJ: **Sphingomyelinase D in sicariid spider venom is a potent insecticidal toxin**. *Toxicon* 2012, **60**(3):265-271.
132. Binford GJ, Cordes MH, Wells MA: **Sphingomyelinase D from venoms of *Loxosceles* spiders: evolutionary insights from cDNA sequences and gene structure**. *Toxicon* 2005, **45**(5):547-560.
133. Ueberheide BM, Fenyö D, Alewood PF, Chait BT: **Rapid sensitive analysis of cysteine rich peptide venom components**. *Proceedings of the National Academy of Sciences* 2009, **106**(17):6910-6915.
134. Arias HR, Blanton MP: **α-Conotoxins**. *The International Journal of Biochemistry & Cell Biology* 2000, **32**(10):1017-1028.
135. Dutertre S, Jin AH, Kaas Q, Jones A, Alewood PF, Lewis RJ: **Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom**. *Molecular & cellular proteomics : MCP* 2013, **12**(2):312-329.
136. Vijay N, Poelstra JW, Kunstner A, Wolf JB: **Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments**. *Mol Ecol* 2013, **22**(3):620-634.
137. Dutertre S, Jin A-H, Vetter I, Hamilton B, Sunagar K, Lavergne V, Dutertre V, Fry BG, Antunes A, Venter DJ *et al*: **Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails**. *Nature Communications* 2014, **5**.