Adrian Jose Jaimes Becerra

Evolução do veneno em cnidários baseada em dados de genomas e proteomas

Venom evolution in cnidarians based on genomes and proteomes data

São Paulo

2015

Adrian Jose Jaimes Becerra

Evolução do veneno em cnidários baseada em dados de genomas e proteomas

Venom evolution in cnidarians based on genomes and proteomes data

Dissertação apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Mestre em Ciências, na Área de Zoologia.

Orientador: Prof. Dr. Antonio C. Marques

São Paulo

2015

Jaimes-Becerra, Adrian J.

Evolução do veneno em cnidários baseada em dados de genomas e proteomas.
103 + VI páginas

Dissertação (Mestrado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Zoologia.

1. Veneno; 2. Evolução; 3. Proteoma. 4. Genoma
I. Universidade de São Paulo. Instituto de Biociências. Departamento de Zoologia.

Comissão Julgadora

_____ _____

Prof(a) Dr(a) Prof(a) Dr(a)

_____

Prof. Dr. Antonio Carlos Marques

Agradecimentos

Eu gostaria de agradecer ao meu orientador Antonio C. Marques, pela confiança desde o primeiro dia e pela ajuda tanto pessoal como profissional durantes os dois anos de mestrado. Obrigado por todo.

Ao CAPES, pela bolsa de mestrado concedida. Ao FAPESP pelo apoio financeiro durante minha estadia em Londres.

Ao Instituto de Biociências da Universidade de São Paulo, pela estrutura oferecida durante a execução desde estudo.

Ao Dr. Paul F. Long pelas conversas, por toda sua ajuda, por acreditar no meu trabalho.

Aos colegas e amigos de Laboratório de Evolução Marinha (LEM), Jimena Garcia, María Mendoza, Thaís Miranda, Amanda Cunha, Karla Paresque, Marina Fernández, Fernanda Miyamura e Lucília Miranda, pela amizade, dicas e ajuda em tudo e por me fazer sentir em casa, muito obrigado mesmo!

Aos meus amigos fora do laboratório, John, Soly, Chucho, Camila, Faride, Cesar, Angela, Camilo, Isa, Nathalia, Susana e Steffania, pelo apoio e por me fazer sentir em casa.

Ao Adriana pelo grande apoio sempre e por me fazer feliz.

Aos meus pais, irmãos e sobrinhos, pela ajuda e apoio em todos os momentos. Los llevaré en mis pensamientos donde quiera que esté.

Gracias a todos.

Índice

Capítulo 1. Introdução Geral 1

O VENENO NO REINO ANIMAL 1

O VENENO NOS CNIDÁRIOS 2

OBJETIVOS GERAIS DO ESTUDO 4

ORGANIZAÇÃO DA DISSERTAÇÃO 4

REFERÊNCIAS 5

Capítulo 2. Comparative proteomics reveals common components of a powerful arsenal in the earliest animal venomous lineage, the cnidarians 10

ABSTRACT 10

RESUMO 11

INTRODUCTION 12

MATERIAL AND METHODS 14

RESULTS 16

DISCUSSION 17

REFERENCES 23

FIGURES 34

TABLES 37

SUPPLEMENTARY MATERIAL 39

Capítulo 3. Evidence of episodic positive selection in the evolution of jellyfish toxins (JFTs) of the cnidarian venom 51

ABSTRACT 51

RESUMO 52

INTRODUCTION 52

MATERIALS AND METHODS 54

RESULTS 55

DISCUSSION 56

REFERENCES 60

FIGURES 65

TABLES 67

SUPPLEMENTARY MATERIAL 69

CAPÍTULO 4. Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from Acropora digitifera (Cnidaria: Anthozoa: Scleractinia) 71

ABSTRACT 71

INTRODUCTION 72

RESULTS 76

DISCUSSION 78

CONCLUSIONS 83

METHODS 83

COMPETING INTERESTS 86

AUTHORS' CONTRIBUTIONS 86

ACKNOWLEDGMENTS 87

REFERENCES 87

FIGURES 94

TABLES 95

CAPÍTULO 5. Considerações Finais 99

REFERÊNCIAS 100

Resumo 102

Abstract 103

CAPÍTULO 1

INTRODUÇÃO

O VENENO NO REINO ANIMAL

Animais venenosos têm sido tema de fascinação ao longo da história humana, evidentemente também pelo perigo inerente associado a esses organismos (Casewell *et al*., 2013; Arbuckle, 2015). Igualmente, animais venenosos têm sido o objeto de numerosos estudos científicos no último século, oferecendo uma visão interessante e por vezes única em diferentes áreas biológicas (Dutertre & Lewis, 2010; King, 2011; Harrison *et al*., 2011; Williams *et al*., 2011, Sunagar *et al*., 2015).

O veneno é definido como uma secreção tóxica que causa lesão fisiológica, sendo transferido passiva ou ativamente de um organismo ao meio interno de outro organismo, por meio de mecanismos de liberação e lesão mecânica (Nelsen *et al*., 2013). Essa definição inclui animais considerados venenosos (por exemplo, escorpiões, serpentes, aranhas e cnidários), bem como animais que não têm sido tradicionalmente reconhecidos como tais (por exemplo, sanguessugas, carrapatos, morcegos hematófagos). Ao se reconhecer analogias evolutivas de recrutamento e utilização de toxinas por parte de todos esses animais, assume-se um grande número de eventos em que o veneno evoluiu de forma independente, já que ocorre em pelo menos trinta linhagens diferentes (Fry *et al*., 2009a; Casewell *et al*., 2013). O melhor reconhecimento sobre essas expressões melhora nossa compreensão dos fatores subjacentes à evolução dos venenos e suas proteínas associadas. Paralelamente, chama a atenção para o grande conjunto de toxinas ainda não estudadas, assumidas como um grande potencial para a descoberta de moléculas bioativas.

O veneno possui múltiplas funções nas diferentes linhagens do reino animal, como predação (Fry *et al*., 2009b; Pekar *et al*., 2014), defesa (Inceoglu *et al*., 2003; Dutertre *et al*., 2014; Grow *et al*., 2015; Nisani & Hayes, 2015), competição intraespecífica (Whittington *et al*., 2009; Macrander *et al*., 2015) e reprodução (Leeming, 2003) tem sido atribuídas nos distintos linhagens estudados. Esta diversidade funcional e taxonômica destaca a importância do veneno como uma inovação evolutiva.

Consequentemente, uma grande variedade de estruturas evoluíram para facilitar a inoculação de venenos, como, bicos, dentes, arpões, nematocistos, probóscides, espinhos, sprays, e aguilhões (Fry *et al*., 2009a; Smith & Wheeler, 2006; Vonk *et al*., 2008; Beckmann & Ozbek, 2012).

A maioria dos venenos animais é um coquetel complexo de compostos bioativos. Os venenos compreendem tipicamente uma mistura de proteínas e peptídeos (referido como toxinas), sais e componentes orgânicos, tais como aminoácidos e neurotransmissores (Fry *et al*., 2009a; Hargreaves *et al*., 2014, Jouiaei *et al*., 2015b). Os componentes proteicos geralmente são os mais abundantes. A composição do veneno geralmente reflete sua função, com venenos defensivos, como em peixes ou abelhas, sendo mais simples e conservados, em que a ação principal frequentemente é uma dor localizada extrema e imediato (Church & Hodgson, 2002; Peiren *et al*., 2005; de Graaf *et al*., 2009). Em contraste, os venenos de predadores são mais complexos e variáveis em composição e efeitos fisiológicos (Fry *et al*., 2009a), e essa complexidade ainda aumenta o potencial de variação na composição do veneno. Essa diversidade de composições resulta em uma variação extrema na toxicidade e no modo de ação do veneno entre táxons próximos (Mackessy, 2010), populações de uma mesma espécie (Calvete *et al*., 2010), sexos diferentes (Menezes *et al*., 2010), variações ontogenéticas na vida de um indivíduo (Andrade & Abe, 1999) e vários outros níveis (Chippaux *et al*., 1991). Esses processos moldam, seja por fatores históricos ou ambientais, o conteúdo do veneno (Mackessy, 2009; Casewell *et al*., 2013).

A importância evolutiva e ecológica do veneno tem sido enfatizada nos últimos anos, bem como sua influência determinante sobre interações interespecíficas (Sunagar *et al*., 2015). Ainda, os sistemas de venenos fornecem modelos sem paralelo para investigar bases moleculares da adaptação, isto é, as inter-relações entre seleção natural e os processos genéticos e moleculares responsáveis por gerar a diversidade molecular e, portanto, a variação na composição das toxinas e sua ação (Wong & Belov, 2012, Casewell *et al*., 2013; Starcevic & Long, 2013).

O VENENO NOS CNIDÁRIOS

Os cnidários, tais como anêmonas-do-mar, corais, águas-vivas e hidras, são a mais antiga linhagem existente de animais venenosos. Desde sua origem no Neoproterozoico, há ~600 milhões de anos, e portanto antes da Irradiação Cambriana,

essa linhagem de anatomia simples presenciou com o surgimento e declínio de inúmeras novas formas de vida com estratégias mais complexas de sensoriamento ambiental, processamento de informações, locomoção e alimentação (Shinzato *et al*., 2007; Van Iten *et al*., 2014). Talvez uma parte significante do sucesso dos cnidários possa ser atribuída justamente à produção de veneno, mediador essencial na interação com presas e predadores morfologicamente mais complexos (Anderluh & Macek, 2002; Saput, 2009; Badre, 2014).

As cnidas são organelas especializadas que definem o filo Cnidaria, capazes de descarregar seu conteúdo interno mediante a ativação dos cnidócitos por estímulos externos, mecânicos ou químicos. Cnidas contêm elementos estruturais e químicos elaborados, que atuam em diferentes funções. Cnidas estão distribuídas em várias partes do corpo dos cnidários, sendo classificadas em três tipos principais, viz., nematocistos, espirocistos e pticocistos (Östman, 2000; Özbek *et al*., 2009). Os nematocistos, especificamente, são encontrados em todos os cnidários, sendo a estrutura primária de inoculação do veneno nos organismos-alvo (Fautin, 2009).

Desde o início do século XX, vários experimentos analíticos e observações clínicas exploraram a diversidade toxicológica dos venenos de cnidários. A diversidade dos componentes do veneno varia desde compostos não proteicos (por exemplo, purinas, aminas biogênicas) até proteínas de peso molecular elevado, tais como enzimas que incluem proteínas lipolíticas e proteolíticas que catabolizam os tecidos da presa, toxinas que formam poros e podem causar morte celular via lise osmótica, e neurotoxinas que exibem atividades rápidas e específicas atuando sobre canais iônicos (Šuput, 2009; Mariottini & Pane, 2013; Badré, 2014; Mariottini, 2014; Jouiaei *et al*., 2015b; Mariottini *et al*., 2015).

Como outros campos da Biologia, investigações sobre venenos foram revolucionadas nos últimos anos com abordagens da biologia de sistemas, i.e. genômica, transcriptômica e proteômica. Estudos de transcriptomas e proteomas têm mostrado que os nematocistos de vários cnidários contêm proteínas, algumas das quais únicas para o grupo, e outra já identificadas previamente para outros animais venenosos, mas poucas toxinas, de fato têm sido caracterizadas (Balasubramanian *et al*., 2012; Weston *et al*., 2013; Li *et al*., 2014; Rachamim *et al*., 2014; Brinkman *et al*., 2015; Gacesa *et al*., 2015; Jouiaei *et al*., 2015a; Macrander *et al*., 2015). Essa complexidade

surpreendente e o enorme potencial dos venenos de cnidários faz surgir várias questões e possibilidades únicas para pesquisas.

OBJETIVOS GERAIS DO ESTUDO

Listas de proteínas de oito espécies correspondentes a quatro classes de cnidários, sequências de nucleótidos e aminoácidos de uma família de toxinas especifica de cnidários e o genoma do coral *Acropora digitifera*, foram usadas com o objetivo de: (i) caracterizar comparativamente a composição dos diferentes venenos entre as classes de cnidários; (ii) inferir a evolução da composição do veneno entre as diferentes classes; (iii) inferir a história evolutiva e as pressões de seleção que influenciaram uma família de toxinas especifica de cnidários, (iv) determinar qual a importância de processos de duplicação gênica na formação da diversidade molecular no veneno da espécie de coral *Acropora digitifera*.

ORGANIZAÇÃO DA DISSERTAÇÃO

Esta tese está organizada em cinco capítulos, sendo o primeiro esta introdução, que expõe as características gerais do estudo, seus objetivos e sua organização.

O Capítulo 2, "Comparative proteomics reveals common components of a powerful arsenal in the earliest animal venomous lineage, the cnidarians", tem como objetivo caracterizar e elucidar a evolução da composição do veneno em Cnidaria por meio da comparação de listas de proteínas resultantes das análises proteômicas de *Chrysaora lactea*, *Tamoya haplonema* e *Chiropsalmus quadrumanus*, geradas neste estudo, e listas de proteínas de estudos publicados anteriormente para *Acropora digitifera* (Gacesa *et al*., 2015), *Olindias sambaquiensis* (Weston *et al*., 2013), *Anemonia viridis*, *Hydra magnipapillata* e *Aurelia aurita* (Rachamim *et al*., 2014). Essa análise corresponde ao estudo comparado mais completo sobre a composição do veneno de cnidários, levantando hipóteses sobre a montagem evolutiva do complexo arsenal bioquímico de cnidários e dos venenos ancestrais desse grupo basal.

O Capítulo 3, "Evidence of episodic positive selection in the evolution of jellyfish toxins of the cnidarian venom", tem como objetivo testar a hipótese de que a variação nessa família de toxinas específica de cnidários tem sido o resultado de um regime de seleção positiva. De fato, nossas análises identificaram um padrão diferente, em que há uma seleção purificadora por longos períodos seguindo longos tempos de

diversificação. Esse padrão é contrastado com a biologia do grupo e a função que nematocistos possuem a história natural de cnidários.

O Capítulo 4, "Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia)", é um artigo que publicamos recentemente e aqui é apresentado como anexo. Este estudo determina a extensão em que a duplicação de genes pode ser considerada como a principal razão para a diversificação de toxinas em Cnidaria. Para tal, comparamos sequências de aminoácidos das toxinas previstas e derivadas do genoma traduzido de *A. digitifera* com toxinas putativas observadas em análises proteômicas de proteínas solúveis que foram obtidas de nematocistos isolados. Concluímos que a duplicação genica desempenha um papel significativo para a diversificação de toxinas nesta espécie de coral, mas não explica a totalidade da diversidade de seu arsenal.

O Capítulo 5 traz as considerações finais deste estudo, destacando os principais resultados obtidos nos capítulos anteriores, e apresenta um marco teórico de questões que surgiram a partir desses resultados e que poderão ser abordadas em trabalhos futuros sobre evolução dos venenos em cnidários.

REFERÊNCIAS

Anderluh, G. & Maček, P. 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). Toxicon 40: 111–124.

Andrade, D.V. & Abe, A.S. 1999. Relationship of venom ontogeny and diet in *Bothrops*. Herpetologica 55:200–204.

Arbuckle, K. 2015. Evolutionary context of venom in animals. Evolution of venemous Animals and Their Toxins 1–23.

Badré, S. 2014. Bioactive toxins from stinging jellyfish. Toxicon 91: 114–125.

Balasubramanian, P.G.; Beckmann, A.; Warnken, U.; Schnölzer, M.; Schüler, A.; Bornberg-Bauer, E.; Holstein, T.W. & Özbek, S. 2012. Proteome of *Hydra* nematocyst. Journal of Biological Chemistry 287: 9672–9681.

Beckmann, A. & Özbek, S. 2012. The nematocyst: A molecular map of the cnidarian stinging organelle. International Journal of Developmental Biology 56: 577–582.

Brinkman, D.L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D. & Mulvenna, J. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16: 1–15.

Calvete, J.J.; Sanz, L.; Cid, P.; de la Torre, P.; Flores-Díaz, M.; Dos Santos, M.C.; Borges, A.; Bremo, A.; Angulo, Y.; Lomonte, B.; Alape-Girón, A. & Gutiérrez, J.M. 2010. Snake venomics of the Central American rattlesnake *Crotalus simus* and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive paedomorphic trend along Crotalus dispersal in South America. Journal of proteome research 9:528–544.

Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A. & Fry, B.G. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends in Ecology and Evolution 28: 219–229.

Chippaux, J.P.; Williams, V. & White. J. 1991. Snake venom variability: Methods of study, results and interpretation. Toxicon 29:1279–303.

Church, J.E. & Hodgson, W.C. 2002. The pharmacological activity of fish venoms. Toxicon 40:1083–1093.

De Graaf, D.C.; Aerts, M.; Danneels, E. & Devreese, B. 2009. Bee, wasp and ant venomics pave the way for a component-resolved diagnosis of sting allergy. Journal of proteomics 72:145–154.

Dutertre, S. & Lewis, R.J. 2010. Use of venom peptides to probe ion channel structure and function. The journal of biological chemistry 285: 13315–13320.

Dutertre, S.; Jin, A.H.; Vetter, I.; Hamilton, B.; Sunagar, K.; Lavergne, V.; Dutertre, V.; Fry, B.G.; Antunes, A.; Venter, D.J.; Alewood, P.F. & Lewis, R.J. 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nature Communications 5:3521.

Fautin, D.G. 2009. Structural diversity, systematics, and evolution of cnidae. Toxicon 54: 1054–1064.

Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.; Lewis, R.J.; Norton, R.S.; Renjifo, C. & de la Vega, R.C.R. 2009a. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10: 483–511.

Fry, B.G.; Wroe, S.; Teeuwisse, W.; van Osch, M.J.P.; Moreno, K.; Ingle, J.; McHenry, C.; Ferrara, T.; Clausen, P.; Scheib, H.; Winter, K.L.; Greisman, L.; Roelants, K.; van derWeerd, L.; Clemente, C.J.; Giannakis, E.; Hodgson, W.C.; Luz, S.; Martelli, P.; Krishnasamy, K.; Kochva, E.; Kwok, H.F.; Scanlon, D.; Karas, J.; Citron, D.M.; Goldstein, E.J.C.; McNaughtan, J.E. & Norman, J.A. 2009b. A central role for venom in predation by *Varanus komodoensis* (Komodo dragon) and the extinct giant Varanus (Megalania) priscus. Proceedings of the National Academy of Sciences of the United States of America 106:8969–74.

Gacesa, R.; Chung, R.; Dunn, S.R.; Weston, A.; Jaimes-Becerra, A.; Marques, A.C.; Morandini, A.; Hranueli, D.; Starcevic, A.; Ward, M. & Long, P.F. 2015. Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from Acropora digitifera (Cnidaria: Anthozoa: Scleractinia). BMC Genomics 16:774.

Grow, N.B. & Wirdateti, N.K.A.I. 2015. Does toxic defence in *Nycticebus* spp. relate to ectoparasites? The lethal effects of slow loris venom on arthropods. Toxicon 95:1–5.

Hargreaves, A.D.; Swain, M.T.; Logan, D.W. & Mulley, J.F. 2014. Testing the Toxicofera: comparative reptile transcriptomics casts doubt on the single, early evolution of the reptile venom system. Toxicon 92:140–156.

Harrison, R.A.; Cook, D.A.; Renjifo, C.; Casewell, N.R.; Currier, R.B. & Wagstaff, S.C. 2011. Research strategies to improve snakebite treatment: challenges and progress. Journal of proteomics 74:1768–1780.

Inceoglu, B.; Lango, J.; Jing, J.; Chen, L.; Doymaz, F.; Pessah, I.N. & Hammock, B.D. 2003. One scorpion, two venoms: prevenom of *Parabuthus transvaalicus* acts as an alternative type of venom with distinct mechanism of action. Proceedings of the National Academy of Sciences of the United States of America 100:922–7.

Jouiaei, M.; Sunagar, K.; Gross, A.F.; Scheib, H.; Alewood, P.F.; Moran, Y. & Fry, B.G. 2015a. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Molecular Biology and Evolution 32: 1598–1610.

Jouiaei, M.; Yanagihara, A.; Madio, B.; Nevalainen, T.; Alewood, P. & Fry, B. 2015b. Ancient venom systems: A review on Cnidaria toxins. Toxins 7: 2251–2271.

King, G.F. 2011. Venoms as a plataform for human drugs: translating toxins into therapeutics. Expert Opinion on Biological Therapy 11: 1469–1484.

Leeming, J. 2003. Scorpions of southern Africa. South Africa: Struik Publishers.

Li, R.; Yu, H.; Xue, W.; Yue, Y.; Liu, S.; Xing, R. & Li, P. 2014. Jellyfish venomics and venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. Journal of Proteomics 106: 17–29.

Mackessy, S.P. 2009. The field of reptile toxinology: snakes, lizards and their venoms. In: Mackessy, S.P. (Ed.), Handbook of venoms and toxins of reptiles. CRC Press/Taylor & Francis Group, Boca Raton, FL 3–23.

Mackessy, S.P. 2010. Evolutionary trends in venom composition in the western rattlesnake (*Crotalus viridis* sensu lato). Toxicity vs. tenderizers. Toxicon 55:1463–1474.

Macrander, J.; Brugler, M.R. & Daly, M. 2015. A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps. BMC Genomics 16: 1–19.

Mariottini, G.L. 2014. Haemolytic venoms from marine cnidarian jellyfish – an overview. Journal of Venom Research 5: 22–32.

Mariottini, G.L.; Bonello, G.; Giacco, E. & Pane, L. 2015. Neurotoxic and neuroactive compounds from Cnidaria: Five decades of research…. and more. Central Nervous System Agents in Medicinal Chemistry 15: 74–80.

Mariottini, G.L. & Pane, L. 2013. Cytotoxic and cytolytic cnidarian venoms. A review on health implications and possible therapeutic applications. Toxins 6: 108–151.

Menezes, M.C.; Furtado, M.F.; Travaglia-Cardoso, S.R.; Camargo, A.C. & Serrano, S.M. 2006. Sex-based individual variation of snake venom proteome among eighteen *Bothrops jararaca* siblings. Toxicon 47:304–312.

Nelsen, D.R.; Nisani, Z.; Cooper, A.M.; Fox, G.A.; Gren, E.C.K. Corbit, A.G. & Hayes, W.K. 2014. Poisons, toxungens, and venoms: Redefining and classifying toxic biological secretions and the organisms that employ them. Biological Reviews 89: 450–465.

Nisani, Z. & Hayes, W.K. 2015. Venom-spraying behavior of the scorpion *Parabuthus transvaalicus* (Arachnida: Buthidae). Behavioural Processes 115:46–52.

Östman, C. 2000. A guideline to nematocyst nomenclature and classification, and some notes on the systematic value of nematocysts. Scientia Marina 64: 31–46.

Özbek, S.; Balasubramanian, P.G. & Holstein, T.W. 2009. Cnidocyst structure and the biomechanics of discharge. Toxicon 54: 1038–1045.

Peiren, N.; Vanrobaeys, F.; de Graaf, D.C.; Devreese, B.; Van Beeumen, J. & Jacobs F.J. 2005. The protein composition of honeybee venom reconsidered by a proteomic approach. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1752: 1–5.

Pekár, S.; Šedo, O.; Líznarová, E.; Korenko, S. & Zdráhal, Z. 2014. David and Goliath: potent venom of an ant-eating spider (Araneae) enables capture of giant prey. Naturwissenschaften 101:533–40.

Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T. & Sher, D. 2014. The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. Molecular Biology and Evolution 32: 740–753.

Starcevic, A. & Long, P.F. 2013. Diversification of animal venom peptides-were jellyfish amongst the first combinatorial chemists?. ChemBioChem 14: 1407–1409.

Sunagar, K.; Morgenstern, D.; Reitzel, A.M. & Moran, Y. (in press). 2015. Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. Journal of proteomics.

Smith, W.L. & Wheeler, W.C. 2006. Venom evolution widespread in fishes: a phylogenetic road map for the bioprospecting of piscine venoms. Journal of heredity 97:206–217.

Šuput, D. 2009. In vivo effects of cnidarian toxins and venoms. Toxicon 54: 1190–1200.

Van Iten, H.; Marques, A.C.; Leme, J.D.M.; Pacheco, M.L.F. & Simões, M.G. 2014. Origin and early diversification of the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's Proterozoic-Cambrian history. Palaeontology 57: 677–690.

Vonk, F.J.; Admiraal, J.F.; Jackson, K.; Reshef, R.; de Bakker, M.A.J.; Vanderschoot, K; Van den Berge, I.; Van Atten, M.; Burgerhout, E.; Beck, A.; Mirtschin, P.J.; Kochva, E.; Witte, F.; Fry, B.G.; Woods, A.E. & Richardson, M.K. 2008. Evolutionary origin and development of snake fangs. Nature 454:630–633.

Weston, A.J.; Chung, R.; Dunlap, W.C.; Morandini, A.C.; Marques, A.C.; Moura-da-Silva, A.M.; Ward, M.; Padilla, G.; da Silva, L.F.; Andreakis, N. & Long, P.F. 2013. Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish *Olindias sambaquiensis*. Toxicon 71: 11–17.

Whittington, C.M.; Koh ,J.M.S.; Warren, W.C.; Papenfuss, A.T.; Torres, A.M.; Kuchel, P.W. & Belov, K. 2009. Understanding and utilising mammalian venom via a *platypus* venom transcriptome. Journal of Proteomics 72:155–64.

Williams, D.J.; Gutiérrez, J.M.; Calvete, J.J.; Wüster, W.; Ratanabanangkoon, K.; Paiva, O.; Brown, N.I.; Casewell, N.R.; Harrison, R.A.; Rowley, P.D.; O'Shea, M.; Jensen, S.D.; Winkel, K.D. & Warrell, D.A. 2011. Ending the drought: new strategies for improving the flow of affordable, effective antivenoms in Asia and Africa. Journal of proteomics 74:1735–1767.

Wong, E.S.W. & Belov, K. 2012. Venom evolution through gene duplications. Gene 496:1–7.

Capítulo 2

Comparative proteomics reveals common components of a powerful arsenal in the earliest animal venomous lineage, the cnidarians

Adrian Jaimes-Becerra[1], Paul F. Long[2-3], Antonio C. Marques[1-4]

[1]Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua Matão, Trav. 14, 101, 05508-090 São Paulo, SP, Brasil.
[2]Department of Chemistry & 9Brazil Institute, King's College London, Strand, London WC2R 2LS, United Kingdom.
[3]Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Av. Prof. Lineu Prestes, 580, B16, 05508-000 São Paulo, SP, Brasil.
[4]Centro de Biologia Marinha, Universidade de São Paulo, Rodovia Manoel Hypólito do Rego, km. 131,5, 11600-000 São Sebastião, Brasil.

ABSTRACT

The evolution of venom, nature's most complex concoction, has underpinned the predatory success and diversification of numerous animal lineages. The nematocysts in the phylum Cnidaria are the most evolutionarily ancient venom apparatus. In cnidarians, as well as in all other venomous animals, the stinging apparatus actively translocate venom compounds away from the originator and into an external target. Cnidarian predators use this exquisite mechanical device to capture and subdue prey. In this study, in order to compare the venom composition and identify recruitment patterns of the families of toxins found in the venom proteomes of the phylum Cnidaria, we select the proteomes of two species from four cnidarian classes (Anthozoa, Hydrozoa, Cubozoa and Scyphozoa). Twenty eight-toxin families types were identified in the venom proteome of the eight species studied. Fifteen of these families have been previously found in the venom of cnidarians. Twelve types of toxin families were shared between the four classes analysed, suggesting common proteome functionalities. Character mapping analysis revealed that at least fifteen-toxin families types were likely recruited into the cnidarian venom proteome before the lineage split between Anthozoa and Medusozoa. Nine of these types (AhpC/TSA, sodium channel inhibitor, phospholipase A_2, phospholipase D, peptidase S1, metalloproteinase, SCRIPs, potassium channel inhibitor and complement C3) have been identified in previous studies. However, the types conotoxins, Flavin monoamine oxidase, Glycosyl hydrolase 56, Latarcin,

Latrotoxin and Snake three-finger were identified in this study for the first time. We demonstrated that more types of toxin families are continuously recruited into venom proteome during the evolutionary process of individual cnidarian classes. This analysis is the most comprehensive comparative study on the cnidarians venom composition. Besides, our study has provided new insights into the evolutionary assembly of the complex biochemical arsenal of cnidarians and has addressed a partial insight into the composition of the earliest cnidarians venoms.

Keywords: cnidarians; venom; ; proteome; evolution

RESUMO

A evolução do veneno, a mistura mais complexa da natureza, tem sustentado o sucesso predatório e a diversificação de numerosas linhagens de animais. Os nematocistos do filo Cnidaria são o aparelho de veneno mais antigo evolutivamente. Em cnidários, como em todos os outros animais venenosos, o aparelho onde é armazenado o veneno expulsa ativamente compostos até um alvo externo. Os cnidários predadores usam este dispositivo mecânico para capturar e subjugar presas. Neste estudo, a fim de comparar a composição do veneno e identificar os padrões de recrutamento das famílias de toxinas encontradas no veneno dos proteomas do filo Cnidaria, nós selecionamos os proteomas de duas espécies para quatro classes do filo (Anthozoa, Hydrozoa, Cubozoa e Scyphozoa). Foram identificados vinte oito tipos de famílias de toxinas no proteoma do veneno das oito espécies estudadas. Quinze destas famílias foram previamente encontradas no veneno de cnidários. Doze tipos de famílias de toxinas são compartilhados entre as quatro classes analisadas, sugerindo funcionalidades comuns no proteoma. A análise de mapeamento de caracteres revelou que pelo menos quinze tipos de famílias de toxinas provavelmente foram recrutados para o proteoma do veneno de Cnidaria antes da divisão entre as linhagens Anthozoa e Medusozoa. Nove destes tipos (AhpC/TSA, inibidores do canal de sódio, fosfolipasas A2, fosfolipasas D, peptidasas S1, metaloproteinasas, SCRIPs, inibidores do canal de potássio e complemento C3) foram identificadas em estúdios prévios. Já os tipos conotoxinas, Flavin monoamino oxidase, Glicosil hydrolase 56, Latarcinas, Latrotoxinas e “três dedos” de serpentes foram identificados neste estudo pela primeira vez. Demonstramos também que mais tipos de famílias de toxinas foram continuamente recrutados no proteoma do veneno durante o processo evolutivo de cada classe de Cnidaria. Essa análise é o estudo

comparado mais completo sobre a composição do veneno de cnidários. Além disso, nosso estudo forneceu novas informações sobre a montagem evolutiva do complexo arsenal bioquímico de cnidários e oferece uma visão parcial da composição dos primeiros venenos em cnidários.

INTRODUCTION

Venoms are composed in large-part of toxic peptides and proteins that cause dose-dependent physiological disruption when delivered by the infliction of a wound into target prey or predator by a venomous animal (Nelsen *et al*., 2014). Venom has been a key innovation in the evolutionary history of an incredibly diverse range of animals, for example, snakes, scorpions, spiders, cephalopods, centipedes, cnidarians and even some mammals. This is because venom systems have evolved independently on at least twenty occasions in extant lineages as an ecological adaptation (Fry *et al*., 2009). Cnidaria are believed to be the most basal Metazoa to be venomous, maybe evolving since Neoproterozoic times, ~650 million years ago, much before the Cambrian radiation (Van Iten *et al*., 2014).

Cnidaria are a diverse phylum comprising over 10.000 predominately marine species, with few species in freshwater and estuarine habitats (Daly *et al*., 2007; Zhang, 2011, but see Okamura *et al*., 2015 for the inclusion of Myxozoa as cnidarians). Cnidaria has two major subphyla: Anthozoa and Medusozoa. Anthozoa includes the sea anemones and both the hard and soft corals (Bridge *et al*., 1992; Marques & Collins, 2004). Medusozoa comprises the classes Staurozoa (e.g. stalked jellyfish), Cubozoa (e.g. box jellyfish), Scyphozoa (e.g. 'true' jellyfish) and Hydrozoa (e.g. *Hydra* and relatives including several species of smaller jellyfish) (Marques & Collins, 2004; Collins *et al*., 2006; Van Iten *et al*., 2014). Anthozoa have basic lifecycles with planula larva and adult polyp stages, whereas Medusozoa have basic lifecycles with at least three generational phases, a free-swimming planula larva, a sessile polyp and a sexual pelagic medusa (Collins, 2002). Although many groups, especially Hydrozoa, have only polyp or jellyfish stages (Collins, 2002; Maronna *et al.*, 2015).

The most evident synapomorphy of Cnidaria is the presence of cnidae, organelles produced by the Golgi apparatus of specialised cells called cnidoblasts (Marques & Collins, 2004; Fautin, 2009; Beckmann & Özbek, 2012). Cnidae are found

in various parts of the body of a cnidarian and are classified into three morphological types, nematocysts, spirocysts and ptychocysts (Östman, 2000; Özbek *et al*., 2009). The nematocysts are universal cnida found in all cnidarians, but are morphologically and functionally heterogeneous (David *et al*., 2008; Fautin, 2009). Despite this, all nematocysts are composed of a capsule containing an inverted, highly folded, and hollow tubule which may be armed with spines (Östman, 2000; David *et al*., 2008; Beckmann & Özbek, 2012). The apex of nematocysts has a cap called operculum (Reft & Daly, 2012). The hydrostatic discharge mechanism of the capsule is activated after chemical or mechanical stimulation of the operculum causing the tubule to became discharged, subsequently causing injection of the venom into the victim (Lotan *et al*., 1995; Olivera, 2002; Fautin, 2009; Özbek *et al*., 2009; Beckmann & Özbek, 2012; Reft & Daly, 2012). The process of nematocyst discharge is one of the fastest known biological phenomena and can be as short as 700 ns with an acceleration of 5.413.000 x g (Nüchter *et al*., 2006; Özbek *et al*., 2009). In addition to prey capture and defence against predation, the venom content of nematocysts is believed to function in the regulation of spatial intraspecific and interspecific competition (Bigger, 1980; Kass-Simon & Scappaticci, 2002).

Increasingly, many transcriptomic and proteomic studies have described toxin protein families in the venoms of cnidarians that are astonishingly similar to the toxins of other venomous animals (Balasubramanian *et al*., 2012; Brinkman *et al*., 2012, 2015; Li *et al*., 2012, 2014; Weston *et al*., 2013; Rachamim *et al*., 2014; Jouiaei *et al*., 2015a, 2015b; Macrander *et al*., 2015). Likewise, the predicted biological activities of these cnidarian toxins encompass all of the major disruptive properties found in other venomous animals including enzymatic, neurotoxic and cytolytic functions (Šuput, 2009; Mariottini & Pane, 2013; Badré, 2014; Mariottini, 2014; Jouiaei *et al*., 2015c; Mariottini *et al*., 2015). Thus, these studies suggest that understanding the mechanisms or factors underpinning toxin diversification in Cnidaria can provide a platform from which the evolution of this trait in higher animals might be more fully explored. However, to achieve this, an understanding of the evolutionary trends of major toxin protein families across the phylum Cnidaria is required. To date, most studies attempting to infer evolutionary aspects of the diversification of toxins of cnidarian venom (Rachamim *et al*., 2014; Brinkman *et al*., 2015; Jouiaei *et al*., 2015b), have

limited taxon sampling, either by the represented classes studied or by all species restricted to a unique class of cnidarians.

The aim of this study is to elucidate the evolution of venom composition among the different classes of cnidarians. To get that, we increase the number of nematocyst proteomes available for comparative analyses. These new data, together with the previously published nematocyst proteomes, were then used as inputs for character-mapping tools, thereby establishing the most complete venom assembly hypothesis to date concurrent with the evolutionary history of this animal group.

MATERIAL AND METHODS

Experimentally derived proteomes

The toxin proteomes of nematocysts isolated from three cnidarians, the scyphozoan *Chrysaora lactea* and two cubozoans, *Tamoya haplonema* and *Chiropsalmus quadrumanus,* samples were taken as in other studies (Weston *et al.*, 2012, 2013). Briefly, animals were collected along coast of São Paulo state (Guarujá County) by bottom shrimp trawls (2 cm mesh size) dragged at 10m deep in Enseada beach on May 7th 2012. Animals were brought to the laboratory and identified based on morphological characters and regional literature (Morandini *et al.*, 2005; Morandini & Marques, 2010; Collins *et al.*, 2011). Intact nematocysts were isolated by modification of the method of Weber *et al*., (1987). The tentacles were homogenized in one cold SuFi solution (300 mM sucrose containing 50% v/v Ficoll-Paque Plus, GE Healthcare). This material was kept at 4°C for 30 min and then passed through a 2 mm diameter sieve. The sample was centrifuged for 10 min at 3000 g at 4°C. The pellet containing intact nematocysts was suspended and washed in cold SuFi solution. The final material was submitted, after microscopic inspection, for lyophilisation. This part was realized in the Laboratório de Bioprodutos of the Institute of Biomedical Science at the University of São Paulo. The proteomic data lists were provided by The Proteomics Unit at King´s College London, again using the methods Weston *et al.* (2012, 2013). To one tube of freeze dried nematocysts, 1ml of protein extraction buffer containing 50mM TEAB, 0.04% SDS, protease inhibitors (Roche) and phosphatase inhibitors (Roche) was added. The reconstituted material was disrupted in a sonic bath (VWR, Lutterworth, UK) for 15 min. The tubes were then centrifuged for 10 minutes at 10,000 x g and 4°C. The

supernatant was decanted and protein concentration determination was performed by Bradford assay. A volume of protein extract equivalent to 15µg of protein was amended to 15µl in extraction buffer before adding 15µl of 2X Laemmli sample buffer, heated for 10 minutes at 95oC and loaded onto a 4-12% NuPAGE gel (Life Technologies) for separation by SDS-PAGE. Electrophoresis was performed using NuPAGE MES (2-(N-morpholino) ethanesulfonic acid) buffer (Life technologies) at 150V for approximately 100 minutes. The entire gel lane was then divided into 15 equal sections, excised and cut into 2mm pieces. In-gel reduction, alkylation, and proteolytic digestion with trypsin were performed for each gel section prior to liquid chromatographic separation and mass spectrometric analysis (Schevchenko *et al*., 1996) as follows. Cysteine residues were reduced with 10mM dithiothreitol and alkylated with 55mM iodoacetamide in 100mM ammonium bicarbonate to form stable carbamidomethyl derivatives. Trypsin digestion was carried out overnight at 37°C in 50mM ammonium bicarbonate buffer and the supernatant was retained. Peptides were extracted from the gel pieces by two washes with 50 mM ammonium bicarbonate and acetonitrile. Each wash involved shaking the gel pieces for 10 minutes. The extract was pooled with the initial digestion supernatant and then lyophilised. Lyophilized extract was reconstituted in 30µl of 50mM ammonium bicarbonate buffer prior to LC-MS/MS analysis with 10µl of the sample injected. To LC-MS/MS analysis, samples were analysed on a Thermo Scientific Orbitrap Velos Pro mass spectrometer coupled to an EASY-nLC II (Proxeon) nano LC system.

Generally, the standard approach used to identify peptides of Rawfile data (MS/MS spectra) from mass spectrometry analysis was through one search strategy in the Tox-Prot UniProtKB/Swiss-Prot database (Jungo *et al*., 2012) via search engine MASCOT (Perkins *et al*., 1999). This algorithm was used with the following variables modifications: methionine oxidation, phosphorylation on S/T/Y, deamidation on N/D, carbamidomethyl cysteine was selected as a fixed modification. A digestion enzyme of trypsin was set allowing up to three missed cleavages. The data were searched with a parent ion tolerance of 5 ppm and a fragment ion tolerance of 0.5 Da. In order to validate protein identifications, the MASCOT result files are loaded in Scaffold 4.3.4 (Proteome Software, Portland, Oregon, USA) (Searle, 2010). Protein lists are generated of Scaffold 4.3.4 based on the following criteria for the identification of proteins: peptide identification was accepted with a greater probability of 80%. In addition, we

accept the identification of proteins with a greater probability of 80% and containing at least one peptide identified. These filters are used, because no significant results were obtained with higher filters. Nevertheless, the relaxation used here does not compromise the analysis, since all inferences are performed at the level of types of toxin families and not from specific proteins. Undoubtedly, these results are most promising for future research.

Proteomic data from the literature

The protein lists of previously published studies from *Acropora digitifera* (Gacesa *et al*., 2015) and *Olindias sambaquiensis* (Weston *et al*., 2013) were used in the analyses. Venom proteomes of the three remaining species (*Anemonia viridis, Hydra magnipapillata,* and *Aurelia aurita*) were obtained from the analysis of raw data from the study of Rachamim and co-workers, 2014. The raw data of this work were obtained by courtesy of Daniel Sher.

Character mapping analysis

This procedure aims to distinguish between different recruitment patterns of each type of toxin protein family in the venom proteome. Data were coded in a matrix of presence (1) or absence (0) of each type of toxin protein family in each species. The reconstruction of ancestral states at different nodes in the topology was performed in Mesquite version 3.04 (Maddison & Maddison, 2015) using the parsimony criterion under the model unordered. This method finds the ancestral states that minimize the number of steps of character change given the tree (in this case pre-established according to previous studies of Marques & Collins, 2004; Collins *et al*., 2006) and observed character distribution.

RESULTS

Proteomic characterisation of venom components in the species studied

Three hundred and eighty proteins were identified of lists generated from the analysis and reanalysis from MS/MS data. The number and details of identified proteins in the nematocyst payload of each organism is shown in the supplementary material (Tables S1-S8). All proteins were assigned to twenty-eigth types of toxin families (Fig. 1, Table 1).

Reconstruction of expression and recruitment patterns of the different toxin families of the cnidarian venom proteome

Fifteen out of the 28 toxin protein families found would correspond to events of single early recruitment (Fig. 2, pattern i). These protein families were recruited, or acquired, early (i.e. before the split between the two subphyla Anthozoa-Medusozoa) in the evolution of the venom proteome of phylum Cnidaria and, subsequently were lost or not in some species of different classes, or even in an entire class. Four toxin protein families were recruited or acquired after the split between Anthozoa and Medusozoa, either in a single species, or sister species or major lineages (like the entire class) (Fig. 2, pattern ii). Nine toxin protein families were recruited or acquired independently in the venom proteome of different species or lineages in the phylum Cnidaria (Fig. 2, pattern iii).

DISCUSSION

Nematocysts containing toxins are a defining morphological characteristic of Cnidaria. To date, transcriptome or proteome data from Cnidaria have identified a number of toxin protein families from discharged nematocysts, especially in species belonging to the subphylum Anthozoa (Anderluh & Maček, 2002; Castañeda & Harvey, 2009; Frazão *et al*., 2012; Jouiaei *et al*., 2015b, 2015c; Macrander *et al*., 2015). In contrast, the venom of nematocysts from the subphylum Medusozoa has been much less well studied, and focused on toxic species for humans (Balasubramanian *et al*., 2012; Brinkman *et al*., 2012, 2015; Weston *et al*., 2013; Badré, 2014; Li *et al*., 2014; Rachamim *et al*., 2014; Jouiaei *et al*., 2015a). This study combines comparative proteomics and character mapping on widely accepted phylogenies to understand the recruitment pattern of 28 toxin families from 8 species and four classes of Cnidaria.

Comparative venom proteomic analysis from species of different Cnidaria classes

Three out of the 8 species examined (*Chrysaora lactea, Tamoya haplonema*, and *Chiropsalmus quadrumanus*) have proteomes described for the first time. Venom data from three species (*Anemonia viridis*, *Hydra magnipapillata*, and *Aurelia aurita*) were published elsewhere (Rachamim *et al*., 2014) and reassessed in this study. These data were combined with our previous published data from *Acropora digitifera* (Gacesa *et al*., 2015) and *Olindias sambaquiensis* (Weston *et al*., 2013), revealing that both

Anthozoa and Medusozoa have complex venoms comprising multiple toxin protein families.

Twelve (~43%) out of the 28 toxin protein families were shared by at least one species in each of the four classes of cnidarians: (Phospholipases A_2 (PLA_2), Phospholipases D, Metalloproteinase, Peptidase S1, Potassium channel inhibitor, Sodium channel inhibitor, complement C3, Conotoxins, Flavin monoamine oxidase, Glycosyl hydrolase 56, Latrotoxin and Snake three finger; Fig.1, Table 1).

Although there are many PLA_2 proteins with non-toxic physiological functions (Six & Dennis, 2000), PLA_2 has been convergently recruited into the venoms of many animal lineages (e.g., reptiles, centipedes, insects, arachnids, cephalopods, and cnidarians, Fry *et al.*, 2009), with many diverse toxic functions for example, like neurotoxic, myotoxic, haemolytic. In Cnidaria, PLA_2 toxins have so far been identified only with haemolytic activity (Hessinger & Lenhoff, 1976; Grotendorst & Hessinger, 2000; Anderluh & Maček, 2002; Talvinen & Nevalainen, 2002; Nevalainen *et al*., 2004; Razpotnik *et al*., 2010). In addition to PLA_2, we also identified phospholipases D family-like proteins in at least one species from the four cnidarian classes (Table 1). Recently, transcripts with similarity to phospholipase D were identified in the transcriptome of the giant jellyfish *Cyanea capillata* (Liu *et al*., 2015). In the venom of brown spiders exhibits a necrotic effect (Chaim *et al*., 2011). Also, this effect has been reported in cnidarians (Burke, 2002; Uri *et al*., 2005).

Most of the metalloproteinases identified in this study belonged to the zinc metalloproteinase family. This family of toxins is an important component found also in the venoms of terrestrial animals such as centipedes, snakes and ticks (Fry *et al.*, 2009; Undheim *et al*., 2014), with a wide range of diverse biological activities culminating in haemorrhage and tissue necrosis (Fox & Serrano, 2005; da Silveira *et al*., 2007). Transcriptomic and proteomic studies have identified zinc metalloprotease in the venoms of the scyphozoans *Stomolophus meleagris* and *Cyanea capillata* (Li *et al*., 2014; Liu *et al*., 2015), the cubozoan *Chironex fleckeri* (Brinkman *et al*., 2015; Jouiaei *et al*., 2015a), and the anthozoan *Anthopleura elegantissima* (Macrander *et al*., 2015). A study of metalloproteases from the scyphozoan *Nemopilema nomurai*, *Rhopilema esculenta*, *Cyanea nozakii*, and *Aurelia aurita* confirmed that the toxicity of these enzymes are related with proteolytic effects (Lee *et al*., 2011).

The peptidase S1 family was the third toxin family found in all cnidarian classes studied. This family is part of the group of serine protease inhibitors and is widely distributed in marine venomous animals including cone snails and cephalopods (Mourão & Schwartz, 2013), as well as terrestrial reptiles (Fry e*t al*., 2009). Recently, transcripts with similarity to serine proteases were identified in the transcriptome of the sea anemone *Anthopleura elegantissima* (Macrander *et al*., 2015). However, to date, the S1 peptidase family of toxins has never been functionally described for cnidarians.

Both sodium (Na) and potassium (K) ion channel inhibitors were identified in all classes examined. These two types of neurotoxins have been widely studied in Anthozoa, especially sea anemones (Moran *et al*., 2009; Šuput, 2009; Turk & Kem, 2009; Frazão *et al*., 2012; Jouiaei *et al*., 2015c; Macrander *et al*., 2015; Mariottini *et al*., 2015). No sea anemones proteomes were included in our analysis. Although neurotoxic effects have been identified in scyphozoans such as *Cyanea nozakii* (Feng *et al*., 2010), *Cyanea capillata* (Helmholz *et al*., 2012), and *Pelagia noctiluca* (Pang *et al*., 1993; Morabito *et al*., 2012), and also in cubozoans such as *Carukia barnesi* (Winkel et al., 2005) and *Malo kingi* (Gershwin, 2007), their toxicities have not been attributed to sodium (Na) and potassium (K) ion channel inhibition. This is the first report of these toxins in non-Anthozoan cnidarians.

Complement C3 family-like proteins were identified in at least one species from all cnidarian classes (Table 1). This protein family was also identified in the venom proteome of the cubozoan *Chironex fleckeri* (Brinkman *et al*., 2015). The toxic effects of Complement C3 family-like proteins is unknown but the presence of these proteins in venoms, that are normally associate with innate immune response, is a fascinating avenue for future research.

Five other toxin protein families were also reported for the first time in Cnidaria. These toxin families included neurotoxins related to three animal lineages: conotoxins (Olivera, 2002), three finger (Fry *et al*., 2003; Kini & Doley, 2010), and latrotoxins (Garb & Hayashi, 2013). Likewise, hyaluronidase-like proteins were also found, but these proteins are common and have non-toxic physiological function in many non-venomous animals. Hence, the presence of this protein does not represent a venom toxin itself, but it is possibly recruited into venoms to increase tissue permeability, making the dispersion of toxins more efficient (Kemparaju & Girish, 2006; Fry *et al*., 2009). The

flavin amino-oxidase family-like proteins (Guo *et al*., 2012) have previously been found extensively in many snakes venoms and display mainly haemolytic activities.

Sixteen protein toxin families were not distributed across all classes. These included three families of pore forming toxins, which was unexpected because this biological activity has been widely studied in Cnidaria (Badré, 2014). We found jellyfish toxin (JFT) family-like proteins in the venom proteomes of the sister classes Cubozoa and Scyphozoa. The JFT family was originally described in the cubozoan *Carybdea alata* and designated CAH1 (Chung *et al*., 2001). Subsequently, JFT family was also identified in many other cubozoans, including *Chironex fleckeri* (Brinkman & Burnell, 2007, 2009; Brinkman *et al*., 2014, 2015), *Carybdea rastonii* (Nagai *et al*., 2000), and *Chiropsoides quadrigatus* (Nagai *et al*., 2002c, as *Chiropsalmus quadrigatus*). Homologues of JFT family have also been identified in the genome, transcriptome and proteome of the scyphozoan *Aurelia aurita* (Brinkman et al., 2014; Rachamim *et al*., 2014), as well as the hydrozoans *Hydra magnipapillata* and *Hydra vulgaris* (Brinkman et al., 2014; Rachamim et al., 2014). Recently, transcriptomic data has also suggested that the JFT family may also be found in the sea-anemones *Aiptasia pallida* and *Anemonia viridis* (Jouiaei et al., 2015b; Rachamim *et al*., 2014). This suggest that JFTs may have originated from the common ancestor of all extant cnidarians, maybe 600 Mya. However, expression patterns have changed over time. This may be reflected in the wide mechanisms of action attributed to JFTs, which are amongst the most dangerous toxins secreted by Medusozoa (Mariottini & Pane, 2013; Badré, 2014; Brinkman et al., 2014; Tibballs *et al*., 2011).

Another family of cytolytic peptides common to cnidarian venoms, especially in Anthozoa, are the actinoporins (Anderluh & Maček, 2002; Mariottini & Pane, 2013). Actinoporins have also been identified in the genome, transcriptome and proteome of the *Hydra magnipapillata* (Glasser *et al*., 2014; Rachamim *et al*., 2014). We identified peptides similar to actinoporins only in the class Cubozoa (Fig. 1, Table 1). Another family of toxic proteins that form pores are membrane attack complex (MAC) proteins. We identified peptides similar to MAC proteins in *Olindias sambaquiensis.* Similar to C3 complement proteins, it is unclear why proteins commonly associated with innate immune responses are also apparently widely distributed in cnidarian venoms. For example, MAC-like toxins have also been identified in sea anemones (Nagai *et al*., 2002b; Oshiro *et al*., 2004) and were recently annotated in the transcriptomes and

proteomes of Hydrozoa and Scyphozoa (Rachamim *et al*., 2014). Other protein toxin families reported to display cytolytic activity identified herein were the latarcins, although not present in the Hydrozoa (Dubovskii *et al*., 2015), and phospholipase B (Bernheimer *et al*., 1986, 1987).

Two lectin families were identified in the cnidarian venom proteomes. C-type lectins were found in one species of Anthozoa, Cubozoa, and Scyphozoa studied, while ficolins were found in species of Cubozoa and Scyphozoa only. Homologous proteins of both these families have only been found previously in the proteome and transcriptome of the scyphozoan *Stomolophus meleagris* (Li *et al*., 2014). Homologues of these families have been widely reported in the venom of snakes and lizards, attributing strong haemolytic activity (Lu *et al*., 2005; Fry *et al*., 2010; Öhler *et al*., 2010).

Peptides with homology to four families of neurotoxins were also restricted to some of the cnidarian classes (Fig. 1, Table 1). The Kunitz type family was expressed in the venom proteome of anthozoan species and in *Aurelia aurita*. Kunitz type family-like proteins have been found in both classes of cnidarians in previous studies (Minagawa *et al*., 1997, 2008; Li *et al*., 2014; Macrander *et al*., 2015). Another family of neurotoxins were the SCRIPs type toxins, found in all classes of cnidarians studied here except Cubozoa, although the toxin has been previously annotated in the proteome of *Chironex fleckeri* (Brinkman *et al.*, 2015). Two protein families reported to be neurotoxic in other venomous animals are reported here for the first time in cnidarians. Magi-1 was identified in Anthozoa, Cubozoa, and Hydrozoa classes (Fig. 1, Table 1). This protein family was first isolated from the venom of Hexathelidae spider. These insecticidal toxins bind to sodium channels and induce flaccid paralysis when injected into lepidopteran larvae (Corzo *et al*., 2003). The other toxin family, called Huwentoxin-1, was detected only in medusozoans (Fig. 1, Table 1). This is a Ca^{2+} channel inhibitor (Diao *et al*., 2003) and is a lethal neurotoxin that binds to the nicotinic acetylcholine receptor and blocks neuromuscular transmission.

Homology to the peroxiredoxin-4 protein from the AhpC/TSA family was found in all but Hydrozoa species (Fig. 1, Table 1). Ruan *et al*., (2014) reported the first peroxiredoxin homologue in the scyphozoan *Cyanea capillata*, attributing to that strong antioxidant functions. Recently, two homologous proteins were detected in the venom proteome from *Chironex fleckeri* (Brinkman *et al*., 2015).

Finally, it is interesting to note the presence of translationally controlled tumour like proteins (TCTP) into venom proteome from both Anthozoa and Hydrozoa (Fig. 1, Table 1). Homologues of this group have been found previously in *Hydra vulgaris* (Yan *et al*., 2000) and recently in *Cyanea capillata* (Liu, *et al*., 2015) but have not been associated with any toxic function. The Natriuretic family of toxic peptide family were identified in the venom proteome of Cubozoa and Scyphozoa (Fig. 1, Table 1). This family has also been recruited into platypus (de Plater *et al*., 1998) and reptile (Fry *et al*., 2005, 2006) venoms and are potent hypotensive toxins.

Taking together the taxonomic distribution and predicted biological activities of the toxins described in this study, one can conclude that the venomous arsenal of the cnidarian subphyla and classes (Fig. 1, Table 1) demonstrate broad convergence on across other animal taxa. Venoms of Medusozoa and Anthozoa do differ in that Anthozoa venoms are composed of more neurotoxins when compared to Medusozoa venoms, in which these toxins are few or absent (Rachamim *et al*., 2014; Brinkman *et al*., 2015). In contrast, cytolytic toxin families are more abundant in Medusozoa (cf. Rachamim *et al*., 2014; Brinkman *et al*., 2015; Jouiaei *et al*., 2015c).

Evolution of the cnidarian venom arsenal

Recruitment patterns of toxin protein families (Fig. 3, Table 2) suggest that the venoms of Medusozoa and Anthozoa ancestors might have been composed of at least fifteen types of toxin families. This shows that many of the toxins families currently identified in cnidarian venoms were recruited in the venom proteome very early in the evolution of the phylum (Figs. 2i, 3). These first venoms probably already had neurotoxic, cytotoxic and others activities, being similar to venom of extant cnidarians.

After separation of the ancestral lineage into Anthozoa and Medusozoa, the expression of certain toxins was interrupted in some clades. For example, the AhpC/TSA family, including peroxiredoxin-4 protein, is not found in any hydrozoan members examined to date. Likewise, the toxin type SCRIPs was lost from the class Cubozoa. In contrast, data from this study suggests that four families of cytolytic toxins were recruited into Medusozoa after the basal diversification event (Fig. 3). For example, MACPF was recruited in the venom proteome of Hydrozoa, supposedly autopomorphic to *Olindias sambaquiensis*. Actinoporins were recruited as a synapomorphy of Cubozoa, present in the two species studied. Lectins and jellyfish

toxin were recruited into venom proteomes of both Cubozoa and Scyphozoa (Figs. 2ii, 3). The Snaclec and Kunitz types of toxin families were recruited independently in two classes, respectively (Figs. 2iii, 3).

This is the most comprehensive comparative study of cnidarian venom composition to date. Our approach is conservative in that the analyses are exclusively based on types of toxin families found in each proteomic profile, and not on specific toxins. Toxic profiles or lists of proteins (Tables S1-S8) are phenotypes, or "morphological representation" of the venom, acting as a character. Toxin composition can vary due to different biotic and abiotic factors (Mackessy, 2009; Casewell *et al*., 2013; Starcevic & Long, 2013). For example, in this study JFTs were found only in the venoms of one Scyphozoa and Cubozoa species. However, previous studies have demonstrated JFTs in the genome and proteome of Hydrozoa, and in the transcriptome of the anthozoan *Anemonia viridis* (Rachamim *et al*., 2014). Patterns like this may suggest either the potential for variation in the venom composition at various levels, such as among major and minor taxonomic groups, or between different parts of the same population of a given species, or during different ages of the animal as well as between different stages of the lifecycle, besides several other factors (Mackessy, 2009). Nevertheless, few reports in the literature have documented variation in the composition of the venom at any level in the phylum Cnidaria (Orts *et al*., 2013; Rachamim *et al*., 2014). Certainly, no study to date has attempted to put into context what the biological consequences of venom variation might be in Cnidaria. In this study, more closely related classes of cnidarians generally tend to have more similar venoms in composition than do more distantly related animals, despite possessing a repertoire of functionally similar toxin protein families (Figs. 1, 3, Table 1). This study provides new insights into the evolutionary assembly of the complex biochemical arsenal of cnidarians and provides a partial view of the composition of the ancestral first venoms in the basal Metazoa. But processes and mechanisms of toxin diversification in venom composition and the functional context that results has yet to be established.

REFERENCES

Anderluh, G. & Maček, P. 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). Toxicon 40: 111–124.

Anderluh, G. & Maček, P. 2003. Dissecting the actinoporin pore-forming mechanism. Structure 11: 1312–1313.

Badré, S. 2014. Bioactive toxins from stinging jellyfish. Toxicon 91: 114–125.

Balasubramanian, P.G.; Beckmann, A.; Warnken, U.; Schnölzer, M.; Schüler, A.; Bornberg-Bauer, E.; Holstein, T.W. & Özbek, S. 2012. Proteome of *Hydra* nematocyst. Journal of Biological Chemistry 287: 9672–9681.

Beckmann, A. & Özbek, S. 2012. The nematocyst: A molecular map of the cnidarian stinging organelle. International Journal of Developmental Biology 56: 577–582.

Bernheimer, A.W.; Linder, R.; Weinstein, S. & Kim, K.S. 1987. Isolation and characterization of a phospholipase B from venom of Collett's snake, *Pseudechis colletti*. Toxicon 25: 547–554.

Bernheimer, A.W.; Weinstein, S. & Linder, R. 1986. Isoelectric analysis of some Australian elapid snake venoms with special reference to phospholipase b and haemolysis. Toxicon 24: 841–849.

Bigger, C.H. 1980. Interspecific and intraspecific acrorhagial aggressive behaviour among sea anemones: a recognition of self and not self. Biological Bulletin 159: 117–134.

Birck, C.; Damian, L.; Marty-Detraves, C.; Lougarre, A.; Schulze-Briese, C.; Koehl, P.; Fournier, D.; Paquereau, L. & Samama, J. P. 2004. A new lectin family with structure similarity to actinoporins revealed by the crystal structure of *Xerocomus chrysenteron* lectin XCL. Journal of Molecular Biology 344: 1409–1420.

Bridge, D.; Cunningham, C.W.; DeSalle, R. & Buss, L.W. 1992. Class-level relationships in the phylum Cnidaria: Evidence from mitochondrial genome structure. Proceedings of the National Academy of Sciences of the United States of America 89: 8750–8753.

Brinkman, D.L. & Burnell, J. 2007. Identification, cloning and sequencing of two major venom proteins from box jellyfish, *Chironex fleckeri*. Toxicon 50: 850–860.

Brinkman, D.L.; Aziz, A.; Loukas, A.; Potriquet, J.; Seymour, J. & Mulvenna, J. 2012. Venom proteome of the box jellyfish *Chironex fleckeri*. PLoS ONE 7 (12): e47866.

Brinkman, D.L. & Burnell, J.N. 2009. Biochemical and molecular characterisation of cubozoan protein toxins. Toxicon 54: 1162–1173.

Brinkman, D.L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D. & Mulvenna, J. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16: 1–15.

Brinkman, D.L.; Konstantakopoulos, N.; McInerney, B.V.; Mulvenna, J.; Seymour, J.E.; Isbister, G.K. & Hodgson, W.C. 2014. *Chironex fleckeri* (box jellyfish) venom proteins: Expansion of a cnidarian toxin family that elicits variable cytolytic and cardiovascular effects. Journal of Biological Chemistry 289: 4798–4812.

Burke, W.A. 2002. Cnidarians and human skin. Dermatologic Therapy 15:18–25.

Calvete, J.J.; Fasoli, E.; Sanz, L.; Boschetti, E. & Righetti, P.G. 2009. Exploring the venom proteome of the western diamondback rattlesnake, *Crotalus atrox*, via snake venomics and combinatorial peptide ligand library approaches. Journal of Proteome Research 8: 3055–3067.

Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A. & Fry, B.G. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends in Ecology and Evolution 28: 219–229.

Castañeda, O. & Harvey, A.L. 2009. Discovery and characterization of cnidarian peptide toxins that affect neuronal potassium ion channels. Toxicon 54: 1119–1124.

Chaim, O.M.; Trevisan-Silva, D.; Chaves-Moreira, D.; Wille, A.C.; Ferrer, V.P.; Matsubara, F.H.; Mangili, O.C.; da Silveira, R.B.; Gremski, L.H.; Senff-Ribeiro, A. & Veiga, S.S. 2011. Brown spider (Loxosceles genus) venom toxins: tools for biological purposes. Toxins 3:309–344.

Chung, J.J.; Ratnapala, L.A.; Cooke, I.M. & Yanagihara, A.A. 2001. Partial purification and characterization of a haemolysin (CAH1) from Hawaiian box jellyfish (*Carybdea alata*) venom. Toxicon 39: 981–990.

Collins, A.G. 2002. Phylogeny of Medusozoa and the evolution of cnidarian life cycles. Journal of Evolutionary Biology 15: 418–432.

Collins, A.G.; Schuchert, P.; Marques, A.C.; Jankowski, T.; Medina, M. & Schierwater, B. 2006. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. Systematic Biology 55: 97–115.

Collins, A.G.; Bentlage, B.; Gillan, W.; Lynn, T.H.; Morandini, A.C. & Marques, A.C. 2011. Naming the Bonaire banded box jelly, *Tamoya ohboya*, n. sp. (Cnidaria: Cubozoa: Carybdeida: Tamoyidae). Zootaxa, 2753: 53-68.

Corzo, G.; Gilles, N.; Satake, H.; Villegas, E.; Dai, L.; Nakajima, T. & Haupt, J. 2003. Distinct primary structures of the major peptide toxins from the venom of the spider *Macrothele gigas* that bind to sites 3 and 4 in the sodium channel. FEBS Letters 547: 43–50.

Da Silveira, R.B.; Wille, A.C.M.; Chaim, O.M.; Appel, M.H.; Silva, D.T.; Franco, C.R.C.; Toma, L.; Mangili, O.C.; Gremski, W.; Dietrich, C.P.; Nader, H.B. & Veiga, S.S. 2007. Identification, cloning, expression and functional characterization of an astacin-like metalloprotease toxin from *Loxosceles intermedia* (brown spider) venom. The Biochemical Journal 406: 355–363.

Daly, M.; Brugler, M.R.; Cartwright, P.; Collins, A.G.; Dawson, M.N.; Fautin, D.G. France, S.C.; McFadden, C.S.; Opresko, D.M.; Rodriguez, E.; Romano, S.L. & Stake, J.L. 2007. The phylum Cnidaria: A review of phylogenetic patterns and diversity 300 years after Linnaeus. Zootaxa 182: 127–182.

David, C.N.; Özbek, S.; Adamczyk, P.; Meier, S.; Pauly, B.; Chapman, J.; Hwang, J.S.; Gojobori, T. & Holstein, T.W. 2008. Evolution of complex structures: minicollagens shape the cnidarian nematocyst. Trends in Genetics 24: 431–438.

De Plater, G.M.; Martin, R.L. & Milburn, P.J. 1998. A C-type natriuretic peptide from the venom of the platypus (*Ornithorhynchus anatinus*): structure and pharmacology. Comparative Biochemistry and Physiology part C: Pharmacology, Toxicology and Endocrinology 120: 99–110.

Diao, J.; Lin, Y.; Tang, J. & Liang, S. 2003. cDNA sequence analysis of seven peptide toxins from the spider *Selenocosmia huwena*. Toxicon 42: 715–723.

Dubovskii, P.V.; Vassilevski, A.A.; Kozlov, S.A.; Feofanov, A.V.; Grishin, E.V. & Efremov, R.G. 2015. Latarcins: versatile spider venom peptides. Cellular and Molecular Life Sciences. 1–22.

Fautin, D.G. 2009. Structural diversity, systematics, and evolution of cnidae. Toxicon 54: 1054–1064.

Feng, J.; Yu, H.; Xing, R.; Liu, S.; Wang, L.; Cai, S.& Li, P. 2010. Isolation and characterization of lethal proteins in nematocyst venom of the jellyfish *Cyanea nozakii* Kishinouye. Toxicon 55: 118–125.

Fox, J.W. & Serrano, S.M.T. 2005. Structural considerations of the snake venom metalloproteinases, key members of the M12 reprolysin family of metalloproteinases. Toxicon 45: 969–985.

Frazão, B.; Vasconcelos, V. & Antunes, A. 2012. Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: An overview. Marine Drugs 10: 1812–1851.

Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.; Lewis, R.J.; Norton, R.S.; Renjifo, C. & de la Vega, R.C.R. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10: 483–511.

Fry, B.G.; Vidal, N.; Norman, J.; Vonk, F.J.; Scheib, H.; Ramjan, S.F.R.; Kuruppu, S.; Fung, K.; Hedges, S.B.; Richardson, M.K.; Hodgson, W.C.; Ignjatovic, V. Summerhayes, R. & Kochva, E. 2006. Early evolution of the venom system in lizards and snakes. Nature 439: 584–588.

Fry, B.G.; Wickramaratana, J.C.; Lemme, S.; Beuve, A.; Garbers, D.; Hodgson, W.C. & Alewood, P. 2005. Novel natriuretic peptides from the venom of the inland taipan (*Oxyuranus microlepidotus*): Isolation, chemical and biological characterisation. Biochemical and Biophysical Research Communications 327: 1011–1015.

Fry, B.G.; Winter, K.; Norman, J.; Roelants, K.; Nabuurs, R.J.; van Osch, M.J.P.; Teeuwisse, W.M.; van der Weerd, L; McNaughtan, .J.E.; Kwok, H.F.; Scheib, H.; Greisman, L.; Kochva, E.; Miller, L.J.; Gao, F.; Karas, J.; Scanlon, D.; Lin, F.; Kuruppu, S.; Shaw, C.; Wong, L. & Hodgson, W.C. 2010. Functional and structural diversification of the *Anguimorpha* lizard venom system. Molecular & Cellular Proteomics 9: 2369–2390.

Fry, B.G.; Wüster, W.; Kini, R.M.; Brusic, V.; Khan, A.; Venkataraman, D. & Rooney, A.P. 2003. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. Journal of Molecular Evolution 57: 110–129.

Garb, J.E. & Hayashi, C.Y. 2013. Molecular evolution of α-latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. Molecular Biology and Evolution 30: 999–1014.

Gacesa, R.; Chung, R.; Dunn, S.R.; Weston, A.; Jaimes-Becerra, A.; Marques, A.C.; Morandini, A.; Hranueli, D.; Starcevic, A.; Ward, M. & Long, P.F. 2015. Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from Acropora digitifera (Cnidaria: Anthozoa: Scleractinia). BMC Genomics 16:774.

Gershwin, L.A. 2007. *Malo kingi*: A new species of Irukandji jellyfish (Cnidaria: Cubozoa: Carybdeidae), possibly lethal to humans, from Queensland, Australia. Zootaxa 68: 55–68.

Glasser, E.; Rachamim, T.; Aharonovich, D. & Sher, D. 2014. Hydra actinoporin-like toxin-1, an unusual haemolysin from the nematocyst venom of *Hydra magnipapillata* which belongs to an extended gene family. Toxicon 91 103–113.

Grotendorst, G.R. & Hessinger, D.A. 2000. Enzymatic characterization of the major phospholipase A2 component of sea anemone (*Aiptasia pallida*) nematocyst venom. Toxicon 38: 931–943.

Guo, C.; Liu, S.; Yao, Y.; Zhang, Q. & Sun, M.Z. 2012. Past decade study of snake venom l-amino acid oxidase. Toxicon 60: 302–311.

Hargreaves, A.D.; Swain, M.T.; Logan, D.W. & Mulley, J.F. 2014. Testing the Toxicofera: comparative reptile transcriptomics casts doubt on the single, early evolution of the reptile venom system. Toxicon 92:140-156.

Helmholz, H.; Wiebring, A.; Lassen, S.; Ruhnau, C. Schuett, C. & Prange, A. 2012. Cnidom analysis combined with an in vitro evaluation of the lytic, cyto- and neurotoxic potential of *Cyanea capillata* (Cnidaria: Scyphozoa). Scientia Marina 76: 339–348.

Hessinger, D.A. & Lenhoff, H.M. 1976. Mechanism of haemolysis induced by nematocyst venom: Roles of phospholipase A and direct lytic factor. Archives of Biochemistry and Biophysics 173: 603–613.

Jouiaei, M.; Casewell, N.; Yanagihara, A.; Nouwens, A.; Cribb, B.; Whitehead, D.; Jackson, T.; Ali, S.; Wagstaff, S.; Koludarov, I.; Alewood, P.; Hansen, J. & Fry, B. 2015a. Firing the sting: Chemically induced discharge of cnidae reveals novel proteins and peptides from box jellyfish (*Chironex fleckeri*) venom. Toxins 7: 936–950.

Jouiaei, M.; Sunagar, K.; Gross, A.F.; Scheib, H.; Alewood, P.F.; Moran, Y. & Fry, B.G. 2015b. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Molecular Biology and Evolution 32: 1598–1610.

Jouiaei, M.; Yanagihara, A.; Madio, B.; Nevalainen, T.; Alewood, P. & Fry, B. 2015c. Ancient venom systems: A review on Cnidaria toxins. Toxins 7: 2251–2271.

Jungo, F., L. Bougueleret, I. Xenarios, & S. Poux, 2012. The UniprotKB/Swiss-prot Tox-Prot program: a central hub of integrated venom protein data. Toxicon 60: 551–557.

Kass-Simon, G. & Scappaticci, J. 2002. The behavioural and developmental physiology of nematocysts. Canadian Journal of Zoology 80: 1772–1794.

Kawashima, Y.; Nagai, H.; Ishida, M.; Nagashima, Y. & Shiomi, K. 2003. Primary structure of echotoxin 2, an actinoporin-like haemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*. Toxicon 42: 491–497.

Kayal, E.; Bentlage, B.; Collins, A.G.; Kayal, M.; Pirro, S. & Lavrov, D.V. 2012. Evolution of linear mitochondrial genomes in medusozoan cnidarians. Genome Biology and Evolution 4: 1–12.

Kemparaju, K. & Girish, K.S. 2006. Snake venom hyaluronidase: A therapeutic target. Cell Biochemistry and Function 24: 7–12.

Kini, R.M. & Doley, R. 2010. Structure, function and evolution of three-finger toxins: Mini proteins with multiple targets. Toxicon 56: 855–867.

Lee, H.; Jung, E.S.; Kang, C.; Yoon, W.D.; Kim, J.S. & Kim, E. 2011. Scyphozoan jellyfish venom metalloproteinases and their role in the cytotoxicity. Toxicon 58: 277–284.

Li, R.; Yu, H.; Xing, R.; Liu, S.; Qing, Y.; Li, K.; Li, B. Meng, X.; Cui, J. & Li, P. 2012. Application of nanoLC-MS/MS to the shotgun proteomic analysis of the nematocyst proteins from jellyfish *Stomolophus meleagris*. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 899: 86–95.

Li, R.; Yu, H.; Xue, W.; Yue, Y.; Liu, S.; Xing, R. & Li, P. 2014. Jellyfish venomics and venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. Journal of Proteomics 106: 17–29.

Liu, G.; Zhou, Y.; Liu, D.; Wang, Q.; Ruan, Z.; He, Q. & Zhang, L. 2015. Global Transcriptome Analysis of the Tentacle of the Jellyfish Cyanea capillata Using Deep Sequencing and Expressed Sequence Tags: Insight into the Toxin- and Degenerative Disease-Related Transcripts. PLoS ONE 10(11): e0142680.

Lotan, A.; Fishman, L.; Loya, Y. & Zlotkin, E. 1995. Delivery of a nematocyst toxin. Nature 375: 456.

Lu, Q.; Navdaev, A.; Clemetson, J.M. & Clemetson, K.J. 2005. Snake venom C-type lectins interacting with platelet receptors. Structure-function relationships and effects on haemostasis. Toxicon 45: 1089–1098.

Mackessy, S.P. 2009. The field of reptile toxinology: snakes, lizards and their venoms. In: Mackessy, S.P. (Ed.), Handbook of venoms and toxins of reptiles. CRC Press/Taylor & Francis Group, Boca Raton, FL 3–23.

Macrander, J.; Brugler, M.R. & Daly, M. 2015. A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps. BMC Genomics 16: 1–19.

Maddison, W.P. & Maddison, D.R. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04 http://mesquiteproject.org. Elsevier.

Mariottini, G.L. 2014. Haemolytic venoms from marine cnidarian jellyfish – an overview. Journal of Venom Research 5: 22–32.

Mariottini, G.L.; Bonello, G.; Giacco, E. & Pane, L. 2015. Neurotoxic and neuroactive compounds from Cnidaria: Five decades of research…. and more. Central Nervous System Agents in Medicinal Chemistry 15: 74–80.

Mariottini, G.L. & Pane, L. 2013. Cytotoxic and cytolytic cnidarian venoms. A review on health implications and possible therapeutic applications. Toxins 6: 108–151.

Maronna, M. M.; Miranda, T. P.; Peña-Cantero, A. L.; Barbeitos, M.S. & Marques, A. C. 2015. Towards a phylogenetic classification of Leptothecata (Cnidaria, Hydrozoa). Scientific Reports.

Marques, A.C. & Collins, A.G. 2004. Cladistic analysis of Medusozoa and cnidarian evolution. Invertebrate Biology 123: 23–42.

Minagawa, S.; Ishida, M.; Shimakura, K.; Nagashima, Y. & Shiomi, K. 1997. Isolation and amino acid sequences of two Kunitz-type protease inhibitors from the sea anemone *Anthopleura* aff. *xanthogrammica*. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology 118: 381–386.

Minagawa, S.; Sugiyama, M.; Ishida, M.; Nagashima, Y. & Shiomi, K. 2008. Kunitz-type protease inhibitors from acrorhagi of three species of sea anemones. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology 150: 240–245.

Morabito, R.; Condello, S.; Currò, M.; Marino, A.; Ientile, R. & La Spada, G. 2012. Oxidative stress induced by crude venom from the jellyfish *Pelagia noctiluca* in neuronal-like differentiated SH-SY5Y cells. Toxicology in Vitro 26: 694–699.

Moran, Y.; Gordon, D. & Gurevitz, M. 2009. Sea anemone toxins affecting voltage-gated sodium channels - molecular and evolutionary features. Toxicon 54: 1089–1101.

Moran, Y.; Praher, D.; Schlesinger, A.; Ayalon, A.; Tal, Y. & Technau, U. 2013. Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. Marine Biotechnology 15: 329–339.

Morandini, A.C.; Ascher, D.; Stampar, S.N. & Ferreira, J.F.V. 2005. Cubozoa e Scyphozoa (Cnidaria: Medusozoa) de águas costeiras do Brasil. Iheringia. Série Zoologia, 95(3): 281-294.

Morandini, A.C. & Marques, A.C. 2010. Revision of the genus Chrysaora Péron & Lesueur, 1810 (Cnidaria: Scyphozoa). Zootaxa, 2464: 1-97.

Mourão, C.B.F. & Schwartz, E.F. 2013. Protease inhibitors from marine venomous animals and their counterparts in terrestrial venomous animals. Marine Drugs 11: 2069–2112.

Nagai, H.; Oshiro, N.; Takuwa-Kuroda, K.; Iwanaga, S.; Nozaki, M. & Nakajima, T. 2002a. A new polypeptide toxin from the nematocyst venom of the Okinawan sea anemone *Phyllodiscus semoni* (Japanese name "unbacho-isoginchaku"). Bioscience, Biotechnology, and Biochemistry 66: 2621–2625.

Nagai, H.; Oshiro, N.; Takuwa-Kuroda, K.; Iwanaga, S.; Nozaki, M. & Nakajima, T. 2002b. Novel proteinaceous toxins from the nematocyst venom of the Okinawan sea anemone *Phyllodiscus semoni* Kwietniewski. Biochemical and Biophysical Research Communications 294: 760–763.

Nagai, H.; Takuwa, K.; Nakao, M.; Ito, E.; Miyake, M.; Noda, M. & Nakajima, T. 2000. Novel proteinaceous toxins from the box jellyfish (sea wasp) *Carybdea rastoni*. Biochemical and Biophysical Research Communications 275: 582–588.

Nagai, H.; Takuwa-Kuroda, K.; Nakao, M.; Oshiro, N.; Iwanaga, S. & Nakajima, T. 2002c. A novel protein toxin from the deadly box jellyfish (sea wasp, Habu-kurage) *Chiropsalmus quadrigatus*. Bioscience, Biotechnology, and Biochemistry 66: 97–102.

Nelsen, D.R.; Nisani, Z.; Cooper, A.M.; Fox, G.A.; Gren, E.C.K. Corbit, A.G. & Hayes, W.K. 2014. Poisons, toxungens, and venoms: Redefining and classifying toxic biological secretions and the organisms that employ them. Biological Reviews 89: 450–465.

Nevalainen, T.J.; Peuravuori, H.J.; Quinn, R.J.; Llewellyn, L.E.; Benzie, J.H.; Fenner, P.J. & Winkel, K.D. 2004. Phospholipase A2 in Cnidaria. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology 139: 731–735.

Nüchter, T.; Benoit, M.; Engel, U.; Ozbek, S. & Holstein, T. 2006. Nanosecond-scale kinetics of nematocyst discharge. Current biology 16: R316–R318.

Öhler, M.; Georgieva, D.; Seifert, J.; Von Bergen, M.; Arni, R.K.; Genov, N. & Betzel, C. 2010. The venomics of *Bothrops alternatus* is a pool of acidic proteins with predominant haemorrhagic and coagulopathic activities. Journal of Proteome Research 9: 2422–2437.

Okamura, B.; Gruhl, A. & Reft A.J. 2015. Cnidarian origins of Myxozoa, pp. 45-68. In: Myxozoan Evolution, Ecology and Development (B Okamura, A Gruhl, JL Bartholomew, eds). Springer, Cham, Switzerland.

Olivera, B.M.; 2002. Conus venom peptides: Reflections from the biology of clades and species. Annual Review of Ecology and Systematics 33: 25–47.

Orts, D.J.B.; Peigneur, S.; Madio, B.; Cassoli, J.S.; Montandon, G.G.; Pimenta, A.M.C.; Bicudo, J.E.P.W.; Freitas, J.C.; Zaharenko, A.J. & Tytgat, J. 2013. Biochemical and electrophysiological characterization of two sea anemone type 1 potassium toxins from a geographically distant population of *Bunodosoma caissarum*. Marine Drugs 11: 655–679.

Oshiro, N.; Kobayashi, C.; Iwanaga, S.; Nozaki, M.; Namikoshi, M.; Spring, J. & Nagai, H. 2004. A new membrane-attack complex/perforin (MACPF) domain

lethal toxin from the nematocyst venom of the Okinawan sea anemone *Actineria villosa*. Toxicon 43: 225–228.

Östman, C. 2000. A guideline to nematocyst nomenclature and classification, and some notes on the systematic value of nematocysts. Scientia Marina 64: 31–46.

Özbek, S.; Balasubramanian, P.G. & Holstein, T.W. 2009. Cnidocyst structure and the biomechanics of discharge. Toxicon 54: 1038–1045.

Pang, K.A. & Schawrtz, M.S. 1993. Guillain-Barré syndrome following jellyfish stings (*Pelagia noctiluca*). Journal of. Neurology, Neurosurgery & Psychiatry 56: 1133–1137.

Peigneur, S.; Billen, B.; Derua, R.; Waelkens, E.; Debaveye, S.; Béress, L. & Tytgat, J. 2011. A bifunctional sea anemone peptide with Kunitz type protease and potassium channel inhibiting properties. Biochemical Pharmacology 82: 81–90.

Perkins, D.N.; Pappin, D.J.C.; Creasy, D.M. & Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20: 3551–3567.

Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T. & Sher, D. 2014. The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. Molecular Biology and Evolution 32: 740–753.

Razpotnik, A.; Križaj, I.; Šribar, J.; Kordiš, D. MačEk, P. Frangež, R.; Kem, W.R. & Turk, T. 2010. A new phospholipase A2 isolated from the sea anemone *Urticina crassicornis* - Its primary structure and phylogenetic classification. FEBS Journal 277: 2641–2653.

Reft, A.J. & Daly, M. 2012. Morphology, distribution, and evolution of apical structure of nematocysts in Hexacorallia. Journal of Morphology 273: 121–136.

Ruan, Z.; Liu, G.; Wang, B.; Zhou, Y.; Lu, J.; Wang, Q.; Zhao, J. & Zhang, L. 2014. First report of a peroxiredoxin homologue in jellyfish: Molecular cloning, expression and functional characterization of CcPrx4 from *Cyanea capillata*. Marine Drugs 12: 214–231.

Searle, B.C. 2010. Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. Proteomics 10: 1265–1269.

Six, D.A. & Dennis, E.A. 2000. The expanding superfamily of phospholipase A2 enzymes: Classification and characterization. Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids 1488: 1–19.

Smith, J.J. & Blumenthal, K.M. 2007. Site-3 sea anemone toxins: Molecular probes of gating mechanisms in voltage-dependent sodium channels. Toxicon 49: 159–170.

Starcevic, A. & Long, P.F. 2013. Diversification of animal venom peptides-were jellyfish amongst the first combinatorial chemists?. ChemBioChem 14: 1407–1409.

Šuput, D. 2009. In vivo effects of cnidarian toxins and venoms. Toxicon 54: 1190–1200.

Talvinen, K. A. & Nevalainen, T.J. 2002. Cloning of a novel phospholipase A2 from the cnidarian *Adamsia carciniopados*. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology 132: 571–578.

Tibballs, J.; Yanagihara, A.; Turner, H. & Winkel, K. 2011. Immunological and toxinological responses to jellyfish stings. Inflammation & Allergy - Drug Targets 10: 438–446.

Turk, T. & Kem, W.R. 2009. The phylum Cnidaria and investigations of its toxins and venoms until 1990. Toxicon 54: 1031–1037.

Undheim, E.B.; Jones, A.; Clauser, K.R.; Holland, J.W; Pineda, S.S.; King, G.F. & Fry, B.G. 2014. Clawing through evolution: Toxin diversification and convergence in the ancient lineage Chilopoda (centipedes). Molecular Biology and Evolution 31: 2124–2148.

Uri, S.; Marina, G. & Liubov, G. 2005. Severe delayed cutaneous reaction due to Mediterranean jellyfish (*Rhopilema nomadica*) envenomation. Contact Dermatitis 52:282–283.

Van Iten, H.; Marques, A.C.; Leme, J.D.M.; Pacheco, M.L.F. & Simões, M.G. 2014. Origin and early diversification of the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's Proterozoic-Cambrian history. Palaeontology 57: 677–690.

Vogel, C.W. & Fritzinger, D.C. 2010. Cobra venom factor: Structure, function, and humanization for therapeutic complement depletion. Toxicon 56: 1198–1222.

Voskoboinik, I.; Dunstone, M.A.; Baran, K.; Whisstock, J.C. & Trapani, J.A. 2010. Perforin: Structure, function, and role in human immunopathology. Immunological Reviews 235: 35–54.

Weston, A.J.; Dunlap, W.C.; Shick, J.M.; Klueter, A.; Iglic, K.; Vukelic, A.; Starcevic, A.; Ward, M.; Wells, M.L.; Trick, C.G. & Long, P.F. 2012. A profile of an endosymbiont-enriched fraction of the coral *Stylophora pistillata* reveals proteins relevant to microbial-host interactions. Molecular & Cellular Proteomics 11. M111.015487.

Weston, A.J.; Chung, R.; Dunlap, W.C.; Morandini, A.C.; Marques, A.C.; Moura-da-Silva, A.M.; Ward, M.; Padilla, G.; da Silva, L.F.; Andreakis, N. & Long, P.F. 2013. Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish *Olindias sambaquiensis*. Toxicon 71: 11–17.

Winkel, K.D.; Tibballs, J.; Molenaar, P.; Lambert, G.; Coles, P.; Ross-Smith, M. Wiltshire, C.; Fenner, P.J.; Gershwin, L.A.; Hawdon, G.M.; Wright, C.E. & Angus, J.A. 2005. Cardiovascular actions of the venom from the Irukandji (*Carukia barnesi*) jellyfish: Effects in human, rat and guinea-pig tissues in vitro and in pigs in vivo. Clinical and Experimental Pharmacology and Physiology 32: 777–788.

Yan, L.; Fei, K.; Bridge, D. & Sarras, M.P. 2000. A cnidarian homologue of translationally controlled tumor protein (P23/TCTP). Development Genes and Evolution 210: 507–511.

Zhang, Z.Q. 2011. Animal biodiversity: An introduction to higher-level classification and taxonomic richness. Zootaxa 12: 7–12.

FIGURES

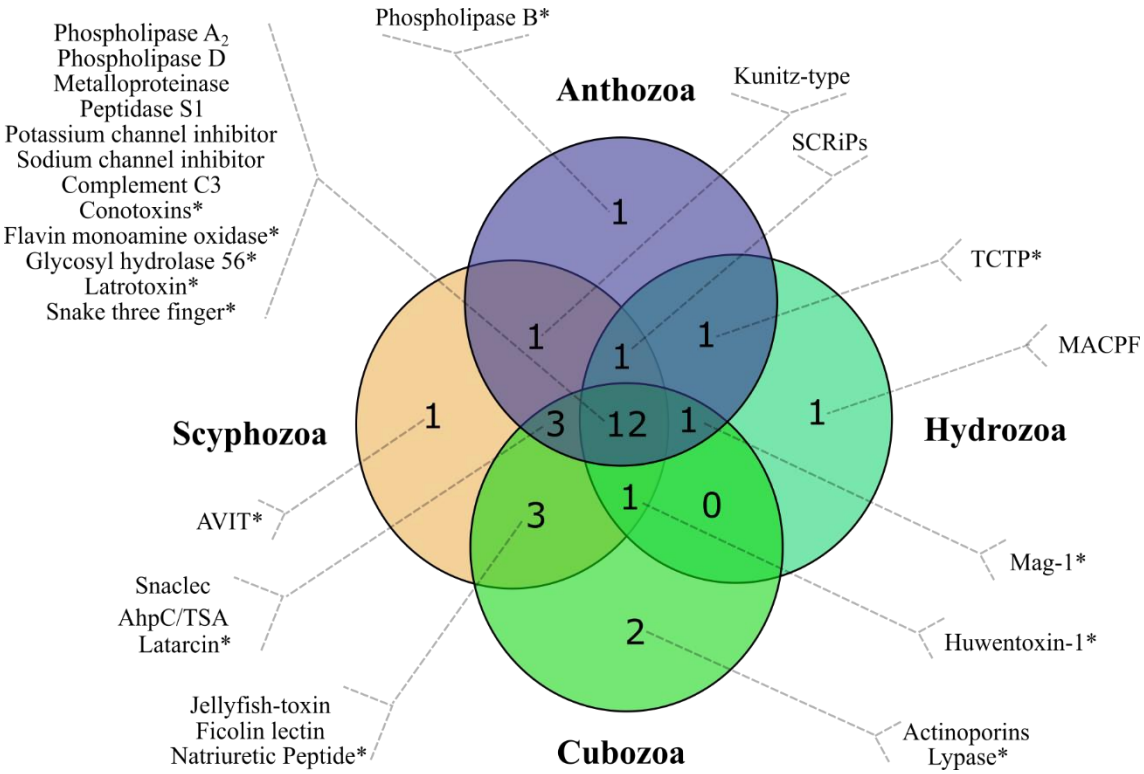

Figure 1. Venn diagram of the venom proteomes of the four-cnidarian classes, showing the number of types of toxin families shared between the classes Cnidaria phylum. The proteins marked with asterisk (*) have never been recorded for cnidarians.

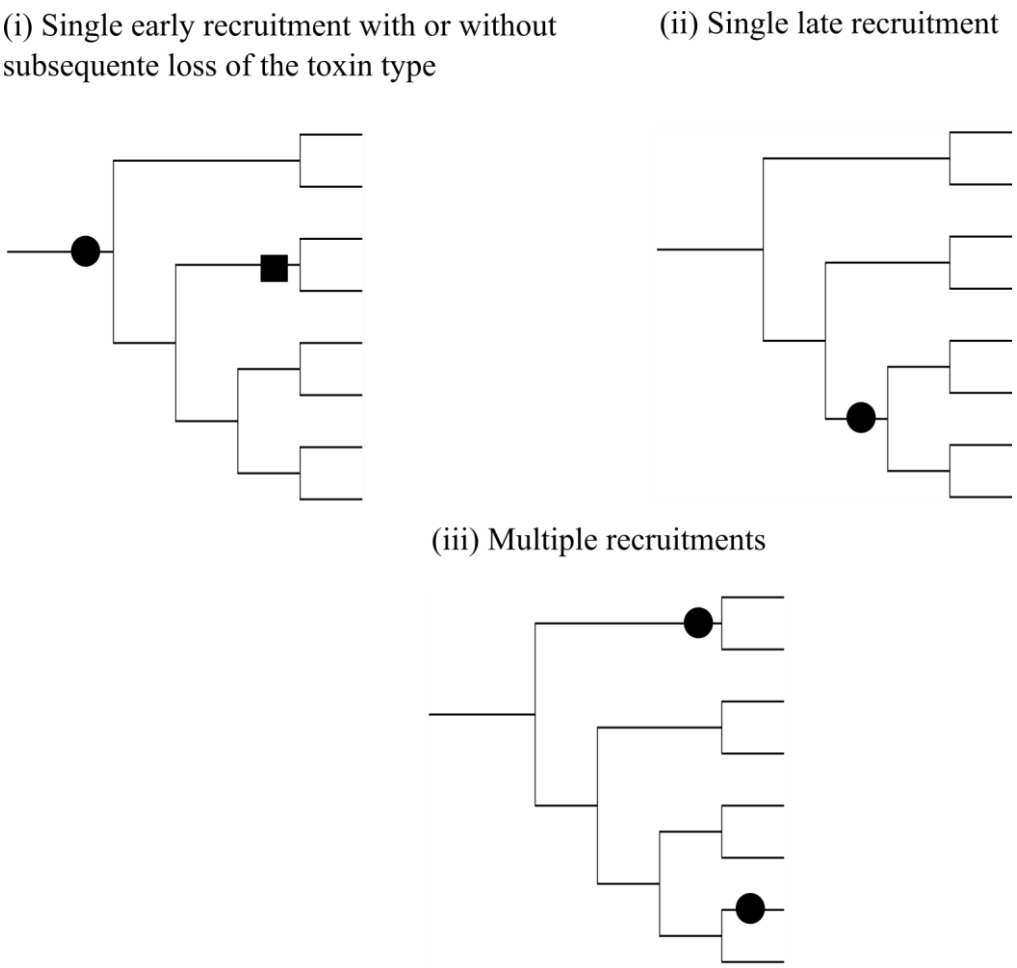

Figure 2. Schematic approach to interpreting patterns of recruitments of toxin families into Cnidaria venom proteome. Black circles represent recruitment events. Black squares represent lost/absence events.

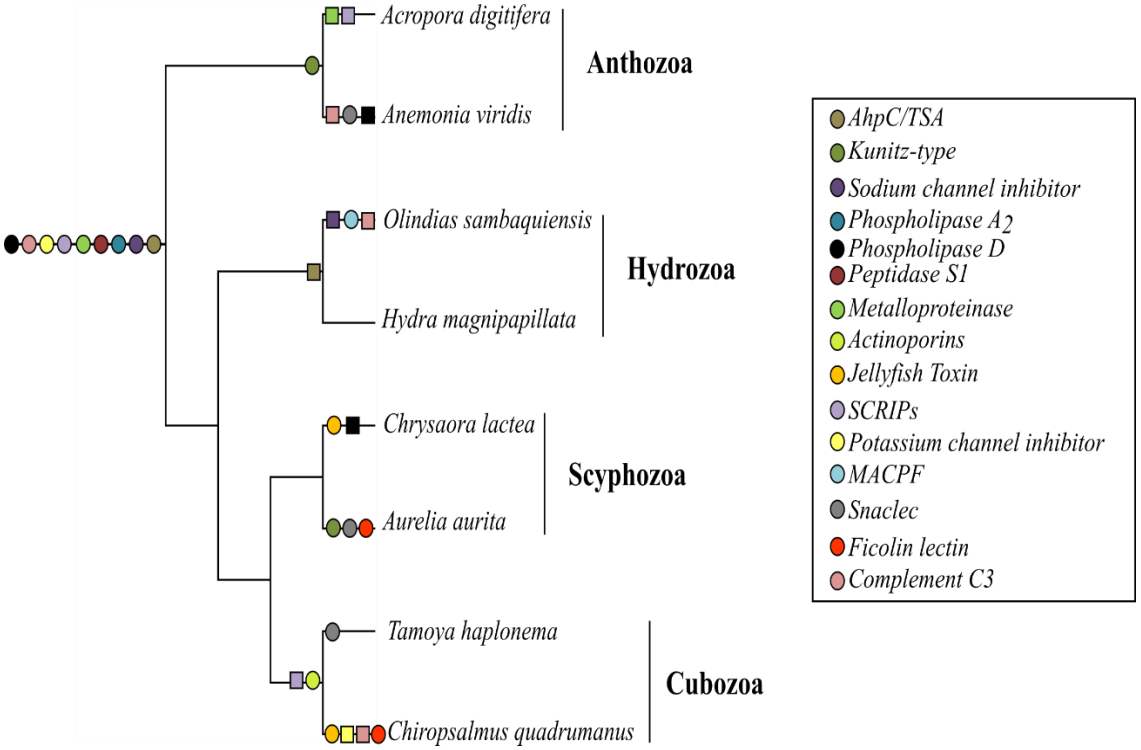

Figure 3. Recruitment pattern of toxin families into Cnidaria venom arsenal based on the cnidarian phylogeny by Marques & Collins (2004) and Collins *et al*. (2006). Circles represent recruitment events. Squares represent loss/absence of toxin families. In this figure are only displayed toxin families that have already been previously recorded for cnidarians.

TABLES

Table 1. Distribution matrix of types of protein families in the venom proteome of cnidarian classes. Expressed (1). Not found (0). The proteins marked with asterisk (*) have never been recorded for cnidarians.

Toxin family	*Acropora digitifera*	*Tamoya haplonema*	*Chiropsalmus quadrumanus*	*Olindias sambaquiensis*	*Chrysaora lactea*	*Anemonia viridis*	*Hydra magnipapillata*	*Aurelia aurita*
Actinoporins	0	1	1	0	0	0	0	0
AhpC/TSA	1	1	1	0	1	1	0	1
Complement C3	1	1	0	0	1	0	1	1
Ficolin lectin	0	0	1	0	0	0	0	1
Jellyfish toxin	0	0	1	0	1	0	0	0
Kunitz-type	1	0	0	0	0	1	0	1
MACPF	0	0	0	1	0	0	0	0
Metalloproteinase	0	1	1	1	1	1	1	1
Peptidase S1	1	1	1	1	1	1	1	1
Phospholipase A2	1	1	1	1	1	1	1	1
Phospholipase D	1	1	1	1	0	0	1	1
Potassium channel inhibitor	1	1	0	1	1	1	1	1
SCRISPs	0	0	0	1	1	1	1	1
Snaclec	0	1	0	0	0	1	0	1
Sodium channel inhibitor	1	1	1	0	1	1	1	1
AVIT*	0	0	0	0	0	0	0	1
Conotoxins*	1	1	1	0	1	1	1	1
Flavin monoamine oxidase*	1	1	0	0	1	1	1	1
Glycosyl hydrolase 56*	0	0	1	0	0	1	1	1
Huwentoxin-1*	0	0	1	1	0	0	1	1
Latarcin*	0	1	0	0	1	1	0	1
Latrotoxin*	0	1	1	1	0	1	1	1
Lipase*	0	1	0	0	0	0	0	0
Magi-1*	1	1	0	0	0	0	1	0
Natriuretic peptide*	0	1	-	0	0	0	0	1
Phospholipase B*	1	0	0	0	0	0	0	0
Snake three-finger*	1	1	1	1	0	1	1	1
TCTP*	1	0	0	0	0	0	1	0

Table 2. Summary of the results of the character mapping analysis. The proteins marked with asterisk (*) have never been recorded for cnidarians.

Toxin Family	Interpretation
Actinoporins	Single late recruitment
AhpC/TSA	Single early recruitment
Complement C3	Single early recruitment
Ficolin lectin	Independent recruitments
Jellyfish toxin	Independent recruitments
Kunitz-type	Independent recruitments
MACPF	Single late recruitment
Metalloproteinase	Single early recruitment
Peptidase S1	Single early recruitment
Phospholipase A2	Single early recruitment
Phospholipase D	Single early recruitment
Potassium channel inhibitor	Single early recruitment
SCRISPs	Single early recruitment
Snaclec	Independent recruitments
Sodium channel inhibitor	Single early recruitment
AVIT*	Single late recruitment
Conotoxins*	Single early recruitment
Flavin monoamine oxidase*	Single early recruitment
Glycosyl hydrolase 56*	Single early recruitment
Huwentoxin-1*	Single late recruitment
Latarcin*	Single early recruitment
Latrotoxin*	Single early recruitment
Lipase*	Independent recruitments
Magi-1*	Independent recruitments
Natriuretic peptide*	Independent recruitments
Phospholipase B*	Independent recruitments
Snake three-finger*	Single early recruitment
TCTP*	Independent recruitments

SUPPLEMENTARY MATERIAL

Table S1. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the anthozoan *Acropora digitifera*.

Peptide	Protein family	Uniprot accession number	Organism with highest similarity
Venom peptide Ocy2	-	P86107	*Opisthacanthus cayaporum*
Scolopendra 5885.28 Da toxin	Scolopendra toxin 4	P0C8C4	*Scolopendra viridicornis nigra*
Mastoparan	MCD	P42716	*Parapolybia indica*
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
Potassium channel toxin alpha-KTx 15.7	Potassium channel inhibitor	Q5K0E0	*Androctonus amoreuxi*
Conotoxin VnMLCL-031	Conotoxin T superfamily	Q9BP53	*Conus ventricosus*
Neurotoxin-like protein 1	Snake three-finger toxin	P84716	*Causus rhombeatus*
Eumenine mastoparan-OD	MCD	P86146	*Orancistrocerus drewseni*
Antimicrobial peptide 143	Short cationic antimicrobial peptide	P0CI90	*Lychas mucronatus*
Conotoxin ViKr92	Conotoxin O1 superfamily	Q3YED8	*Conus virgo*
Kunitz-type serine protease inhibitor superbin-1	Venom Kunitz-type	B5KL38	*Austrelaps superbus*
Toxin To6	Sodium channel inhibitor	P84685	*Tityus obscurus*
U16-lycotoxin-Ls1a	U16-lycotoxin	B6DD52	*Lycosa singoriensis*
Candiduxin-1	Snake three-finger toxin	Q8AY53 (+4)	*Bungarus candidus*
Augerpeptide-s6a	-	P0C1T6	*Terebra subulata*
Kunitz-type serine protease inhibitor B5	Venom Kunitz-type	A8Y7P5	*Daboia siamensis*
Potassium channel toxin TdiKIK	Long chain scorpion toxin	Q0GY43	*Tityus discrepans*
Kunitz-type serine protease inhibitor 3	Venom Kunitz-type	P00992	*Vipera ammodytes ammodytes*
Toxin ICK-7	Magi-1	W4VRY7	*Trittame loki*
U35-theraphotoxin-Cj1a	-	B1P1J4	*Chilobrachys guangxiensis*
Acidic phospholipase A2 PL-I	Phospholipase A2	C1IC46	*Walterinnesia aegyptia*
Basic phospholipase A2 RV-4	Phospholipase A2	Q02471	*Daboia siamensis*
Calglandulin	Calmodulin	Q8AY75 (+8)	*Bothrops insularis*
Translationally-controlled tumor protein homolog	TCTP	G3LU44	*Loxosceles intermedia*
Factor V activator RVV-V alpha	Peptidase S1	P18964 (+3)	*Daboia siamensis*
Thrombin-like enzyme bhalternin	Peptidase S1	P0CG03	*Bothrops alternatus*
Venom serine proteinase-like protein 1	Peptidase S1	Q6T6S7	*Bitis gabonica*
Phospholipase D LapSicTox-alphaII1	Arthropod phospholipase D	C0JB07 (+1)	*Loxosceles apachea*

Phospholipase B	Phospholipase B-like	F8S101	*Crotalus adamanteus*
Venom dipeptidyl peptidase 4	Peptidase S9B	B2D0J4	*Apis mellifera*
A.superbus venom factor 1	Venom complement C3 homolog	Q0ZZJ6	*Austrelaps superbus*

Table S2. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the cubozoan *Tamoya haplonema*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Alpha-latroinsectotoxin-Lt1a	Latrotoxin	Q02989	*Latrodectus tredecimguttatus*
Potassium channel toxin alpha-KTx 3	Potassium channel inhibitor	P0C8R1	*Androctonus amoreuxi*
Snake venom serine proteinase 1	Peptidase S1	J3S3W5	*Crotalus adamanteus*
L-amino-acid oxidase	Flavin monoamine oxidase	B5AR80	*Bothrops pauloensis*
Lipolysis-activating peptide 1-alpha chain	long (3 C-C) scorpion toxin superfamily	D9U2A4	*Lychas mucronatus*
M-zodatoxin-Lt4b	Latarcin	Q1ELU4	*Lachesana tarabaevi*
Potassium channel toxin TdiKIK	long chain scorpion toxin	Q0GY43	*Tityus discrepans*
U16-lycotoxin-Ls1a	U16-lycotoxin	B6DD52	*Lycosa singoriensis*
Conotoxin Bu2	Conotoxin O1 superfamily	P0CY61	*Conus bullatus*
Conotoxin ar11a	Conotoxin I1 superfamily	P0C607	*Conus arenatus*
Potassium channel toxin alpha-KTx 4.6	Potassium channel inhibitor	P0CB56	*Tityus stigmurus*
Phospholipase D LlSicTox-alphaIII	Arthropod phospholipase D	Q1KY79	*Loxosceles laeta*
Conotoxin Ml15b	Conotoxin O2 superfamily	C8CK75	*Conus miles*
Neurotoxin LmNaTx34.1	Sodium channel inhibitor	P0CI81	*Lychas mucronatus*
Zinc metalloproteinase-disintegrin-like BfMP	Venom metalloproteinase (M12B)	A8QL48	*Bungarus fasciatus*
Echotoxin-2	Sea anemone actinoporin	Q76CA2	*Monoplex parthenopeus*
Beta-insect depressant toxin LqhIT2	Sodium channel inhibitor	Q26292	*Leiurus quinquestriatus hebraeus*
Insecticidal toxin LaIT1	-	P0C5F2	*Liocheles australasiae*
Beta-mammal/insect toxin Lqhb1	Sodium channel inhibitor	P0C5H3	*Leiurus quinquestriatus hebraeus*
Neurotoxin 3FTx-RK	Snake three-finger toxin	P0C554	*Bungarus fasciatus*
Conotoxin Lt5.9	Conotoxin T superfamily	Q1A3Q7	*Conus litteratus*
Conotoxin PnMLKM-011	Conotoxin M superfamily	Q9BPI1 (+1)	*Conus pennaceus*
Conotoxin Cal5a L3	-	D2Y169	*Conus californicus*
Phospholipase D LlSicTox-alphaIV1i	Arthropod	C0JB23	*Loxosceles laeta*

	phospholipase D		
Eumenine mastoparan-OD	MCD	P86146	*Orancistrocerus drewseni*
Natriuretic peptide Oh-NP	natriuretic peptide	D9IX98	*Ophiophagus hannah*
Phospholipase D SpaSicTox-betaIF1	Arthropod phospholipase D	C0JB52	*Sicarius patagonicus*
Beta-insect depressant toxin BjIT2	Sodium channel inhibitor	P24336	*Hottentotta judaicus*
Putative potassium channel blocker TXKS1	-	Q95P89	*Mesobuthus martensii*
Acidic phospholipase A2 2	Phospholipase A2	Q9I968	*Protobothrops mucrosquamatus*
Natriuretic peptide Mf-NP	Natriuretic peptide	B8K1V9	*Micrurus fulvius*
Snaclec alboaggregin-A subunit alpha'	Snaclec	P81112	*Cryptelytrops albolabris*
Snake venom metalloproteinase	Venom metalloproteinase (M12B)	Q8JJ51	*Crotalus molossus molossus*
Phospholipase D LlSicTox-alphaIV2i*	Arthropod phospholipase D	C0JB25	*Loxosceles laeta*
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
A.superbus venom factor 1	Venom complement C3 homolog	Q0ZZJ6	*Austrelaps superbus*
Mucroporin-like peptide	Short cationic antimicrobial peptide	D9U2B8	*Lychas mucronatus*
Hainantoxin-XVI-5	magi-1	D2Y274	*Haplopelma hainanum*
Conotoxin Tx8.1	Conotoxin S superfamily	B2CI27	*Conus textile*
Venom carboxylesterase-6	lipase	B2D0J5	*Apis mellifera*
Alpha-fibrinogenase	Peptidase S1	Q8JH85	*Macrovipera lebetina*

Table S3. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the cubozoan *Chiropsalmus quadrumanus*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
Beta-theraphotoxin-Ps1a	huwentoxin-1	P84510	*Paraphysa scrofa*
Conotoxin Cal14.1c	-	D2Y102	*Conus californicus*
Omega-conotoxin-like Bu1	conotoxin O1 superfamily	P0CY60	*Conus bullatus*
Conotoxin VnMKLT2-011-013	Conotoxin O1 superfamily	Q9BP93-95	*Conus ventricosus*
Toxin Td12	Sodium channel inhibitor	Q1I172	*Tityus discrepans*
Conotoxin Bu2	Conotoxin O1 superfamily	P0CY61	*Conus bullatus*
Phi-liotoxin-Lw1a	-	P0DJ08	*Liocheles waigiensis*
Hainantoxin-XX.3	-	D2Y2D0	*Haplopelma hainanum*
Kappa-conotoxin-like SmIVB	conotoxin A superfamily	P0CE76	*Conus stercusmuscarum*
Beta-mammal toxin Cn2	Sodium channel inhibitor	P01495	*Centruroides noxius*
Neurotoxin LmNaTx1	Sodium channel inhibitor	D9U297	*Lychas mucronatus*

Probable weak neurotoxin 3FTx-Lio1	snake three-finger toxin	A7X3M9	*Erythrolamprus poecilogyrus*
Lipolysis-activating peptide 1-beta chain	lipo-B	P84809	*Buthus occitanus tunetanus*
Hainantoxin-XVIII-5	-	D2Y2N9	*Haplopelma hainanum*
U1-lycotoxin-Ls1b	U1-lycotoxin	B6DCK1	*Lycosa singoriensis*
Disintegrin Eo1 subunit 1	Disintegrin	Q3BER2	*Echis ocellatus*
U3-lycotoxin-Ls1x	U3-lycotoxin	B6DCP7 (+6)	*Lycosa singoriensis*
Phospholipase A2	phospholipase A2	A7LCJ2	*Urticina crassicornis*
Zinc metalloproteinase-disintegrin-like mikarin	Venom metalloproteinase (M12B)	P0DJ43	*Micropechis ikaheka*
Veficolin-1	Ficolin lectin	E2IYB3	*Varanus komodoensis*
Snake venom serine protease 3	Peptidase S1	O13063	*Trimeresurus gramineus*
Echotoxin-2	sea anemone actinoporin	Q76CA2	*Monoplex parthenopeus*
Phospholipase D SpeSicTox-betaIB4	Arthropod phospholipase D	C0JB34	*Sicarius peruensis*
Zinc metalloproteinase-disintegrin-like crotastatin	Venom metalloproteinase (M12B)	Q076D1	*Crotalus durissus terrificus*
Toxin CfTX-2	jellyfish toxin	A7L036	*Chironex fleckeri*
Hyaluronidase-3	Glycosyl hydrolase 56	A3QVN4	*Cerastes cerastes*
Alpha-latroinsectotoxin-Lt1a	Latrotoxin	Q02989	*Latrodectus tredecimguttatus*

Table S4. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the hydrozoan *Olindias sambaquiensis*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Phospholipase A2 acanmyotoxin-3	Phospholipase A2	P85061	*Acanthophis sp.*
Acrorhagin-1	-	Q3C258	*Actinia equina*
Snake venom metalloproteinase aculysin-1	Venom metalloproteinase (M12B)	Q9W7S2	*Deinagkistrodon acutus*
Snake venom metalloproteinase bothrojaractivase	Venom Metalloproteinase (M12B)	P0C7A9	*Bothrops jararaca*
Kappa-4-bungarotoxin	Three-finger toxin	O12961	*Bungarus multicinctus*
Cytotoxic linear peptide	Scorpion NDBP 5	H2CYR5	*Pandinus cavimanus*
Alpha-latrocrustotoxin-Lt1a	Latrotoxin	Q9XZC0	*Latrodectus tredecimguttatus*
Putative metalloprotease	-	F1CJ78	*Hottentotta judaicus*
Metalloproteinase	Disintegrin domain	E9JG55	*Echis coloratus*
Natriuretic peptide Oh-NP	Natriuretic peptide	D9IX98	*Ophiophagus hannah*
Neublin-like protein	-	Q3SAY4	*Oxyuranus scutellatus*
Potassium channel toxin alpha-Katz 12.3	Potassium channel inhibitor	P0C185	*Tityus costatus*
Acidic phospholipase A2 Ts-A1	Phospholipase A2	Q6H3D0	*Trimeresurus stejnegeri*
Basic phospholipase A2 DsM-S1	Phospholipase A2	A8CG84	*Daboia siamensis*
Zinc metalloproteinase/disintegrin	Venom metalloproteinase	Q90220	*Gloydius halys*

	(M12B)		
Putative toxin	-	C5J8C9	*Opisthacanthus cayaporum*
Serine proteinase 8	Peptidase S1	F8S120	*Crotalus adamanteus*
Serine protease HS112	Peptidase S1	Q5WVSP	*Bothrops jararaca*
Phospholipase D SdSicTox-betaIF1	Arthropod phospholipase D	C0JB55	*Sicarius cf. damarensis*
U14-Theraphotoxin-Cj1b	Huwentoxin-1	B1P1E7	*Chilobrachys guangxiensis*
k -Theraphotoxin-Cj1b	Huwentoxin-1	B1P1A0	*Chilobrachys guangxiensis*
Toxin AvTX-60A	EGF-like domain/ MACPF domain	Q76DT2	*Actineria villosa*
Turripeptide Gkn9.1	Pg turripeptide superfamily	P0C848	*Gemmula kieneri*
Toxin PsTX-60A	EGF-like domain/ MACPF domain	P58911	*Phyllodiscus semoni*
Turritoxin UID-02	-	D5KXG9	*Gemmula speciosa*
Venom allergen 5	CRISP	A9YME1	*Microctonus hyperodae*
Venom allergen 5.01	-	G7Y9P2	*Clonorchis sinensis*
Venom dipeptidylpeptidase IV	-	A6MJH9	*Notechis scutatus*
Venom serine protease 34	Peptidase S1	Q8MQS8	*Apis mellifera*

Table S5. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the scyphozoan *Chrysaora lactea*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Potassium channel toxin alpha-KTx 6.14	Potassium channel inhibitor	P84864	*Hadrurus gertschi*
Venom peptide Ocy2	-	P86107	*Opisthacanthus cayaporum*
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
Phi-liotoxin-Lw1a	-	P0DJ08	*Liocheles waigiensis*
Toxin Td5	Sodium channel inhibitor	Q1I169	*Tityus discrepans*
Conotoxin Bu2	Conotoxin O1 superfamily	P0CY61	*Conus bullatus*
Iota-conotoxin-like r11c	conotoxin I1 superfamily	Q7Z096	*Conus radiatus*
U16-lycotoxin-Ls1a	U16-lycotoxin	B6DD52	*Lycosa singoriensis*
Acidic phospholipase A2 PL1	Phospholipase A2	F8QN52	*Vipera renardi*
Basic phospholipase A2 PeBP(R)-I/II	Phospholipase A2	Q2PG81	*Protobothrops elegans*
Basic phospholipase A2 myotoxin III	Phospholipase A2	P20474	*Bothrops asper*
Cathelicidin-NA antimicrobial peptide	Cathelicidin	B6S2X0	*Naja atra*
M-zodatoxin-Lt4a	Latarcin	Q1ELU5	*Lachesana tarabaevi*
Cysteine-rich venom protein LIO1	CRISP	Q2XXQ0	*Erythrolamprus poecilogyrus*
Snake venom serine protease KN2	Peptidase S1	Q71QJ0 (+3)	*Trimeresurus stejnegeri*
Toxin CfTX-2	jellyfish toxin	A7L036	*Chironex fleckeri*
Zinc metalloproteinase/disintegrin	Venom metalloproteinase (M12B)	P15503 (+1)	*Trimeresurus gramineus*
L-amino-acid oxidase	Flavin monoamine oxidase	P0DI84	*Vipera ammodytes ammodytes*
Cobra venom factor	venom complement C3 homolog	Q91132	*Naja kaouthia*

Table S6. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the anthozoan *Anemonia viridis*.

Peptide	Protein family	Uniprot accession number	Organism with highest similarity
Snake venom serine protease Dav-PA	Peptidase S1	Q9I8X1(+1)	*Deinagkistrodon acutus*
Phospholipase A1	Lipase family	P0CH86	*Vespula squamosa*
Phospholipase A2	Phospholipase A2	P0C8L9	*Hadrurus gertschi*
L-amino-acid oxidase	Flavin monoamine oxidase	O93364	*Crotalus adamanteus*
Potassium channel toxin alpha-KTx 2.6	Potassium channel inhibitor	P59849	*Centruroides limbatus*
Snake venom serine proteinase 11	peptidase S1	J3S832 (+4)	*Crotalus adamanteus*
Alpha-latrocrustotoxin-Lt1a	Latrotoxin	Q9XZC0	*Latrodectus tredecimguttatus*
Zinc metalloproteinase/disintegrin	Venom metalloproteinase (M12B)	O57413 (+1)	*Protobothrops mucrosquamatus*
Exendin-2-long	Glucagon	C6EVG2/P04204	*Heloderma suspectum cinctum*
Zinc metalloproteinase/disintegrin	Venom metalloproteinase (M12B)	Q14FJ4	*Echis ocellatus*
Venom allergen 5	CRISP	A9QQ26	*Lycosa singoriensis*
Hyaluronidase	Glycosyl hydrolase 56	A3QVN2 (+3)	*Echis ocellatus*
Delta-conotoxin-like Ai6.1	conotoxin O1 superfamily	P0CB09	*Conus ammiralis*
Cysteine-rich venom protein tripurin	CRISP	P81995	*Cryptelytrops purpureomaculatus*
M-zodatoxin-Lt4b	Latarcin	Q1ELU4	*Lachesana tarabaevi*
Factor V activator RVV-V alpha	peptidase S1	P18964 (+2)	*Daboia siamensis*
Myotoxin-1	crotamine-myotoxin	P24331 (+2)	*Crotalus durissus terrificus*
Zinc metalloproteinase leucurolysin-B	venom metalloproteinase (M12B)	P86092	*Bothrops leucurus*
Snaclec bitiscetin subunit alpha	Snaclec	Q7LZK5	*Bitis arietans*
33kDa venom protein	-	Q7M3V1	*Chelonus sp. nr. curvimaculatus*
Kunitz-type serine protease inhibitor	Venom Kunitz-type	E5AJX3 (+1)	*Vipera nikolskii*
Hyaluronidase	Glycosyl hydrolase 56	R4J7Z9	*Loxosceles intermedia*
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
Toxin Tst2	Sodium channel inhibitor	P68410 (+1)	*Tityus serrulatus*
Basic phospholipase A2 DsM-b1/DsM-b1'	Phospholipase A2	A8CG82 (+3)	*Daboia siamensis*
Snaclec alboaggregin-A subunit beta	Snaclec	P81113	*Trimeresurus albolabris*
Potassium channel toxin alpha-KTx 9.11	Potassium channel inhibitor	B3EWX9 (+2)	*Mesobuthus gibbosus*
U16-lycotoxin-Ls1a	U16-lycotoxin	B6DD52	*Lycosa singoriensis*
Phospholipase A2	phospholipase A2	F8J2D2	*Drysdalia coronoides*
Alpha-toxin Ts5	Sodium channel inhibitor	P46115	*Tityus serrulatus*

Hongotoxin-5	Potassium channel inhibitor	P59851	*Centruroides limbatus*
Muscarinic m2-toxin	snake three-finger toxin	P60237	*Dendroaspis angusticeps*
Potassium channel toxin alpha-KTx	Potassium channel inhibitor	P84777	*Tityus discrepans*
Phospholipase A2	phospholipase A2	P86524	*Acanthophis antarcticus*
Gamma-conotoxin-like TxVIIA	conotoxin O2 superfamily	P24160	*Conus textile*
Kunitz-type serine protease inhibitor hainantoxin-XI	venom Kunitz-type	D2Y2C2 (+15)	*Haplopelma hainanum*
U2-agatoxin-Ao1o	U2-agatoxin	Q5Y4X1	*Agelena orientalis*
Conotoxin VnMKLT1-01122	conotoxin O1 superfamily	Q9BPA4	*Conus ventricosus*

Table S7. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the hydrozoan *Hydra magnipapillata*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Ophiophagus venom factor	Venom complement C3 homolog	I2C090	*Ophiophagus hannah*
Phospholipase D SpeSicTox-betaIF1	Arthropod phospholipase D	C0JB53	*Sicarius peruensis*
A.superbus venom factor 1	venom complement C3 homolog	Q0ZZJ6 (+1)	*Austrelaps superbus*
U4-ctenitoxin-Pk1a	spider toxin Tx2	P83896	*Phoneutria keyserlingi*
Conotoxin flf14a	Conotoxin L superfamily	P84705	*Conus anabathrum floridanus*
Latartoxin-1b	spider toxin CSTX superfamily	B3EWF3	*Lachesana tarabaevi*
Basic phospholipase A2 LmTX-I	phospholipase A2	P0C942 (+4)	*Lachesis muta muta*
Snake venom serine protease NaSP	Peptidase S1	A8QL53 (+1)	*Naja atra*
Putative endothelial lipase	Lipase	J3RZ81	*Crotalus adamanteus*
Basic phospholipase A2	Phospholipase A2	P19000	*Laticauda laticaudata*
L-amino-acid oxidase	Flavin monoamine oxidase	Q6WP39 (+1)	*Trimeresurus stejnegeri*
Translationally-controlled tumor protein homolog	TCTP	J3SFJ3 (+1)	*Crotalus adamanteus*
Venom prothrombin activator pseutarin-C non-catalytic subunit	multicopper oxidase	Q7SZN0	*Pseudonaja textilis*
Toxin Isom2	Sodium channel inhibitor	P0C5H2 (+1)	*Isometrus vittatus*
Phospholipase A1 1	Lipase	P0DMB4	*Vespa affinis*
Toxin ICK-7	magi-1	W4VRY7	*Trittame loki*
Toxin-like protein 14	-	L0GCW8	*Urodacus yaschenkoi*
Toxin CpTx1	spider toxin CSTX superfamily	D5GSJ8	*Cheiracanthium punctorium*
Basic phospholipase A2 homolog promutoxin	phospholipase A2	Q2PWA3	*Protobothrops mucrosquamatus*
Conotoxin VnMKLT1-01122	conotoxin O1 superfamily	Q9BPA4	*Conus ventricosus*

Basic phospholipase A2 ammodytoxin A	phospholipase A2	P00626 (+2)	*Vipera ammodytes ammodytes*
Techylectin-like protein	fibrinogen C-terminal domain	P85031	*Phoneutria nigriventer*
Omega-ctenitoxin-Pn2a	Tx3	P81789	*Phoneutria nigriventer*
Alpha-mammal toxin Lqq5	Sodium channel inhibitor	P01481 (+1)	*Leiurus quinquestriatus quinquestriatus*
Cytotoxin homolog S3C2	snake three-finger toxin	P19003	*Aspidelaps scutatus*
U14-hexatoxin-Mg1a	-	Q75WG6	*Macrothele gigas*
Conotoxin Pu3.6	conotoxin M superfamily	P0CH21	*Conus pulicarius*
Hyaluronidase	Glycosyl hydrolase 56	A3QVN2 (+3)	*Echis ocellatus*
Alpha-latroinsectotoxin-Lt1a	Latrotoxin	Q02989	*Latrodectus tredecimguttatus*
Zinc metalloproteinase carinactivase-1 catalytic subunit	venom metalloproteinase (M12B)	Q9PRP9	*Echis carinatus*
Conotoxin Bu2	conotoxin O1 superfamily	P0CY61	*Conus bullatus*
Cobra venom factor	venom complement C3 homolog	Q91132	*Naja kaouthia*
Exendin-2-long	Glucagon	C6EVG2 (+1)	*Heloderma suspectum cinctum*
Venom protein 7.1	-	P0CJ03	*Lychas mucronatus*
M-zodatoxin	cytoinsectotoxin	P0CAZ2 (+13)	*Lachesana tarabaevi*
Hyaluronidase 1	Glycosyl hydrolase 56	P85841	*Tityus serrulatus*
Snaclec purpureotin subunit alpha	Snaclec	P0DJL2	*Cryptelytrops purpureomaculatus*
Mu-theraphotoxin-Hh2a	huwentoxin-1	P83303	*Haplopelma schmidti*
Stonustoxin subunit alpha	SNTX/VTX toxin	Q98989	*Synanceia horrida*
Phospholipase A1	Lipase	P0CH87	*Vespa crabro*
Acidic phospholipase A2	phospholipase A2	P20476(+3)	*Trimeresurus gramineus*
Venom prothrombin activator omicarin-C catalytic subunit	peptidase S1	Q58L95	*Oxyuranus microlepidotus*
Snake venom serine proteinase	peptidase S1	F8S116 (+2)	*Crotalus adamanteus*
Toxin F-VIII	snake three-finger toxin	P01404	*Dendroaspis angusticeps*
Zinc metalloproteinase/disintegrin ussurin	venom metalloproteinase (M12B)	Q7SZD9	*Gloydius ussuriensis*
Snake venom 5'-nucleotidase	5'-nucleotidase	B6EWW8 (+1)	*Gloydius brevicaudus*
Conodipine-M alpha chain	phospholipase A2	Q9TWL9	*Conus magus*
Ponericin-G4	-	P82417(+1)	*Pachycondyla goeldii*
Zinc metalloproteinase-disintegrin-like agkihagin	venom metalloproteinase (M12B)	Q1PS45	*Deinagkistrodon acutus*
Heteroscorpine-1	long chain scorpion toxin	P0C2F4	*Heterometrus laoticus*
Conotoxin VnMEKL-012	conotoxin O2 superfamily	Q9BPC3 (+1)	*Conus ventricosus*
Long neurotoxin LNTX22	snake three-finger toxin	Q2VBP5	*Ophiophagus hannah*
U2-agatoxin-Ao1o	U2-agatoxin	Q5Y4X1	*Agelena orientalis*
Conotoxin MaI51	conotoxin O2 superfamily	Q3YEF4	*Conus marmoreus*

Alpha-latrotoxin-Lh1a	Latrotoxin	G0LXV8 (+1)	*Latrodectus hasseltii*
Phospholipase A2	phospholipase A2	P0C8L9	*Hadrurus gertschi*
Linear conopeptide	-	P0DKQ7	*Conus consors*
Snaclec coagulation factor IX/factor X-binding protein subunit A	Snaclec	Q71RR4	*Trimeresurus stejnegeri*
Zinc metalloproteinase-disintegrin-like ohanin	venom metalloproteinase (M12B)	A3R0T9	*Ophiophagus hannah*
Snake venom serine proteinase 11	peptidase S1	J3S832 (+4)	*Crotalus adamanteus*
U8-lycotoxin-Ls1k	U8-lycotoxin	B6DCZ1	*Lycosa singoriensis*
Inactive hyaluronidase B	Glycosyl hydrolase 56	Q5D7H4	*Vespula vulgaris*
Zinc metalloproteinase/disintegrin VMP-II	Venom metalloproteinase (M12B)	C9E1R9	*Crotalus viridis viridis*
Conotoxin Gla(1)-TxVI	conotoxin O2 superfamily	P58922	*Conus textile*
Kappa-5-bungarotoxin	snake three-finger toxin	O12962	*Bungarus multicinctus*
Potassium channel toxin alpha-KTx 15.7	Potassium channel inhibitor	Q5K0E0	*Androctonus amoreuxi*
Basic phospholipase A2 homolog zhaoermiatoxin	phospholipase A2	P84776 (+4)	*Protobothrops mangshanensis*
Basic phospholipase A2 homolog ammodytin L	phospholipase A2	P17935	*Vipera ammodytes ammodytes*
Augerpeptide Hhe9a	-	P0CI14	*Hastula hectica*
Thrombin-like enzyme bhalternin	peptidase S1	P0CG03	*Bothrops alternatus*
Kappa-scoloptoxin-Ssm2c	-	I6R1R7	*Scolopendra mutilans*
Acidic phospholipase A2 Tgc-E6	Phospholipase A2	A8E2V8	*Trimeresurus gracilis*
Snake venom metalloproteinase atroxase	Venom metalloproteinase (M12B)	Q91401	*Crotalus atrox*
Mu-theraphotoxin-Hhn2b	Huwentoxin-1	D2Y1X7	*Haplopelma hainanum*
Snake venom serine protease KN13	peptidase S1	Q71QH6 (+2)	*Trimeresurus stejnegeri*
Omega-scoloptoxin-Ssm2a	-	I6S390	*Scolopendra mutilans*
Hainantoxin-II-5	huwentoxin-2	D2Y216	*Haplopelma hainanum*
Cystatin-POGU1	cystatin	Q2XXN5	*Pogona barbata*
Alpha-latrocrustotoxin-Lt1a	latrotoxin	Q9XZC0	*Latrodectus tredecimguttatus*
Cysteine-rich venom protein VAR11	CRISP	Q2XXP1 (+2)	*Varanus varius*

Table S8. Potential toxin components of venoms identified through similarity searching of MS/MS events relating to peptides isolated following the proteomic analysis. Venom proteomic profile from the scyphozoan *Aurelia aurita*.

Peptide	**Protein family**	**Uniprot accession number**	**Organism with highest similarity**
Conotoxin Cal14.1c	-	D2Y102	*Conus californicus*
Latartoxin-2c	Spider toxin CSTX superfamily	B3EWF6	*Lachesana tarabaevi*
Potassium channel toxin alpha-KTx 2.6	Potassium channel inhibitor	P59849	*Centruroides limbatus*

U4-ctenitoxin-Pk1a	spider toxin Tx2	P83896	*Phoneutria keyserlingi*
Zinc metalloproteinase-disintegrin-like brevilysin H6	venom metalloproteinase (M12B)	P0C7B0 (+1)	*Gloydius brevicaudus*
Snaclec alboaggregin-A subunit beta	Snaclec	P81113	*Trimeresurus albolabris*
Snake venom serine protease 1	peptidase S1	O13059 (+1)	*Trimeresurus gramineus*
L-amino-acid oxidase	Flavin monoamine oxidase	F8S0Z5 (+2)	*Crotalus adamanteus*
Alpha-latrotoxin-Lh1a	Latrotoxin	G0LXV8	*Latrodectus hasseltii*
Cytotoxin 1	Snake three-finger toxin	P01471	*Hemachatus haemachatus*
Venom allergen 2	Ant venom allergen 2/4	P35776	*Solenopsis richteri*
Conotoxin Bt11.4	Conotoxin I1 superfamily	P0C609	*Conus betulinus*
Cysteine-rich venom protein tripurin	CRISP	P81995	*Cryptelytrops purpureomaculatus*
Omega-agatoxin-1A	Type I omega-agatoxin	P15969	*Agelenopsis aperta*
Zinc metalloproteinase-disintegrin-like bothrojarin-2	venom metalloproteinase (M12B)	Q0NZX9	*Bothrops jararaca*
Ponericin-G4	-	P82417 (+1)	*Pachycondyla goeldii*
Venom protein 55.1	diuretic hormone class 2	P0CJ11	*Lychas mucronatus*
U4-agatoxin-Ao1a	-	Q5Y4U5	*Agelena orientalis*
Ryncolin-4	ficolin lectin	D8VNT0	*Cerberus rynchops*
Snake venom serine protease ussurase	peptidase S1	Q8UUJ1 (+2)	*Gloydius ussuriensis*
Reticulocalbin-2	CREC	J3S9D9	*Crotalus adamanteus*
Alpha-latrocrustotoxin-Lt1a	latrotoxin	Q9XZC0	*Latrodectus tredecimguttatus*
Kappa-theraphotoxin-Cj1b	huwentoxin-1	B1P1A0	*Chilobrachys guangxiensis*
U6-lycotoxin-Ls1a	U6-lycotoxin	B6DCV1 (+6)	*Lycosa singoriensis*
Potassium channel toxin alpha-KTx 4.4	Potassium channel inhibitor	P60210	*Tityus obscurus*
Short neurotoxin homolog	snake three-finger toxin	P43445	*Bungarus multicinctus*
Potassium channel toxin alpha-KTx 4.3	Potassium channel inhibitor	P59925	*Tityus discrepans*
Zinc metalloproteinase-disintegrin-like ammodytagin	venom metalloproteinase (M12B)	P0DJE2	*Vipera ammodytes ammodytes*
Phospholipase D LvSicTox-alphaIC1bi	Arthropod phospholipase D	C0JAZ4 (+3)	*Loxosceles variegata*
Putative beta-neurotoxin RjAa14F	Sodium channel inhibitor	F2YLA3	*Rhopalurus junceus*
Mastoparan	MCD	P42716	*Parapolybia indica*
Conotoxin LiCr173	conotoxin O3 superfamily	Q3YED3	*Conus lividus*
Weak neurotoxin D2B	snake three-finger toxin	P0CAR4	*Micrurus pyrrhocryptus*
Basic phospholipase A2 F15	phospholipase A2	P0CAS5	*Crotalus durissus terrificus*
Hainantoxin-XVIII-5	-	D2Y2N9	*Haplopelma hainanum*
Opistoporin-1	Long chain multifunctional peptide (group 2)	P83313	*Opistophthalmus carinatus*
Thrombin-like enzyme CPI-enzyme 2	peptidase S1	O42207	*Gloydius ussuriensis*
M-zodatoxin-Lt8i	cytoinsectotoxin	P0CAZ2	*Lachesana tarabaevi*

		(+13)	
Natriuretic peptide NsNP-b	natriuretic peptide	Q3SAE7	*Notechis scutatus scutatus*
Thrombin-like enzyme bilineobin	peptidase S1	Q9PSN3	*Agkistrodon bilineatus*
Phospholipase D LlSicTox-alphaIV3	arthropod phospholipase D	C0JB29	*Loxosceles laeta*
Conotoxin 3	conotoxin O1 superfamily	Q5K0C7 (+1)	*Conus virgo*
M-zodatoxin-Lt4a	Latarcin	Q1ELU5 (+1)	*Lachesana tarabaevi*
Disintegrin ocellatusin	disintegrin	Q3BER1	*Echis ocellatus*
Snake venom serine protease Dav-PA	peptidase S1	Q9I8X1 (+1)	*Deinagkistrodon acutus*
Kunitz-type serine protease inhibitor bitisilin-3	venom Kunitz-type	Q6T269	*Bitis gabonica*
Goannatyrotoxin-Vere1	NPY	E2E4L2	*Varanus eremius*
L-amino-acid oxidase	Flavin monoamine oxidase	Q6WP39 (+1)	*Trimeresurus stejnegeri*
Long neurotoxin 3FTx-Oxy1	snake three-finger toxin	A7X4Q3 (+5)	*Oxyuranus microlepidotus*
Potassium channel toxin alpha-KTx 4.1	Potassium channel inhibitor	P46114	*Tityus serrulatus*
Peroxiredoxin-4	AhpC/TSA	P0CV91	*Crotalus atrox*
Conotoxin p21a	-	P86500	*Conus purpurascens*
Conotoxin AbVIF	conotoxin O1 superfamily	Q9TVK3	*Conus abbreviatus*
Hainantoxin-XVIII-3	-	D2Y2H1	*Haplopelma hainanum*
Poly-His-poly-Gly peptide 1	pHpG	P0DL07 (+1)	*Bothrops cotiara*
Cobra venom factor	venom complement C3 homolog	Q91132	*Naja kaouthia*
Long neurotoxin MS2	snake three-finger toxin	P86096 (+1)	*Micrurus surinamensis*
Phospholipase D SpeSicTox-betaIIA3i	Arthropod phospholipase D	C0JB56 (+9)	*Sicarius peruensis*
Astacin-like metalloprotease toxin 2	peptidase M12A	C9D7R2	*Loxosceles intermedia*
U1-theraphotoxin-Hh1b	huwentoxin-2	P82960	*Haplopelma schmidti*
Conotoxin BeB54	conotoxin O2 superfamily	Q3YEF9	*Conus betulinus*
Zinc metalloproteinase-disintegrin-like MTP9	venom metalloproteinase (M12B)	F8RKV9	*Drysdalia coronoides*
Phospholipase D LsaSicTox-alphaIB2ii	arthropod phospholipase D	C0JAU8	*Loxosceles sabina*
Acidic phospholipase A2 BITP01A	phospholipase A2	Q8QG87	*Bothrops insularis*
Alpha-latrotoxin-Lhe1a	Latrotoxin	P0DJE3	*Latrodectus hesperus*
Putative endothelial lipase	Lipase	J3RZ81	*Crotalus adamanteus*
Venom allergen 5	CRISP	P85840	*Tityus serrulatus*
Venom protein 59.1	IGFBP N-terminal domain	P0CJ12	*Lychas mucronatus*
Ponericin-W-like 32.1	Medium-length antimicrobial peptide (group 3)	P0CI91 (+1)	*Lychas mucronatus*
Snake venom serine proteinase 11	peptidase S1	J3S832 (+4)	*Crotalus adamanteus*
Potassium channel toxin alpha-KTx 19.1	Potassium channel inhibitor	P83407	*Mesobuthus martensii*
Basic phospholipase A2 homolog 1	phospholipase A2	P82114	*Bothrops moojeni*
Cysteine-rich venom protein	CRISP	A8S6B6 (+9)	*Austrelaps superbus*
Snake venom serine protease KN4 homolog	peptidase S1	Q71QJ4	*Trimeresurus stejnegeri*
Phospholipase D SpeSicTox-betaIIA1	Arthropod phospholipase D	C0JB67	*Sicarius peruensis*

Disintegrin VLO5B	disintegrin	P0C6B0 (+2)	*Macrovipera lebetina obtusa*
Helofensin-2	beta-defensin	D2X5W3 (+1)	*Heloderma suspectum cinctum*
Potassium channel toxin alpha-KTx 6.12	Potassium channel inhibitor	P0C166	*Anuroctonus phaiodactylus*
Conotoxin LiC42	conotoxin O1 superfamily	Q3YEG6	*Conus lividus*
Toxin ICK-2	-	W4VRY9	*Trittame loki*
Phi-liotoxin-Lw1a	-	P0DJ08	*Liocheles waigiensis*
Potassium channel toxin Meg-beta-KTx1	long chain scorpion toxin	A0A059UI30	*Mesobuthus gibbosus*
Snake venom serine protease 2	peptidase S1	Q9DF67	*Protobothrops jerdonii*
Neurotoxin 3FTx-RI	snake three-finger toxin	P0C555 (+1)	*Bungarus fasciatus*
Icarapin-like	-	Q5BLY4	*Apis mellifera*
Acidic phospholipase A2 5	phospholipase A2	P59171	*Echis ocellatus*
Exendin-2-long	Glucagon	C6EVG2 (+1)	*Heloderma suspectum cinctum*
Potassium channel toxin alpha-KTx 13.3	Potassium channel inhibitor	P84630	*Tityus pachyurus*
Toxin MIT1	AVIT (prokineticin)	P25687	*Dendroaspis polylepis polylepis*
Mu-theraphotoxin-An1a	huwentoxin-2	B3A0P0	*Acanthoscurria natalensis*
Conotoxin Cal14.1a	-	D2Y100	*Conus californicus*
Kunitz-type serine protease inhibitor microlepidin-2	venom Kunitz-type	Q6ITB4 (+3)	*Oxyuranus microlepidotus*
Acidic phospholipase A2 1	phospholipase A2	Q9W7J4	*Pseudonaja textilis*
Spiderine-2a	-	P86718 (+1)	*Oxyopes takobius*
Venom allergen 4	Ant venom allergen 2/4	P35777	*Solenopsis invicta*
Phospholipase A2 homolog	phospholipase A2	P0DMT1	*Echis pyramidum leakeyi*
Hyaluronidase A	Glycosyl hydrolase 56	P49370 (+1)	*Vespula vulgaris*
Cytotoxin 11	snake three-finger toxin	P62390	*Naja annulifera*
Hainantoxin-XII.3	huwentoxin-1	D2Y2H7	*Haplopelma hainanum*
Turripeptide GpIAa	turripeptide	P0C1X3	*Gemmula periscelida*
U19-ctenitoxin-Pn1a	-	P83997	*Phoneutria nigriventer*
Cytotoxin 6	snake three-finger toxin	Q98965	*Naja atra*
Hainantoxin-II-5	huwentoxin-2	D2Y216	*Haplopelma hainanum*
Alpha-mammal toxin Lqq5	Sodium channel inhibitor	P01481	*Leiurus quinquestriatus quinquestriatus*

Capítulo 3

Evidence of episodic positive selection in the evolution of jellyfish toxins (JFTs) of the cnidarian venom

Adrian Jaimes-Becerra[1], Paul F. Long[2-3], Antonio C. Marques[1-4]

[1]Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua Matão, Trav. 14, 101, 05508-090 São Paulo, SP, Brasil.
[2]Department of Chemistry & 9Brazil Institute, King's College London, Strand, London WC2R 2LS, United Kingdom.
[3]Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Av. Prof. Lineu Prestes, 580, B16, 05508-000 São Paulo, SP, Brasil.
[4]Centro de Biologia Marinha, Universidade de São Paulo, Rodovia Manoel Hypólito do Rego, km. 131,5, 11600-000 São Sebastião, Brasil.

ABSTRACT

The jellyfish toxins (JFTs) are a family of pore-forming proteins and are among the most dangerous toxins secreted by representatives of the phylum Cnidaria (sea anemones, jellyfish and hydroids). Besides, these toxins are known for a diversity of immunological and toxicological activities. Originally, the JFTs were reported to be limited to Cubozoa and Scyphozoa, but homologs have been recently described in other classes. However, structure-function relationships, phyletic distributions, and the molecular evolutionary regimes of toxins are poorly understood in Cnidaria, despite being the oldest venomous animal lineage. Therefore, we have made a comprehensive evolutionary analysis of JFT genes. The phylogenetic analyses identified three major clusters of JFT genes, originated from an early duplication for group III and group I/group II splitting, with group I originated from the latest duplication event near the origin of Cubozoa. Using codon based tests of positive selection, we indicated that despite long evolutionary history, this toxin family remain conserved under strong influence of negative selection. This finding is in striking contrast to the rapid evolution of toxin families in evolutionarily younger lineages, such as cone snails and advanced snakes. Besides, the results suggested branch and site specific shifts in selection within group II. We reported a correlation between the accumulation of episodically adaptive sites and the increased toxicity in group II.

Keywords: cnidarians; jellyfish toxins; venom; evolution; selection

RESUMO

As toxinas específicas de cnidários (JFTs) são uma família de proteínas formadoras de poros e estão entre as toxinas mais perigosas secretadas por representantes do filo Cnidaria (anêmonas do mar, medusas e hidróides). Além disso, são conhecidas por desempenhar uma diversidade de atividades imunológicas e toxicológicas. Originalmente, as JFTs foram relatadas como estando limitadas a Cubozoa e Scyphozoa, mas homólogos foram recentemente descritos para outras classes. No entanto, relações de estrutura-função, distribuições filéticas e os regimes evolutivos moleculares de toxinas são pouco conhecidos em Cnidaria, apesar de ser a mais antiga linhagem de animais venenosos. Assim, fizemos uma análise evolutiva abrangente dos genes JFT. As análises filogenéticas identificaram três grandes grupos de genes JFT, originados a partir de uma duplicação precoce para a divisão dos grupos I/II e grupo III, com o grupo I originado do evento de duplicação mais recente perto da origem de Cubozoa. Usando testes baseados em códons de seleção positiva, nós mostramos que, apesar da longa história evolutiva, esta família de toxinas permanece conservada sob forte influência de seleção negativa. Este resultado está em contraste com a rápida evolução das famílias de toxinas em linhagens evolutivamente mais novas, tais como as do caracol-do-cone e das serpentes. Além disso, os resultados sugerem mudanças específicas nos sítios e ramos da seleção dentro do grupo II, o que pode resultar em um aumento da toxicidade deste grupo.

INTRODUCTION

Venoms are chemical complex secretions, formed by a diversity of molecules, such proteins, peptides, polyamines, and salts, produced by some animals in response to different stimuli (Fry *et al*., 2009). Venoms may be passively or actively inoculated into another organism through delivery mechanisms and mechanical injury causing dose-dependent physiological injury (Nelsen *et al*., 2013). Several animal lineages have become venomous independently and even share molecular similarities in toxin repertoire, presumably derived from independent recruitment of related genes during venom evolution (Fry *et al*., 2009). Genes encoding venom toxins are evolutionarily significant because these genes directly affect organismal fitness and ecological adaptation. Notably, molecular evolution of venoms is highly dynamic and appears to be shaped by frequent gene duplications followed by rapid hyper-mutation and adaptive selection (Wong & Belov, 2012; Casewell *et al*., 2013, Gacesa *et al*., 2015), as well as

other molecular mechanisms such as alternative primary mRNA splicing, or a variety of post-translational modifications (Starcevic & Long, 2013).

Cnidarians might be the earliest venomous animal lineage, originated ca. 650 Ma ago (Van Iten *et al*., 2014). The phylum encompasses two major lineages, Anthozoa (e.g., sea anemones and corals) and Medusozoa, the latter comprising the classes Staurozoa (stalked jellyfish), Cubozoa (box jellyfish), Scyphozoa ('true' jellyfish), and Hydrozoa (hydroids) (Marques & Collins, 2004; Collins *et al.,* 2006, Van Iten *et al.*, 2006, 2014). All cnidarians are venomous (Turk & Kem, 2009; Badré, 2014; Jouiaei *et al*., 2015) and have a specialised venom delivery system named nematocyst, organelles contained in cnidocytes that define the phylum (Marques & Collins, 2004). For all these reasons, cnidarians make a key group to understand the origin and diversification of the venom trait.

However, even though there have been numerous studies characterising venoms of several animal lineages (e.g., Fry *et al*., 2010; Casewell *et al*., 2012; Jackson *et al*., 2013; Undheim *et al*., 2013, 2014; Baumann *et al*., 2014; Lomonte *et al*., 2014), very few deal with the complement, function, evolution, and diversification of cnidarian venoms (Jouiaei *et al.,* 2015b). Despite these deficiencies, it is known that some cnidarians, such as box jellyfishes, have some of the world's most lethal venoms (Badré, 2014; Mariottini, 2014). This has been attributed to jellyfish toxins (JFTs), which are among the most dangerous toxins secreted by jellyfish, exhibiting many toxicological activities (Tibballs *et al*., 2011; Brinkman *et al*., 2014, 2015; Rachamim *et al*., 2014; Jouiaei *et al*., 2015a, 2015b). This protein family was first discovered in the *Carybdea alata* species (Chung *et al*., 2001). To date has been reported in many other cubozoans (Nagai *et al*., 2000; Brinkman *et al*., 2015) and species of the Scyphozoa and Hydrozoa classes too (Brinkman *et al*., 2014; Rachamim *et al*., 2014). JFTs are associated with potent haemolytic activity and pore formation in mammalian cells and with inflammation, cardiovascular collapse and dermonecrosis (Nagai *et al*., 2002; Brinkman & Burnell, 2008; Yanagihara & Shohet, 2012).

In this study, we assessed the evolutionary history and selection pressures influencing JFTs at gene and protein levels. We test the hypothesis whether JFTs' variation has been driven by positive selection regime. Finally, we infer how the

observed mutations affect the tertiary structure of JFT and how this could have influenced the evolution of cnidarians.

MATERIALS AND METHODS

Sequence retrieval and alignment

We compiled a dataset of 24 nucleotide sequences from eight species of cnidarians (Table S1) to assess the molecular evolution of JFTs. Fifteen nucleotide and protein sequences were downloaded from the National Center for Biotechnology Information (NCBI: http://www.ncbi.nlm.nih.gov). These sequences (identified by GenBank accession numbers in table S1) were retrieved by BLAST. Nine nucleotide and protein sequences were directly obtained by courtesy of colleagues of a recent published study (Rachamim *et al*. 2014; marked with asterisk in table S1). Altogether, the complete dataset is herein referred to as large. Nucleotide and amino acid sequences were aligned using MAFFT 7 (Katoh & Standley, 2013). A subset of n=15 sequences, i.e., sequences only obtained from GenBank, was separately analysed (referred as small dataset). Differences between large and small datasets may give clues about the sensitivity of the results in relation to sequence sampling.

Evolutionary analyses

Best-fit model of nucleotide substitution and amino acid replacement for both datasets were determined according to Akaike's information criterion using jModeltest 2.1 (Darriba *et al*., 2012) and Prottest 3.0 (Darriba *et al*., 2011), respectively. Phylogenetic relationships among cnidarian JFTs for large and small datasets were inferred by maximum likelihood (ML) approach. Optimal ML phylogenetic tree was obtained using RAxML version 8 (Stamatakis, 2014) and node support was evaluated with 1000 bootstrapping replicates. Maximum Likelihood (ML) models of coding sequence evolution implemented by Codeml in PAML v.4 (Yang, 2007) were used to test the hypothesis that functional diversification of JFT genes was driven by positive selection. PAML compares ML estimates of the ratios of non-synonymous (d_N) to synonymous (d_S) substitutions (called omega (ω)) across an alignment with a predefined distribution and uses empirical bayesian methods to identify individual positively selected sites (Nielsen & Yang, 1998; Yang & Bielawski, 2000). The evidence for positive selection on JFT genes in cnidarians was evaluated for both large and small datasets by

employing branch, sites and branch-sites models (Goldman & Yang, 1994; Nielsen & Yang, 1998; Yang, 2000; Yang & Nielsen, 2002; Zhang *et al*., 2005).

Statistical differences among models were estimated by likelihood ratio test (LRT). LRT statistics equals two twice the difference between the log likelihood scores of the models being evaluated and compared with the X^2 distribution with appropriate degree of freedom, the difference in the number of parameters between the two models. The bayes empirical bayes (BEB) approach (Yang *et al*., 2005) was used to identify amino acids under positive selection by calculating the posterior probabilities (PP) of a particular amino acid and its relation to a given selection class (neutral, conserved, or highly variable). Sites with greater PP (PP ≥ 95 %) of belonging to the ''ω > 1 class'' were inferred to be positively selected. Additionally, MEME (Murrell *et al*., 2012) was used to detect episodically diversifying sites.

Structural and sequence analysis

In order to construct tertiary (3D) structural models of five JFTs, homology modelling was performed using I-TASSER server (Zhang, 2008). Amino acid sequences of JFTs were submitted to I-TASSER server for modelling based on multiple-threading alignments and iterative template fragment assembly simulations. The sites under positive selection were mapped on the best structural models. The 3D structures were visualised in PyMOL Molecular Graphics System, Version 1.3 Schrodinger, LLC. Prediction of solvent accessibility of each of the amino acids sites was estimated using I-TASSER server. Conserved motifs in amino acid sequences were detected using MEME suite web server (Bailey *et al.* 2009). SMART (Simple Modular Architecture Research Tool) (Letunic *et al*. 2014) was used for identification of domain architecture in JFTs.

RESULTS

Evolutionary analyses

Phylogenetic analyses identified three major clusters of JFTs genes (Fig. 1), supported by high bootstrapping values, for large and small datasets. The first group I represented by JFT genes from the species belonging only to Cubozoa, the second group II represented by genes from species belonging only to Scyphozoa and Cubozoa, while group III encompasses the remaining JFTs - like paralogue genes from *Hydra*

species and the remaining Scyphozoa. Evolution of JFTs exhibits a divergence pattern similar to that of speciation of the phylum.

One-ratio branch model estimates of ω were 0.18 and 0.14 for small and large datasets, respectively, and correspond to a significantly worse fit to the free ratio branch model (Table 1). Two-ratio branch model for both datasets, allowing branch leading to group II to have a different ω, were not significantly different from one-ratio model (Table 1). This pattern indicated an overall history of purifying selection on JFTs, with no temporal variation along branches. Comparisons of M1a and M2a sites models indicated a non-significant variation in ω among codon sites for large and small alignments (Table 1). M7 and M8 site models were significantly different for small alignment ($P<0.05$) and one site undergoing positive selection was identified by Bayes empirical Bayes (BEB) (posterior probability $\omega > 1 = 0.669$). Moreover, M8 model did not result in a better fit to the data than M7 model for the large alignment. MA branch-sites model for large alignment examining evidence for selection at specific sites along group II branch was not significant and different from null model, fixing this branch to $\omega=1$. Branch-sites model for small dataset did estimate 16% of sites having $\omega>1$ along group II branch and BEB identified 11 codons as possible targets of positive selection, but only three with posterior probabilities 0.96-0.98: 194A, 399S and 510W (Table 1). MEME, allowing ω to vary across codons as well as across branches of the phylogeny, was applied to small alignment. This resulted in thirteen codons likely under positive selection (Table 2).

Structural analysis

The quality of the modelled proteins was estimated regarding confidence scores generated by I-TASSER software, which ranged from -1.39 to -2.87. Model 3D constructed from *Chironex fleckeri* toxin B sequence was chosen to map the sixteen positively selected sites (68S 78G 121A 177S 194A 217F 219N 399S 419R 481N 510W 515A 516D 542S 561S) identified by BEB and MEME (Fig. 2). This model obtained one score ≥-1.5, which indicates that the quality prediction is correct (Roy *et al.*, 2010).

DISCUSSION

This study represents the most comprehensive evolutionary analysis based on currently available data of jellyfish toxins. Phylogenetic approach allows to characterise

three distinct groups of JFT genes (groups I-III) (Fig. 1), originated from an early duplication for group III and group I/group II splitting, with group I originated from the latest duplication event near the origin of Cubozoa, therefore encompassing only cubozoan genes. All groups have at least two JFT genes for most species, suggesting that additional gene duplication have resulted in different JFT genes repertoire in different cnidarians. Despite having only twenty-four sequences from eight species, groups II and III of JFT genes diverged in accordance with the species tree, suggesting divergent evolution rather than the more widely accepted theory that toxin gene families have arisen by convergent evolution. Additionally, the phylogenetic tree suggests that the three groups of JFT genes might have experienced birth-and-death model of evolution (Nei *et al*., 1997) because, apart from gene duplications, group-specific gene losses may have occurred as well as clustering of orthologs instead of paralogues. Similar groupings were shown previously based on 14 and 19 JFT sequences (Brinkman *et al*., 2014; Rachamim *et al*., 2014). Gene duplication that gives rise to new toxins within a toxin protein family has also been very recently inferred in the coral *Acropora digitifera* (Gacesa *et al*., 2015) and in other non-cnidarian venomous animals (Fry *et al*., 2009; Casewell *et al*., 2013).

Most genes typically evolve under negative selection (Endo *et al*., 1986). However, most animal toxins evolve under significant influence of positive selection, driven by an ongoing predator-prey arms race (Casewell *et al*., 2012; Sunagar *et al*., 2012; Brust *et al*., 2013). This model seems to be almost universal amongst venomous taxa. Contrastingly, in this study, selective regime on JFT genes was predominantly purifying or of negative selection, based on estimated overall ω value <1 for both alignments (Table 1). This pattern may be related to functional constraints of these proteins, perhaps because the maintenance of this cnidarian cytotoxin would have important fitness consequences. An alternative explanation is that JFT could play a role in nematocyst development, in addition to the intrinsic cnidarian toxicity, resulting in bifunctional proteins (Rachamim *et al*., 2014), thereby constraining a molecular arms race between cnidarians and prey. Moreover, considering that JFTs have the non-specific function of pore-forming toxins, JFTs probably do not experience a predator-prey chemical arms race, ensuring important amino acid residues remain non-synonymous, evolving under constraints of negative selection. Other toxin protein

families found in cnidarians have also been suggested to be maintained under negative selection (Rachamim *et al*., 2014; Jouiaei *et al*., 2015).

It has been noted that positive selection is more readily identified in smaller alignments, although counter intuitively, including additional sequences may cause sites to no longer be detected (Yokoyama *et al*., 2008; Chen & Sun, 2011). This phenomenon could be readily explained by purifying selection in some lineages, masking the signal of positive selection in others. Besides that branch model estimates ω for each lineage as an average of all branches and site models as an average over all sites, often episodic adaptations that affect only few amino acids and/or lineages are not identified. In the present study, a small alignment and branch site model was used to avoid these caveats. The results suggested branch and site specific shifts in selection within group II. Interestingly, analysis of the other groups (I and III) obtained no significant results. It has been suggested that one important difference between groups II and I, is the presence of a N-terminal short propart (5-7 residues) with one dibasic proteolytic cleavage site (RR/RK) between the signal peptide and mature protein (Brinkman *et al*., 2014). Many functions can be attributed to this region, including preventing unwanted toxicity to the host cell prior to propart cleavage and toxin activation (Bravo *et al*., 2002; Tang *et al*., 2003; Chen & Inouye, 2008; Wong *et al*. 2013). It is possible that toxins of group II could be more toxic than group I, thus the propart may provide additional protection to the cnidarian host organism during toxin production. Also, differences in the cardiovascular and cytolytic activities associated with toxins of groups I and II have been experimentally determined previously, suggesting an evolutionary adaptive diversification (Brinkman *et al*., 2014).

Results of structural homology predicted by I-TASSER server were indicated between the five protein JFTs and pore forming structure toxins, mainly CRY proteins, a protein family from *Bacillus thuringiensis* that kills infected host insects (Pardo-López *et al*., 2013). Interestingly, structures of CRY proteins and JFTs are highly similar in the N-terminal domain (Domain I), formed by seven or eight helix bundles with a central hydrophobic helix surrounded by outer helices. This N-terminal domain I is presumably inserted into the target membrane to form part of the pore that eventually is lethal to the target epithelial cells of the insect's gut (Schnepf *et al*., 1998). Therefore, it is plausible that this region of JFT proteins is involved in pore formation too. A search of protein sequences of JFT in Simple Modular Architecture Research Tool (SMART) (Letunic *et*

al., 2014) allows the identification of similar domains to delta-Endotoxin (insecticide), N-terminal domain with an e-value=2.00e-04. The functional role(s) of the C-terminal domain in JFT is not clear and further functional studies are necessary.

Bayes Empirical Bayes methods (BEB) (Yang *et al*., 2005) identified eleven codon positions, three of which were highly significant (Table 2). Additionally, MEME identified thirteen codon positions (Table 3), suggesting that positions found in both methods could be particularly important for adaptive processes in JFTs. Remarkably, six out of sixteen episodically diversifying sites are on the exposed surface (Table 4), while seven are buried. Thus, evolution through rapid accumulation of variation in exposed residues (RAVER), in which the molecular surface of the toxin accumulates a bulk of variations under significant influence of positive Darwinian selection and preserves core residues involved in stability and/or catalytic activity (Sunagar *et al*. 2013), seems unclear in this cnidarian toxin family. However, it should be noted that recently RAVER was identified in other families of toxins from the class Anthozoa (Jouiaei *et al*., 2015a).

It is unclear what is the significance of the positively selected sites for JFTs in group II, since these residues are located throughout the protein (Fig. 3), possibly including all functional and non-functional parts thereof. Few sites are found in three highly conserved regions (sequence motifs) identified by MEME algorithm (Table 4) (Bailey & Elkan, 1994), so are probably functionally and structurally important for this toxin family.

Concluding, episodes of diversified evolution affecting a small subset of branches at individual sites seems to be the model that best fits the evolution of JFTs. The dynamic nature of the venom in cnidarians possibly derived from an alternation of events with evolutionary chemical arms race scenarios and purifying selection for long periods, contrasts to patterns found in "younger" lineages, such snakes, where rapid evolution of genes under positive selection is widespread. Data presented herein supports a role for divergent evolution, which is contrary to current thinking that argues in favour of convergent evolution of the venom trait. However, perhaps a role for divergent evolution in diversification of selected toxin protein families now warrants fresh reappraisal and scrutiny.

REFERENCES

Badré, S. 2014. Bioactive toxins from stinging jellyfish. Toxicon 91:114–125.

Bailey, T.L.; Boden, M.; Buske, F.A; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W. & Noble, W.S. 2009. MEME Suite: Tools for motif discovery and searching. Nucleic Acids Research 37:202–208.

Bailey, T.L. & Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the International Conference on Intelligent Systems for Molecular Biology 2:28–36.

Baumann, K.; Casewell, N.R.; Ali, S.A.; Jackson, T.N.W.; Vetter, I.; Dobson, J.S.; Cutmore, S.C.; Nouwens, A.; Lavergne, V. & Fry, B.G. 2014. A ray of venom: Combined proteomic and transcriptomic investigation of fish venom composition using barb tissue from the blue-spotted stingray (*Neotrygon kuhlii*). Journal of Proteomics 109:188–198.

Bravo, A.; Sánchez, J.; Kouskoura, T. & Crickmore, N. 2002. N-terminal activation is an essential early step in the mechanism of action of the Bacillus thuringiensis Cry1Ac insecticidal toxin. The Journal of Biological Chemistry 277:23985–23987.

Brinkman, D. & Burnell, J. 2008. Partial purification of cytolytic venom proteins from the box jellyfish, *Chironex fleckeri* Toxicon 51:853–863.

Brinkman, D.L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D. & Mulvenna, J. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16:1–15.

Brinkman, D.L.; Konstantakopoulos, N.; McInerney, B.V.; Mulvenna, J.; Seymour, J.E.; Isbister, G.K. & Hodgson, W.C. 2014. *Chironex fleckeri* (box jellyfish) venom proteins: Expansion of a cnidarian toxin family that elicits variable cytolytic and cardiovascular effects. The Journal of Biological Chemistry 289:4798–4812.

Brust, A.; Sunagar, K.; Undheim, E.A.B.; Vetter, I.; Yang, D.C.; Yang, D.C.; Casewell, N.R.; Jackson, T.N.W.; Koludarov, I.; Alewood, P.F.; Hodgson, W.C.; Lewis, R.J.; King, G.F.; Antunes, A.; Hendrikx, I. & Fry, B.G. 2013. Differential evolution and neofunctionalization of snake venom metalloprotease domains. Molecular & Cellular Proteomics 12:651–63.

Casewell, N.R.; Huttley, G.A. & Wüster, W. 2012. Dynamic evolution of venom proteins in squamate reptiles. Nature Communications 3:1066.

Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A. & Fry, B.G. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends in Ecology & Evolution 28:219–229.

Chen, J. & Sun, Y. 2011. Variation in the analysis of positively selected sites using nonsynonymous/synonymous rate ratios: An example using influenza virus. PLoS One 6 (5):e19996.

Chen, Y.J. & Inouye, M. 2008. The intramolecular chaperone-mediated protein folding. Current Opinion in Structural Biology 18:765–770.

Chung, J.J.; Ratnapala, L.A.; Cooke, I.M. & Yanagihara, A.A. 2001. Partial purification and characterization of a hemolysin (CAH1) from Hawaiian box jellyfish (*Carybdea alata*) venom. Toxicon 39:981–990.

Collins, A.G.; Schuchert, P.; Marques, A.C.; Jankowski, T.; Medina, M. & Schierwater, B. 2006. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. Systematic Biology 55:97–115.

Darriba, D.; Taboada, G.L.; Doallo, R. & Posada, D. 2011. ProtTest-HPC: Fast selection of best-fit models of protein evolution. Lecture Notes in Computer Science. 6586:177–184.

Darriba, D.; Taboada, G.L.; Doallo, R. & Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772.

Endo, T.; Ikeo, K. & Gojobori, T. 1986. Large-Scale Search for Genes on Which Positive Selection May Operate. Molecular Biology and Evolution 13:685–690.

Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.A; King, G.F.; Nevalainen, T.J.; Norman, J.A; Lewis, R.J.; Norton, R.S.; Renjifo, C. & de la Vega, R.C.R. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10:483–511.

Fry, B.G.; Winter, K.; Norman, J.A; Roelants, K.; Nabuurs, R.J.A; van Osch, M.J.P.; Teeuwisse, W.M.; van der Weerd, L.; McNaughtan, J.E.; Kwok, H.F.; Scheib, H.; Greisman, L.; Kochva, E.; Miller, L.J.; Gao, F.; Karas, J.; Scanlon, D.; Lin, F.; Kuruppu, S.; Shaw, C.; Wong, L. & Hodgson, W.C. 2010. Functional and structural diversification of the Anguimorpha lizard venom system. Molecular Cellular and Proteomics 9:2369–2390.

Goldman, N. & Yang, Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Molecular Biology and Evolution 11:725–736.

Jackson, T.N.W.; Sunagar, K.; Undheim, E.A.B.; Koludarov, I.; Chan, A.H.C.; Sanders, K.; Ali, S.A.; Hendrikx, I.; Dunstan, N. & Fry, B.G. 2013. Venom down under: Dynamic evolution of Australian elapid snake toxins. Toxins 5:2621–2655.

Jouiaei, M.; Sunagar, K.; Gross, A.F.; Scheib, H.; Alewood, P.F.; Moran, Y. & Fry, B.G. 2015a. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Molecular Biology and Evolution 32: 1598–1610.

Jouiaei, M.; Yanagihara, A.; Madio, B.; Nevalainen, T.; Alewood, P. & Fry, B. 2015b. Ancient venom systems: A review on Cnidaria toxins. Toxins 7: 2251–2271.

Katoh, K. & Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30:772–780.

Letunic, I.; Doerks, T. & Bork, P. 2014. SMART: recent updates, new developments and status in 2015. 2015. Nucleic Acids Research 43:257–260.

Lomonte, B.; Tsai, W.C.; Ureña-Diaz, J.M.; Sanz, L.; Mora-Obando, D.; Sánchez, E.E.; Fry, B.G.; Gutiérrez, J.M.; Gibbs, H.L.; Sovic, M.G. & Calvete, J.J. 2014. Venomics of new world pit vipers: Genus-wide comparisons of venom proteomes across agkistrodon. Journal of Proteomics 96:103–116.

Mariottini, G.L. 2014. Haemolytic venoms from marine cnidarian jellyfish – an overview. Journal of Venom Research 5: 22–32.

Marques, A.C. & Collins, A.G. 2004. Cladistic analysis of Medusozoa and cnidarian evolution. Invertebrate Biology 123: 23–42.

Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K. & Kosakovsky, P.S.L. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genetics 8: e1002764.

Nagai, H.; Takuwa, K.; Nakao, M.; Ito, E.; Miyake, M.; Noda, M. & Nakajima, T. 2000. Novel proteinaceous toxins from the box jellyfish (sea wasp) *Carybdea rastoni*. Biochemical and Biophysical Research Communications 275: 582–588.

Nagai, H.; Takuwa-Kuroda, K.; Nakao, M.; Oshiro, N.; Iwanaga, S. & Nakajima, T. 2002. A novel protein toxin from the deadly box jellyfish (sea wasp, Habu-kurage) *Chiropsalmus quadrigatus*. Bioscience, Biotechnology, and Biochemistry 66: 97–102.

Nei, M.; Gu, X. & Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proceedings of the national academy of sciences 94:7799–7806.

Nelsen, D.R.; Nisani, Z.; Cooper, A.M.; Fox, G.A.; Gren, E.C.K. Corbit, A.G. & Hayes, W.K. 2014. Poisons, toxungens, and venoms: Redefining and classifying toxic biological secretions and the organisms that employ them. Biological Reviews 89: 450–465.

Nielsen, R. & Yang, Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929–936.

Pardo-López, L.; Soberón, M. & Bravo, A. 2013. *Bacillus thuringiensis* insecticidal three-domain Cry toxins: Mode of action, insect resistance and consequences for crop protection. FEMS Microbiology Reviews 37:3–22.

Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T. & Sher, D. 2014. The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. Molecular Biology and Evolution 32: 740–753.

Roy, A.; Kucukural, A. & Zhang, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 5:725–738.

Schnepf, E.; Crickmore, N.; Van Rie, J.; Lereclus, D.; Baum, J.; Feitelson, J.; Zeigler, D.R. & Dean, D.H. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiology and Molecular Biology Reviews 62:775–806.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Starcevic, A. & Long, P.F. 2013. Diversification of animal venom peptides-were jellyfish amongst the first combinatorial chemists?. ChemBioChem 14: 1407–1409.

Sunagar, K.; Jackson, T.N.W.; Undheim, E.A.B.; Ali, S.A.; Antunes; A. & Fry, B.G. 2013. Three-fingered RAVERs: Rapid Accumulation of Variations in Exposed Residues of snake venom toxins. Toxins 5:2172–2208.

Sunagar, K.; Johnson, W.E.; O'Brien, S.J.; Vasconcelos, V. & Antunes, A. 2012. Evolution of CRISPs associated with toxicoferan-reptilian venom and mammalian reproduction. Molecular Biology and Evolution 29:1807–1822.

Tang, B.; Nirasawa, S.; Kitaoka, M.; Marie-Claire, C. & Hayashi, K. 2003. General function of N-terminal propeptide on assisting protein folding and inhibiting catalytic activity based on observations with a chimeric thermolysin-like protease. Biochemical and Biophysical Research Communications 301:1093–1098.

Tibballs, J.; Yanagihara, A.; Turner, H. & Winkel, K. 2011. Immunological and toxinological responses to jellyfish stings. Inflammation & Allergy - Drug Targets 10: 438–446.

Turk, T. & Kem, W.R. 2009. The phylum Cnidaria and investigations of its toxins and venoms until 1990. Toxicon 54: 1031–1037.

Undheim, E.B.; Jones, A.; Clauser, K.R.; Holland, J.W; Pineda, S.S.; King, G.F. & Fry, B.G. 2014. Clawing through evolution: Toxin diversification and convergence in the ancient lineage Chilopoda (centipedes). Molecular Biology and Evolution 31: 2124–2148.

Undheim, E.A B.; Sunagar, K.; Herzig, V.; Kely, L.; Low, D.H.W.; Jackson, T.N.W.; Jones, A.; Kurniawan, N.; King, G.F.; Ali, S.A.; Antunes, A.; Ruder, T. & Fry, B.G. 2013. A Proteomics and Transcriptomics investigation of the venom from the Barychelid spider Trittame loki (brush-foot trapdoor). Toxins 5:2488–2503.

Van Iten, H.; Marques, A.C.; Leme, J.D.M.; Pacheco, M.L.F. & Simões, M.G. 2014. Origin and early diversification of the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's Proterozoic-Cambrian history. Palaeontology 57: 677–690.

Wong, E.S.W. & Belov, K. 2012. Venom evolution through gene duplications. Gene 496:1–7.

Wong, E.S.W.; Hardy, M.C.; Wood, D.; Bailey, T. & King, G.F. 2013. SVM-Based Prediction of Propeptide Cleavage Sites in Spider Toxins Identifies Toxin Innovation in an Australian Tarantula. PLoS One 8(7): e66279.

Yanagihara, A.A & Shohet, R.V. 2012. Cubozoan venom-induced cardiovascular collapse is caused by hyperkalemia and prevented by zinc gluconate in mice. PLoS One. 7(12):e51368.

Yang, Z. & Bielawski, J.R. 2000. Statistical methods for detecting molecular adaptation. Trends in Ecology and Evolution 15:496–503.

Yang, Z. & Nielsen, R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Molecular Biology and Evolution 19:908–917.

Yang, Z.; Wong, W.S.W. & Nielsen, R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Molecular Biology and Evolution 22:1107–1118.

Yang, Z. 2000. Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. Journal of Molecular Evolution 51:423–432.

Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24:1586–1591.

Yokoyama, S.; Tada, T.; Zhang, H. & Britt, L. 2008. Elucidation of phenotypic adaptations: Molecular analyses of dim-light vision proteins in vertebrates. Proceedings of the national academy of sciences 105:13480–13485.

Zhang, J.; Nielsen R. & Yang, Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution 22:2472–2479.

Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40.

FIGURES

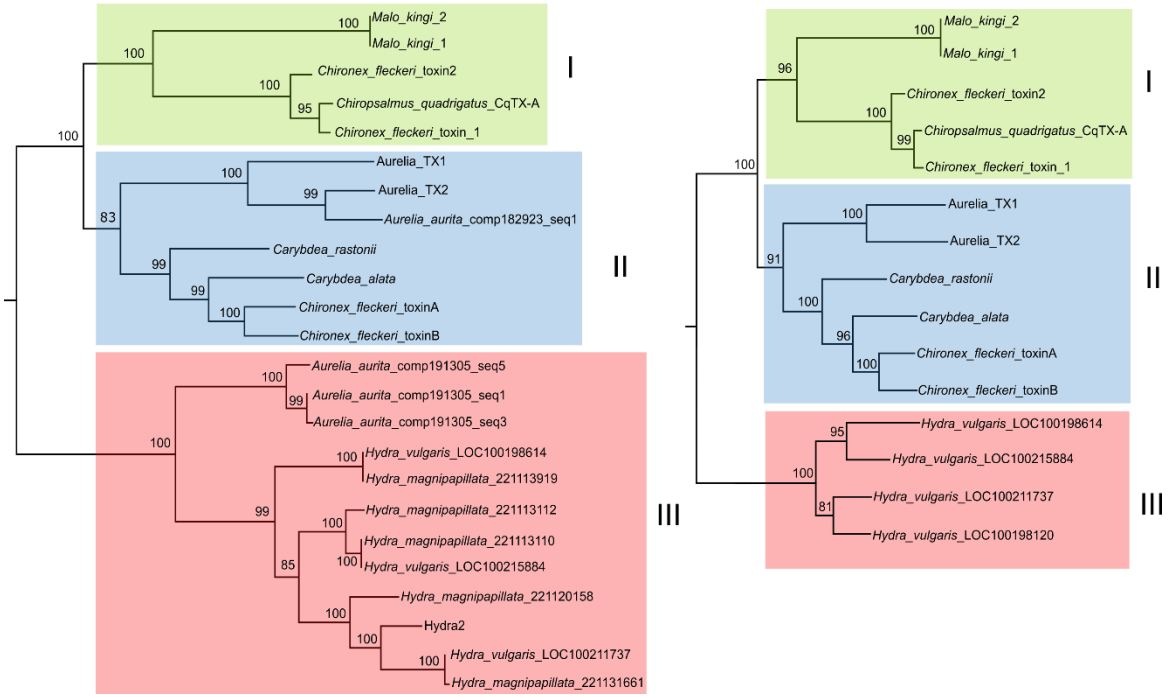

Figure 1. Phylogenetic relationships of JFT genes, based on ML analysis on the nucleotide-coding sequences, bootstrap values based on 1,000 replicates, only nodes > 50% are shown. Large dataset (24 JFTs) on left; small dataset (15 JFTs) on right.

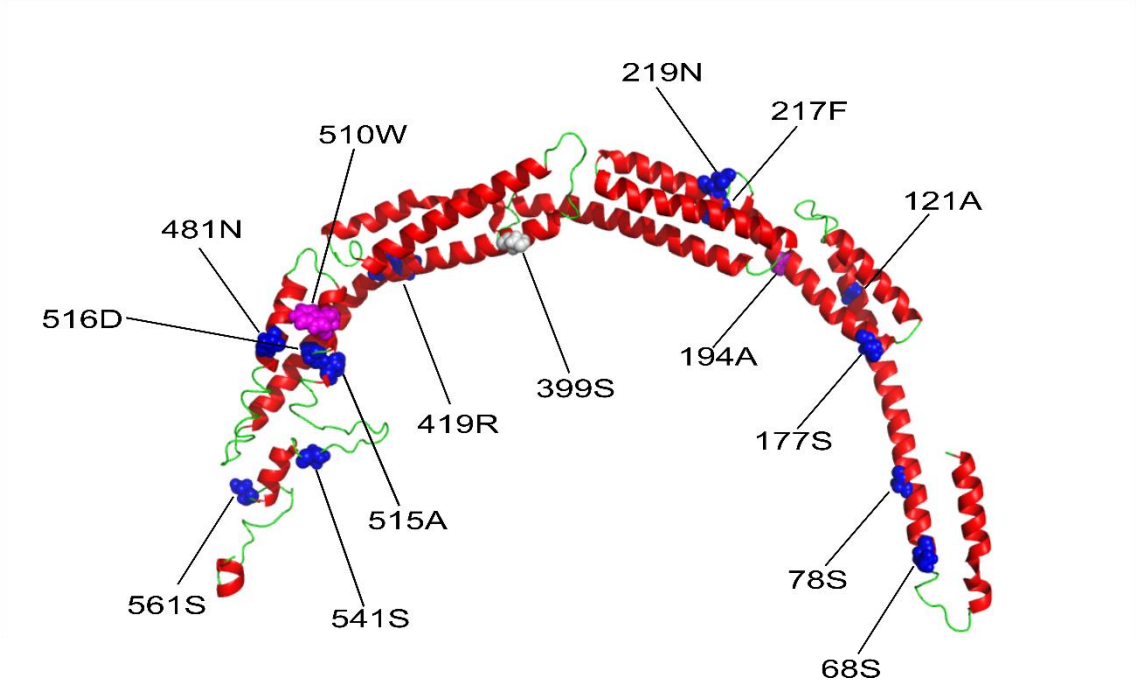

Figure 2. Location of positively selected sites in JFT 3D structure. Sites under positive selection identified using BEB and MEME in the small dataset highlighted in the predicted *Chironex fleckeri* toxin B structure with their side chains shown (*blue spheres*). Purple spheres represents the two sites identified by BEB. Gray spheres represent the only site recovered by the two methods.

TABLES

Table 1. Summary of Codon Substitution Model examined in Codeml for JFTs, with estimates of ω, model parameters, likelihood values, probabilities and positively selected sites.

Model	Parameter estimates	LnL	2 ΔlnL (P value)	Positively selected sites
Branch "large"				
Free ratio		-29627.635760	293.014548 (P < 0.001**)	None
One ratio	ω =0.1831	-29774.143034	0.009977 (P>0.05)	None
Two ratios	*ω0=0.1830; ωII=0.2855*	-29774.138046		None
Fixed two ratios	*ω0=0.1830; ωII=1*	-29774.151102		None
Branch "small"				
Free ratio	ω =0.1455	-22014.069429	117.203336 (P< 0.001**)	None
One ratio		-22072.671097	0.77906 (P>0.05)	None
Two ratios	*ω0=0.1457; ωII=999*$^{+}$	-22072.281567		None
Fixed two ratios	*ω0=0.1457; ωII=1*	-22072.344853		None
Sites "large"				
M1a	P(0,1): 0.00001, 0.99999 *ω*(0,1): 0.0000, 1.00000	-1653.610964		Not allowed
M2a	p(0,1, 2): 0.00000, 0.99999, 0.00001; *ω*(0,1,2): 0.00000, 1.00000, 1.00000	-1653.610654	M2a versus M1a 0 (P=1)	None
M7	p = 5.55877; q = 22.81060	-1600.284171		Not allowed
M8	p_0=0.99999;p=.5.55878; q= 22.81061 (p_1 = 0.00001); *ω*= 2.52155	-1600.284480	M8 versus M7 0 (P=1)	None
Sites "small"				
M1a	P(0,1): 0.92911, 0.07089 *ω*(0,1): 0.12181, 1.00000	-12733.165934		Not allowed
M2a	p(0,1, 2): 0.92911, 0.07089, 0.000; *ω*(0,1,2): 0.112, 1.000, 35.731	-12733.165938	M2a versus M1a 0 (P=1)	Not significant
M7	p = 2.54981; q = 15.79368	-12691.323872		Not allowed
M8	p_0=0.97253;p=3.01868; q=20.3820 (p_1 = 0.02747); *ω*= 1.000	-12686.447241	M8 versus M7 9.75363 (P<0.05*)	Not significant
Branch-sites "large"				
MA(II branch)	p(0,1,2a+2b)=0.0000,0.67760.32240;*ω*(2a+2b,foreground):999^{+}	-1649.989632		Not significant
MA(II branch) (Fixed)	Foreground *ω* = 1	-1602.683371	0 (P=1)	Not allowed
Branch-sites "small"				
MA(II branch)	p(0,1,2a+2b)=0.77569,0.05958,0.16473 ;*ω*(2a+2b,foreground):999^{+}	-12725.825558	10.716399 (P<0.05*)	188 V 191 G 194 A*** 297 S 326 F 332 A 348 K 393 Y 399 S*** 510 W*** 584 W
MA(II branch) (Fixed)	Foreground *ω* = 1	-12731.183757		

+ The w ratio will be estimated to be infinity; * Significant (P<0.05); ** Extremely significant (P<0.01); *** Positively selected sites significant.

Table 2. Codons in the small dataset of JFT alignment identified as positively selected using MEME method with statistical support (0.05 significance), BEB posterior probability values if also identified in Codeml models positively selected.

Codon position	MEME (p value < 0.05)	Codeml MA-BEB pr (ω > 1)
68S	0.0143	-
78G	0.0417	-
121A	0.0146	-
177S	0.0294	-
217F	0.0070	-
219N	0.0443	-
399S	0.0356	0.973
419R	0.0455	-
481N	0.0191	-
515A	0.0464	-
516D	0.0009	-
542S	0.0138	-
561S	0.0091	-

Table 3. Surface accessibility (ASA) of positively selected sites in JFTs. Residues with an ASA ratio greater than 50% are considered to be exposed, while those with an ASA ratio less than 20% are considered to be buried in the surrounding medium (ASA of 21%–39%: cannot be assigned to buried/exposed classes; ASA of 40%-50% are likely to have exposed side chains).

Site	BEB	MEME	ASA	Inference
68S	X	✓	20%	Buried
78G	X	✓	40%	Exposed
121A	X	✓	60%	Exposed
177S	X	✓	10%	Buried
194A	✓	X	0	Buried
217F	X	✓	10%	Buried
219N	X	✓	40%	Exposed
399S	✓	✓	30%	-
419R	X	✓	40%	Exposed
481N	X	✓	40%	Exposed
510W	✓	X	0	Buried
515A	X	✓	20%	Buried
516D	X	✓	40%	Exposed
542S	X	✓	20%	Buried
561S	X	✓	30%	-

Table 4. Positively selected sites on sequence motifs identified by MEME algorithm.

Codon position	MEME Number Motif
68S	-
78G	-
121A	-
177S	-
217F	3
219N	3
399S	-
419R	-
481N	-
515A	-
516D	-
542S	-
561S	-

SUPPLEMENTARY MATERIAL

Table S1. Sequences used in this study. The proteins marked with asterisk (*) were directly obtained from a recent published study (Rachamim *et al*. 2014).

Taxa	Accession Number
Hydra vulgaris	XM_012704160.1
Hydra vulgaris	XM_012699441.1
Hydra vulgaris	XM_002156821.3
Hydra vulgaris	XM_012700622.1
*Hydra_magnipapillata*_221113110	*
*Hydra_magnipapillata*_221113112	*
*Hydra_magnipapillata*_221113919	*
*Hydra_magnipapillata*_221120158	*
*Hydra_magnipapillata*_221131661	*
*Aurelia_aurita*_comp182923_c1_seq1	*
*Aurelia_aurita*_comp191305_c1_seq1	*
*Aurelia_aurita*_comp191305_c1_seq3	*
*Aurelia_aurita*_comp191305_c1_seq5	*
*Aurelia*_TX1	AFK76348.1
*Aurelia*_TX2	AFK76349.1
Carybdea rastonii	BAB12728.1
Carybdea alata	BAB12727.1
*Malo_kingi*_1	ACX30670.1
*Malo_kingi*_2	ACX30671.1

*Chironex_fleckeri*_toxin1	ABS30940.1
*Chironex_fleckeri*_toxin2	ABS30941.1
*Chiropsoides_quadrigatus*_CqTX-A	BAB82520.1
*Chironex_fleckeri*_toxinA	AFQ00676.1
*Chironex_fleckeri*_toxinB	AFQ00677.1

CAPÍTULO 4

Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia)

Ranko Gacesa[1], Ray Chung[2], Simon R. Dunn[3], Andrew J. Weston[4], Adrian Jaimes-Becerra[5], Antonio C. Marques[5,6], André C. Morandini[5], Daslav Hranueli[7], Antonio Starcevic[7], Malcolm Ward[2] and Paul F. Long[1, 8-10*]

[1]Institute of Pharmaceutical Science, King's College London, 150 Stamford Street, London SE1 9NH, United Kingdom.
[2]Proteomics Facility, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AF, United Kingdom.
[3]Coral Reefs Ecosystems Laboratory, School of Biological Sciences, The University of Queensland, St. Lucia, QLD 4072, Australia.
[4]Mass Spectrometry Laboratory, UCL School of Pharmacy, 29/39 Brunswick Square, London WC1N 1AX, United Kingdom.
[5]Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua Matão, Trav. 14, 101, 05508-090 São Paulo, SP, Brasil.
[6]Centro de Biologia Marinha, Universidade de São Paulo, Rodovia Manoel Hypólito do Rego, km. 131,5, 11600-000 São Sebastião, Brasil.
[7]Section for Bioinformatics, Department of Biochemical Engineering, Faculty of Food Technology and Biotechnology, University of Zagreb, Pierottijeva 6, 10000 Zagreb, Croatia.
[8]Department of Chemistry & [9]Brazil Institute, King's College London, Strand, London WC2R 2LS, United Kingdom.
[10]Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Av. Prof. Lineu Prestes, 580, B16, 05508-000 São Paulo, SP, Brasil.
[*]Corresponding author: Dr Paul F. Long, email: paul.long@kcl.ac.uk

ABSTRACT

Background: Gene duplication followed by adaptive selection is a well-accepted process leading to toxin diversification in venoms. However, emergent genomic, transcriptomic and proteomic evidence now challenges this role to be at best equivocal to other processes. Cnidaria are arguably the most ancient phylum of the extant metazoa that are venomous and such provide a definitive ancestral anchor to examine the evolution of this trait.

Methods: Here we compare predicted toxins from the translated genome of the coral Acropora digitifera to putative toxins revealed by proteomic analysis of soluble proteins discharged from nematocysts, to determine the extent to which gene duplications contribute to venom innovation in this reef-building coral species. A new bioinformatics

tool called HHCompare was developed to detect potential gene duplications in the genomic data, which is made freely available (https://github.com/rgacesa/HHCompare).

Results: A total of 55 potential toxin encoding genes could be predicted from the A. digitifera genome, of which 36 (65 %) had likely arisen by gene duplication as evinced using the HHCompare tool and verified using two standard phylogeny methods. Surprisingly, only 22 % (12/55) of the potential toxin repertoire could be detected following rigorous proteomic analysis, for which only half (6/12) of the toxin proteome could be accounted for as peptides encoded by the gene duplicates. Biological activities of these toxins are dominated by putative phospholipases and toxic peptidases.

Conclusions: Gene expansions in A. digitifera venom are the most extensive yet described in any venomous animal, and gene duplication plays a significant role leading to toxin diversification in this coral species. Since such low numbers of toxins were detected in the proteome, it is unlikely that the venom is evolving rapidly by prey driven positive natural selection. Rather we contend that the venom has a defensive role deterring predation or harm from interspecific competition and overgrowth by fouling organisms. Factors influencing translation of toxin encoding genes perhaps warrants more profound experimental consideration.

Keywords: coral, nematocyst, venom, proteome, evolution, hidden Markov model (HMM)

INTRODUCTION

Venoms are usually complex mixtures of peptides and proteins colloquially known as toxins. These toxins can disrupt cellular functions or physiological processes, but venoms differ from poisons in that the venom must be delivered through specialised anatomical structures, such as fangs or stinging devices, that inflict a wound to the target prey or predator. This generally accepted definition includes also that toxins are biosynthesised and the venom then secreted from specialised glands (Fry *et al*., 2009). However, this definition falls short for a group of venomous invertebrates called the cnidarians that do not have any glandular tissues for toxin secretion. Instead, venom is produced by the Golgi apparatus of specialized gland cells (cnidoblasts) that further develop into a delivered cellular organelle called the cnida, contained in cells called

cnidocytes, which is unique to cnidarians and a defining characteristic of this phylum (Marques & Collins, 2004; Morandini *et al*., 2014).

The Cnidaria has two major linages; the Anthozoa (sea anemones and corals) and Medusozoa, comprising the classes Staurozoa (stalked jellyfish), Cubozoa (box jellyfish), Scyphozoa ('true' jellyfish) and Hydrozoa (*Hydra* and relatives including several species of smaller jellyfish) (Marques & Collins, 2004; Van Iten *et al*., 2014, for a recent review). Human envenomation by cnidarians is common and, although seldom life-threatening, fatal contact with certain jellyfish such as the cubozoan *Chironex fleckeri* (the Australian Sea Wasp) is well documented in both the scientific literature and lay press (Fenner & Harrison, 2000). There have been numerous studies characterising the venoms of many animals, but until relatively recently the toxin component and function of cnidarian venoms was poorly studied and near completely unknown (Turk & kem, 2009). Still now, patterns for cnidarian venoms are variable and fuzzy. We have used a high throughput proteomics approach to characterise putative toxins from the nematocysts (a type of cnida) of the coral *Stylophora pistillata* (Weston *et al*., 2012) and the hydrozoan jellyfish *Olindias sambaquiensis* (Weston *et al*., 2013). The biological diversity and sequence similarity between these toxins and those of completely unrelated animals was astounding, suggesting that at least some universal molecular processes leading to toxin diversification might be shared between basal metazoans and diverging lineages of venomous animals.

It is conventionally accepted that venom systems arose by a 'birth and death' process following convergent recruitment of ancestral genes that originally encoded non-toxic physiological functions (Nei *et al*., 1997). These genes underwent duplication followed by rapid hyper-mutation independently in different animals to evolve proteins with cytotoxic functions when expressed in venom gland tissues (Kordis & Gubensek, 2000; Casewell *et al*., 2011). Adaptive selection has retained the most useful paralog genes, which in turn has given rise to larger toxin-specific gene families, for example: phospholipase A2, serine proteases, C-type lectins and coagulation factor V that are present in many venomous animals (Gutierrez & Lomonte, 1995; Kini, 2005; Minh *et al*., 2005; Ogawa *et al*., 2005; Reza *et al*., 2006). Toxins evolved additionally as more species-restricted gene families such as the snake three finger toxins (Fry *et al*., 2003), scorpion cysteine-enriched toxins (Zhijian *et al*., 2006) and the conotoxins of marine cone snails (Duda & Palumbi, 1999). This 'birth and death' hypothesis has been

recently refined, based upon genome sequence data from the non-venomous Burmese python *Python molurus bivittatus* (Reyes-Velasco *et al*., 2014). Using tissue specific gene expression profiling, evidence provides that some genes encoding physiological functions are orthologs of toxin encoding genes which are differentially expressed in many different tissue types of the python. Specific recruitment of such orthologs into venom gland tissue followed by 'birth and death' evolution would result in paralogs where one copy would now encode a toxic function. This explanation might, therefore, account for the large gene expansions seen in venom gland transcripts of xenophidian snakes (Fry, 2005) and that observed in the genome sequence of the highly venomous King Cobra *Ophiophagus hannah* (Vonk *et al*., 2013). Reverse recruitment of toxin encoding genes into non-venom gland tissue with reverse conversion of the gene products back to a physiological role has also been predicted from phylogenetic analyses (Casewell *et al*., 2013) as demonstrated by comparative transcriptome analysis of toxin gene paralogs in venom gland and other tissues of the venomous snake *Bothrops jararaca* (Junqueira-de-Azevedo *et al*., 2015).

Comparative transcriptomics of venomous and non-venomous 'reptiles' has, however, has cast doubt on the extent to which recruitment and reverse recruitment processes play in the evolution of venom systems (Hargreaves *et al*., 2014). The 'restriction hypothesis' confirms previous findings that toxin orthologs are expressed in many tissues of non-venomous 'reptiles', including salivary glands, suggesting that toxin orthologs have not been recruited but had already existed in glandular tissues (Vonk *et al*., 2013; Casewell *et al*., 2012; Junqueira-de-Azevedo *et al*., 2015) . Following gene duplication, paralogs can evolve so that expression of one copy, now encoding a toxic function, is restricted to the venom gland, whilst the original copy encoding a physiological role remains expressed in other tissues (Hargreaves *et al*., 2014). The extent to which gene duplication has impacted on venom innovation has also been challenged because, although gene duplication in cone snail (Chang & Duda, 2012) and snake toxins (Doley et al., 2009) may occur at a fast rate, gene duplication in eukaryotes is generally considered a rare event (Lynch & Conery, 2000). In addition, evaluation of transcriptomic data, together with sequence analysis of the duck-billed platypus (*Ornithorhynchus anatinus*) genome affirms that gene duplication did not contribute significantly to toxin diversification in this venomous mammal (Wong *et al*., 2012a).

Other molecular processes that could lead to toxin diversification *in lieu* of gene duplications have been proposed. For example, although experimentally not proven, exon shuffling of primary mRNA transcripts has been suggested as a mechanism to account for active site variation in amino acid sequences of venom gland serine proteases in the snake *Macrovipera schweizeri* (Siigur et al., 2001). Likewise, homologous recombination at the DNA or RNA levels might account also for sequence variation in Class P-I and P-II snake venom metalloproteinases (SVMP) in *Bothrops neuwiedi* (Moura-da-Silva *et al*., 2011). However, such arguments have been based on mapping to sequences outside of known exon splicing sites in cDNA encoding a different SVMP class, prepared from the venom transcript of a taxonomically distant snake (Sanz *et al*., 2012). Hence, the extent to which toxin diversification can be attributed to processes of gene recruitment and duplication, or indeed recombination and alternative splicing of DNA or RNA, remain largely unexplored. This is principally due to a lack of sequenced genomes of venomous animals from which either true gene duplicates can be identified, or onto which RNA and peptide sequences can be mapped. In direct contrast, post-translational processes including amino acid modifications and protein splicing have both been unequivocally demonstrated to increase conotoxin diversity in marine cone snail venoms (Dutertre *et al*., 2012).

The sequenced genomes of three cnidarians are currently available; these are *Nematostella vectensis* (Putnam *et al*., 2007), *Hydra magnipapillata* (Chapman *et al*., 2010) and *Acropora digitifera* (Shinzato *et al*., 2011). There are also numerous transcriptome libraries for many cnidarians and, in addition to the nematocyst proteomes we have published (Weston *et al*., 2012, 2013) the proteome of *H. magnipapillata* has likewise been reported that includes a description of putative toxins (Balasubramanian *et al*., 2012). We have made freely available annotation of the predicted proteome of *Acropora digitifera* at ZoophyteBase (http://bioserv7.bioinfo.pbf.hr/Zoophyte/registration/login.jsp). A search of this database revealed that the predicted toxins of *A. digitifera* are highly homologous to those toxins of many taxonomically distant venomous animals (Dunlap *et al*., 2013). Having existed since at least the Pre-Cambrian era, Cnidaria are possibly the oldest lineage of extant animals to have evolved means to inject toxins into their prey (Van Iten *et al*., 2013, 2014). If one assumes a single early evolutionary origin of toxin genes, Cnidaria provide a unique ancestral anchor to explore common mechanisms of toxin

innovation that may have subsequently radiated to evolve independently in other venomous animals (Starcevic & Long, 2013). To access the extent to which gene duplication is the primary driver for toxin diversification in the Cnidaria, we herein compare the amino acid sequences of predicted toxins derived from the translated genome of *A. digitifera* to that of putative toxins observed by proteomic analysis of soluble proteins discharged from isolated nematocysts.

RESULTS

Identification of potential toxin encoding genes in the *A. digitifera* genome: The translated genome of *A. digitifera* was searched for homology to known animal toxins in the UniProtKB/Swiss-Prot Tox-Prot dataset. The BLAST search used an e-value cut-off selection criterion of $1.0e^{-5}$ and recovered 950 potential animal toxin homologs. To discriminate potential coral specific toxins from coral proteins with physiological functions, these 950 hits were further filtered using an iterative five step process adapted from previously published methods for Cnidaria toxin identification (Rachamim *et al*., 2015; Brinkman *et al*., 2015). Firstly, only sequences with Reciprocal Blast Best Hit (RBBH) or relaxed RBBH (using the top five BLAST hits for reciprocal BLAST) to sequences in the UniProtKB/Swiss-Prot Tox-Prot dataset with query coverage above 70% were retained. Secondly, BLASTp comparisons were performed against the entire UniProt database supplemented with additional cnidarian protein sequences (Oliveira *et al*., 2012) and, against a customized database constructed using only cnidarian protein sequences contained within UniProt. Only RBBH or relaxed RBBHs hits were retained having a cut-off e-value of less than 1.0e−5 for sequences from both databases. Thirdly, sequences were then manually validated for consistency, and all sequences giving higher scores to non-toxin protein family hits in the cnidarian supplemented UniProt database were discarded. Fourthly, sequences with two or more potential transmembrane domains, or having domain architectures different from known toxins, and Gene Ontology (GO) term assignments unlikely to be related to toxins were also excluded from further examination. Finally, the retained sequences were compared by BLASTp to the translated *A. digitifera* genome, and those with peptide sequences coverage greater than 75 % and e-value homology below $1.0e^{-20}$ were predicted to be bona fide coral specific toxins. A total of 55 potential toxins could be recovered following this five stage filtering process. These 55 potential toxins are shown in Table

1, together with a expectation of likely biological function by inference to a known animal toxin with closest peptide sequence homology. Nearly a quarter (13/55) of the potential A. digitifera toxins shared most similar sequence homology to that of other known cnidarians toxins.

Identification of potential gene duplicates: Evaluation of the role that gene duplication plays in the evolution of toxin diversity requires phylogenetic analysis of sequence data to identify related paralogs from many closely related species. No such data exists for coral species; hence, potential gene duplicates were used as the most likely sequences to be best related to true paralogs. Gene duplicates were identified using a newly developed HMM-HMM based hierarchical clustering tool called HHCompare. Clustering was also performed using standard Maximum Likelihood and Maximum Parsimony phylogenetic methods. All three methods grouped together all of sequences related by identical function (Fig. 1), although there was a slight difference in the number of groups generated by the different methods (Additional file 1:http://www.biomedcentral.com/1471-2164/16/774/additional). Tajima's Test of Neutrality showed that all three methods grouped together all of the same sequences that were related by function (Fig. 1), although there was a slight difference in the number of clusters generated by the different methods (Supplementary Data File 1). Tajima's test of neutrality was performed on each group containing more than 2 domain sequences and, in all cases produced a D statistic greater or equal to 4, indicating balancing selection. When taking the results from the three clustering methods together (Fig. 1), the positioning of 36/55 (65 %) sequences within specific groups inferred that these sequences had arisen following gene duplication events. These 36 sequences could be divided amongst 13 groups with predominantly cytotoxic or toxic protease activities. The remaining 19 sequences could not be grouped and were regarded as singlets, again with mainly cytotoxic activities, possibly involved in affecting haemostasis, immune function, neurotoxicity or toxin maturation.

Identification of potential toxins in the proteome of *A. digitifera* nematocysts: Mass spectral data of peptide fragments obtained from tryptic digests of soluble proteins extracted from discharged nematocysts were first matched for identity to the predicted toxins of A. digitifera (Table 1). Stringent identity criteria of two peptide matches at greater than 95 % sequence similarity were selected that gave just 12 homologous

matches, representing 22 % (12/55) of the potential toxins in the translated genome sequence. A MASCOT search (i.e., two peptide matches with >95 % sequence similarity) of the spectral data for matches to the predicted proteome of Symbiodinium clade B1 was performed to also identify any endosymbiotic algal peptide sequences with homology to predicted A. digitifera toxins. No potential contaminating Symbiodinium clade B1 proteins were identified despite using a BLAST search with a stringent e-value cut off selection criterion of 1.0e−20. The venom toxins of A. digitifera had a relatively narrow profile of predicted biological activities such to include phospholipases and pore forming toxins, toxic peptides and peptides predicted to disrupt haemostasis or immune function. Metalloproteases and other peptidases possibly involved in venom toxin maturation were also annotated as part of the expected toxin proteome. Of the 36 peptides attributed to gene duplication, 6 were detected in the proteome which represented 50 % (6/12) of the total peptides in the expressed venom. Manual validation of mass spectra for annotation of 19 potentially unique A. digitifera coral toxins was assessed by searching the PRIDE proteomics data repository (http://www.ebi.ac.uk/pride/archive/) for the dataset named 'Acropora_Digitifera_Toxins', with sequences in FASTA format are also available from ZoophyteBase (http://bioserv7.bioinfo.pbf.hr/Zoophyte/registration/login.jsp, Dunlap *et al*., 2013).

DISCUSSION

Toxin diversification in venoms is widely accepted to have arisen by convergent recruitment of genes that have evolve independently within the glandular tissues of diverse animal lineages, following common molecular processes of DNA sequence duplication and deletion (Nei *et al*., 1997; Kordis & Gubensek, 2000; Fry *et al*., 2009; Casewell *et al*., 2011; Hargreaves *et al*., 2014Reyes-Velasco *et al*., 2014; Junqueira de Azevedo *et al*., 2015;). Yet, the concept that gene recruitment, sequence duplication and sequence deletion alone are sufficient to explain the surprising chemical diversity of toxins in venoms is increasingly being challenged as genome, transcriptome and proteome data from venomous animals are becoming available (Temple-Smith, 1973; Wong *et al*., 2012; Dutertre *et al*., 2013). Cnidaria is likely to be the most basal of the extant metazoans to be venomous, so we used *Acropora digitifera* for which we had already annotated the predicted proteome (Dunlap *et al*., 2013) to evaluate the extent to

which gene duplication could account for toxin diversification in this reef-building coral.

Here, a BLAST homology search of the *A. digitifera* predicted proteome against the UniProtKB/Swiss-Prot Tox-Prot dataset, followed by a stringent five step process to exclude proteins with possible physiological functions (Oliveira et al., 2012; Brinkman *et al*., 2015; Rachamim *et al*., 2015) uncovered 55 potential toxins with homology to animal toxins (Table1). This was a low number of potential toxin encoding genes in comparison to that of the two venomous vertebrates for which genome sequences are presently available. Such, there were 107 potential toxin encoding genes identified by similarity to known toxins encoded in the genome of the Duck-Billed Platypus *Ornithorhynchus anatinus* (Wong *et al*., 2012) and, 69 predicted toxin encoding genes with homology to toxin families were identified in the genome sequence of the King Cobra *Ophiophagus hannah* (Vonk *et al*., 2013). However, there was a disparity between the higher numbers of predicted toxin encoding genes that had arisen from likely duplication events identified in this study (36/55, 65 %) as compared to much lower numbers of gene duplicates in the Duck-billed platypus and King cobra genomes. Of the 107 platypus genes with significant sequence similarity to known toxins, only 16 (15 %) were likely to have evolved subsequent to a duplication event; this low number would suggest that the venom of the platypus is diversifying slowly and likely under negative selection. Indeed, the 16 gene duplicates were not members of any major known lethal toxin gene families, and so the venom is unlikely to be under strong adaptive (i.e., positive) evolutionary pressure, thereby producing venom of low potency (Wong *et al*., 2012). This would agree with the likely purpose attributed to platypus venom, which is to incapacitate rather than to kill mating competitors (Temple-Smith, 1973), a widespread common sexual selection pattern among mammals. In contrast, the 69 potential toxin encoding genes predicted in the genome of the King cobra have undergone massive expansion, with 30 (i.e., 43 %) likely to have arisen following gene duplication. Of these 30 duplicates, 25 were concentrated in just three major lethal toxin gene families, namely the three-finger toxins, phospholipase A2 and snake venom metalloproteinase enzymes (Vonk *et al*., 2013). This high number of gene duplications is consistent with natural selection for specific prey, which requires highly toxic and lethal venom that is evolving quickly to adapt to molecular co-evolution of prey resistance (Casewell *et al*., 2013).

Evaluation of the role gene duplication plays in the evolution of toxin diversity in basal Metazoa requires bioinformatics methods to identify putative gene paralogs. There are currently two standard approaches based on either comparing the positions of paralogs on phylogenetic tree relationships or by assessing the degree of identity between sequences using BLAST similarity searching methods. Both methods require genomic, transcriptomic or proteomic data obtained from many closely related species in order to identify related paralogs. There are sequenced genomes only for three distantly related cnidarians available in the public domain, and so, tree and BLAST based approaches to identify paralogs is not dependable. Currently available clustering methods such as cd-hit and BLASTClust (ftp://ftp.ncbi.-nih.gov/blast/documents/blastclust.html) from the NCBI– BLAST package (Altschul *et al*., 1990) can be used to infer potential orthology, but do not provide an evolutionary perspective, and such fall short in precision because they use BLAST-like algorithms. Comparison of similarity between groups of potential orthologs based on generating and then comparing hidden Markov models (HMMs) does allow inference of evolutionary distance. However, there are currently no tools available that compare HMMs and then cluster orthologous proteins to allow potential paralogs to be detected within ortholog clusters. For this reason we have developed a new tool called HHCompare. It implements well tested HHsuit programs for HMM generation and HMM vs HMM comparisons (Biegert, *et al*., 2006). HHCompare then uses iterative pairwise HMM vs HMM comparisons to generate related ortholog groups based on high HMM-HMM similarity (e-value cut-off less than $1.0e^{-20}$) and then generates relationship trees to cluster the orthologous groups, thereby allowing potential orthologs in and between cluster groups to be detected. In this study, such a low e-value cut-off would only cluster extremely similar orthologous proteins, and so this approach was considered a proxy for likely gene duplication in the absence of sequences from closely related species. The strength of this clustering compared favourably against two standard methods of approach (Additional file 1: http://www.biomedcentral.com/1471-2164/16/774/additional). The 55 predicted toxins encoded in the A. digitifera genome formed 13 clusters with two or more sequences and 19 singlets (Fig. 1). This requires that an astounding 65 % of the predicted venom of A. digitifera had likely arisen subsequent to gene duplication, which is far greater than the total expansion of toxin genes reported in the King cobra venom (43 % (Vonk *et al*., 2013)). This degree of duplication is nearly equivalent to gene expansions reported for specific toxin families

in other venomous animals. Conotoxin genes are thought to be the most rapidly evolving in the Metazoa with 70 % of the A-superfamily of conotoxin genes having been established by gene duplication (Chang & Duda, 2012). In sharp contrast, genes encoding the sphingomyelinase D toxin in sicariid spiders are believed to be composed of only 4.4 % of gene duplicates (Binford *et al*., 2009). To our knowledge, A. digitifera has the greatest percentage of toxin encoding gene duplications yet reported in the genome of any venomous animal to date.

To assess what adaptive selective pressures might drive and maintain such massive gene expansions in A. digitifera, the expressed venom proteome was determined empirically using high throughput mass spectrometric protein analysis. When matched predicted toxins against the translated proteome sequence, and surprisingly only 22 % (12/55) of the predicted proteome could be identified using strict spectral identification parameters. Although peptides likely to be products from gene duplicates accounted for 50 % (6/12) of the toxic proteome, the high number of potential toxins not detected in the venom proteome might reflect poor promotor recognition and therefore weak expression of very recently duplicated genes such that protein abundance is less than the detection limits of the proteomics method (Force *et al*., 1991). Such a high number of gene duplicates would suggest that the venom is evolving rapidly under adaptive, positive selection. However, with so many of the gene duplicates not seemingly expressed in the empirically determined proteome would, in fact, indicate contrarily that the venom of A. digitifera has low toxicity since it is evolving gradually under negative selection. This is in broad agreement with data comparing multiple alignments of amino acid sequences and calculations of amino acid substitution rates, particularly for the sea anemone peptide neurotoxins and pore-forming toxins, which show these cnidarian toxins are under negative selection and thus are highly conserved (Jouiaei *et al*., 2015). Likewise, critical examination of the evolution of three species across cnidarian lineages (the anthozoan sea anemone Anemonia viridis (Actinaria), the scyphozoan jellyfish Aurelia aurita and the hydrozoan Hydra magnipapillata) agrees also with our data that venom of the anthozoan Acropora digitifera (Scleractinia) shows little evidence for diversification through positive selection (Rachamim *et al*., 20150. The putative biological activities of the toxins in both the predicted and observed A. digitifera venom were dominated by cytotoxic phospholipases and pore forming toxins (Table 1 and Fig. 1). This is not unusual

compared to the known or predicted pharmacological effects of toxins in other cnidarian venoms. For example, in anthozoans, of which sea anemone venoms are the most widely studied in all of the Cnidaria, their venoms are composed mainly of pore forming toxins and peptide neurotoxins (Radwan *et al*., 2004). Other anthozoan venoms are less widely studied, but our proteomic analysis of toxins from the coral Stylophora pistillata (Scleractinia) predicts that in this coral species venoms are also composed predominantly of cytotoxic peptides and neurotoxins (Weston *et al*., 2012). The venoms of hydrozoans, such as those of the genus Millepora (commonly known as 'fire corals' and well known for human envenomation causing sever irritation) and Hydra, are composed mainly of cytolysins, phospholipase and haemolytic enzymes (Radwan *et al*., 2004). A. digitifera does feed on microscopic phytoplankton and zooplankton, however, like all of the reef-building corals, A. digitifera has evolved an endosymbiotic metabolic partnership with photosynthetic dinoflagellates of the genus Symbiodinium (Dinophyceae) which is essential for survival in the nutrient-poor waters of tropical marine environments (Muscatine, 1990; Stanley, 2006). The biological relevance of a largely cytolytic toxic arsenal could reflect a possible defensive role to deter fish predation and death by fouling organisms, including attack by coral-excavating sponges (Clionidae) which are strong competitors of corals for space on the reef shelf (Hutchings, 1986; Zundelevich *et al*., 2007; Carballo *et al*., 2013). Biochemical studies to assign specific pore-forming activities to the *A. digitifera* cytolysins will require in future a comprehensive comparative review of pore-forming toxins in Cnidaria to better understand the provenance and biological relevance of these toxins to the life history strategy of these animals (Glasser *et al*., 2014). It is a well accepted concept that toxin gene acquisition follows duplication of genes encoding non-toxic physiological functions (Fry *et al*., 2009; Nei *et al*., 1997). It follows that the toxin encoding genes that were considered as singlets in this study would have most likely have arisen following gene duplication that occurred in the very distant past such that strict evidence for duplication events could not be detected with the methods employed here. Developing an evolutionary clock to determine if the timing of gene duplication events and emergence of specific toxin gene families is correlated with a transition of cnidarians from sessile animals in photoautotrophic symbiosis to free living heterotrophic lineages is worthy of future research.

CONCLUSIONS

This is the first study to combine genome analysis and proteomics data to critically examine venom innovation in the Cnidaria and the relevance of gene duplication in toxin diversification in particular. After filtering proteins with likely non-toxic physiological function, 55 potentially unique coral toxins have been described. Exploring selection pressures and processes driving the evolution of venom is problematic in Cnidaria since few genomes of related species have been sequenced. Here we exemplify a new bioinformatics tool called HHCompare that overcomes the severity of this impediment. Using this tool, predicted toxin encoding genes of the coral A. digitifera could be divided into orthologous groups that are the closest representation to gene duplicates currently possible, which is consistent with groupings determined by conventional phylogenetic methods. Of the 55 toxins, 36 (65 %) are likely established by gene duplication, which represents the largest gene expansion as a percentage proportion of all toxin encoding genes identified in the genome in any venomous animal reported to date. Only 22 % of these peptides were detected in the expressed proteome of discharged nematocysts, suggesting that the venom had evolved for predator defence rather than an offensive role for prey capture. Biochemical validation of toxin activities is now warranted so that full annotation of A. digitifera coral specific toxins can be deposited in publically available protein databases. Gene expansion by gene duplication appears crucial to toxin evolution in the basal Metazoa such as exemplified by the Cnidaria. Factors influencing translation of these gene products to enhance venom potency provides a fascinating avenue for further study.

METHODS

Isolation of nematocysts from coral: Fragments of 3 colonies of the hermatypic coral *A. digitifera* were collected from reef flat sites adjacent to Heron Island Research Station (S 23° 13′ 30″, E 15° 11′ 54″), Great Barrier Reef, Australia in November 2013 and were immediately snap frozen in LN_2 for transport to the laboratory. The coral fragments were airbrushed on ice with 30 mL of Ca^{2+} free artificial seawater (pH 8.2/ 32 ppt) for tissue removal. The homogenised tissue slurry (2 mL) was placed on top of a dilution gradient density column consisting of 2 mL each of 30%, 50%, 70% and 90% (w/v) polyvinyl pyrrolidone (Percoll®: Sigma) in artificial seawater and cooled on ice for 20 min prior to centrifugation at 4 °C for 10 min at 280 x *g*. Following

centrifugation, the 50 - 70% layer that contained the highest concentration of undischarged nematocysts was collected and then freeze dried (MicroModulo-230 freeze drier in combination with a RVT100 refrigerated vapourtrap, Thermo Savant). Corals were collected under permit G12/35434.1 issued by the Great Barrier Reef Marine Park Authority and coral material was transferred to the United Kingdom in accordance with CITES institutional permits AU053 and BG029.

Proteomics: To extract soluble proteins, 500 µL of extraction buffer containing 50 mM triethylammonium bicarbonate, 0.04% (w/v) SDS, 1 x Complete Mini Protease Inhibitors (Roche) and 1 x Complete Mini Phosphatase Inhibitors reagent (Roche) was added to the freeze-dried coral material. The material was vortex for 1 min and then placed on ice for 1 min; the procedure was repeated 10 times. The material was disrupted with a probe sonicator (model: VC250, Sonics & Materials Inc.) whilst on ice for a total of 15 sec using a duty cycle of 40% and an output of 3. The material was centrifuged at 13000 x *g* for 15 min at 4 ^{o}C. The protein concentration in the cleared supernatant was then measured by Nanodrop spectroscopy (Thermo Scientific) by averaging the results of three determinations. An extract portion containing 30 µg of soluble protein in 2 x Laemmli buffer (Sigma) was heated for 10 min at 90 ^{o}C and loaded onto a 4-12% (w/v) NuPAGE Novex Bis-Tris gel (Life Technologies) for separation by 1D SDS-PAGE electrophoresis using 2-(*N*-morpholino)-ethanesulfonic acid (MES) buffer alongside Novex SeeBlue Plus2 pre-stained molecular weight standards. Electrophoresis was carried out at 150 V for approximately 100 min. The gel was fixed, Coomassie Blue-stained, de-stained and visualised by scanned image. The entire gel lane was sectioned into 15 equal portions and each section was divided into 2 mm^2 pieces for in-gel digestion. Briefly, cysteine residues were reduced with 10 mM dithiothrietol and alkylated with 55 mM iodoacetamide in 100 mM ammonium bicarbonate to form stable carbamidomethyl derivatives. Trypsin (Promega) solution was added to the gel sections at 13 ng/µL in 50 mM ammonium bicarbonate and digestion was carried out at 37 ^{o}C overnight. The supernatant was retained and the peptides were extracted from the gel sections by two washes with 50 mM ammonium bicarbonate and acetonitrile. Each wash involved shaking for 15 min before collecting the peptide extract and pooling with the initial supernatant. Pooled peptide extracts were then lyophilised. Lyophilised peptides were re-suspended in 30 µL of 50 mM ammonium bicarbonate per gel section prior to LC-MS/MS analysis with 10 µL of each

sample injected. Samples were analysed sequentially beginning with the largest molecular weight region on a Thermo Fisher Scientific Orbitrap Velos Pro mass spectrometer coupled to an EASY-nLC II (Thermo Fisher Scientific) nano-liquid chromatography system. Samples were trapped on a 0.1 x 20 mm EASY-Column packed with C18-bonded ultrapure silica, 5 µm (Thermo Fisher Scientific) and separated on a 0.075 x 100 mm EASY-Column packed with C18-bonded ultrapure silica, 3 µm (Thermo Fisher Scientific). Columns were equilibrated in 95% buffer A (99.9% deionised water, 0.1% formic acid) and 5% buffer B (99.9% acetonitrile, 0.1% formic acid). Peptides were resolved over 50 min at a flow rate of 300 nL/min with a gradient of 5% to 40% buffer B for 40 min followed by a gradient of 40% to 80% buffer B for 5 min and held at 80% buffer B for a further 5 min. Mass spectra ranging from 400 to 1800 Da (m/z) were acquired in the Orbitrap at a resolution of 30,000 and the 20 most intense ions were subjected to MS/MS by CID fragmentation in the ion trap selecting a threshold of 5,000 counts. The isolation width of precursor selection was 2 units and the normalised collision energy for peptides was 35. Automatic gain control settings for FTMS survey scans were 10^6 counts and FT MS/MS scans were 10^4 counts. Maximum acquisition time was 500 ms for survey scans and 250 ms for MS/MS scans. Charge-unassigned and +1 charged ions were excluded for MS/MS analysis. Raw MS data were processed for database spectral matching using Proteome Discoverer (Thermo Scientific) software. MASCOT was used as the search algorithm with the variable modifications: carbamidomethylation of cysteine and oxidation of methionine. A digestion enzyme of trypsin was set allowing up to three missed cleavages. A parent ion tolerance of 10 ppm and a fragment ion tolerance of 0.5 Da were used.

Bioinformatics: The peptide sequences for the approximately 5000 toxins deposited in the UniProtKB/Swiss-Prot Tox-Prot dataset (www.uniprot.org/program/Toxins , (Jungo *et al*., 2012)) were downloaded in FASTA format. Likewise, the predicted proteomes of A. digitifera (http://marinegenomics.oist.jp/genomes/downloads?project_id=3 , (Shinzato *et al*., 2011)) and Symbiodinium clade B1 (http://marinegenomics.oist.jp/genomes/downloads?project_id=21 , (Shoguchi *et al*., 2013)) were also downloaded in FASTA format and the three datasets were used as query searches for MS/MS spectra. All dataset search results were reviewed by loading the Mascot result files into Scaffold 4 (www.proteomesoftware.com). BLASTp searches were performed to assess local similarities between sequences in the A. digitifera and Symbiodinium clade B1 datasets and the UniProtKB/Swiss-Prot Tox-Prot dataset using

program version 2.2.27+ from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast,executables/blast+/2.2.27/ , (Altschul *et al*., 1990)). The outputs from these comparisons were parsed and filtered using a custom assembled program written in Python (www.python.org) to select for high scoring segment pairs with e-values selected with a cut-off value below $1.0e^{-5}$. Sequences of high scoring segment pairs were filtered to remove proteins with likely physiological functions involving Reciprocal Blast Best Hit (RBBH) analysis (Altenhoff & Dessimoz, 2009) domain architecture prediction using InterProScan5 (http://www.ebi.ac.uk/Tools/pfa/iprscan5/), a search of gene ontology terms (http://www.ebi.ac.uk/QuickGO/) and prediction of transmembrane domains using TMHMM Server 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). Grouping of the truncated high scoring segment pairs used our new Hidden Markov Model (HMM) based comparative software designated 'HMMCompare' that is assembled in Python. 'HMMCompare' is freely available at http://bioserv.pbf.hr/HHCompare-master.zip and is implemented using programs from the HHsuite version 2.0 compiled for the Debian based Linux OS (http://wwwuser.gwdg.de/~compbiol/data/hhsuite/releases/ , (Biegert *et al*., 2006). Multiple alignments of the truncated sequences were constructed using ClustalW version 2.1 compiled for the Debian based Linux OS (ftp://ftp.ebi.ac.uk/pub/software/clustalw2/2.1/). Phylogenetic clustering was also performed using Maximum Likelihood and Maximum Parsimony methods in MEGA 6.0 (Tamura *et al*., 2013) with multiple alignments generated using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/). The clusters were tested for neutral evolution using Tajima's Test of Neutrality (Tajima, 1989) implemented in MEGA 6.0.

COMPETING INTERESTS

The author(s) declare that they have no competing interests

AUTHORS' CONTRIBUTIONS

SRD collected the coral specimens and prepared the nematocysts. RC, AJW and MW carried out the proteomics. RG, DH and AS carried out the bioinformatics analysis. AJB, AC Marques and AC Morandini participated in the data interpretation. PFL conceived the study, and participated in its design and coordination and wrote the

manuscript. All authors helped to draft the manuscript and have approved the final version.

ACKNOWLEDGMENTS

We are indebted to Dr. Walter C. Dunlap and Prof[a]. Dr[a]. *Ana Maria Moura da Silva* for their scientific expertise and guidance during the course of this research; we thank also Prof. Dr. Adalberto Pessoa Júnior and Prof. Dr. Gabriel Padilla for reviewing the manuscript. This work was supported by the United Kingdom Medical Research Council (MRC grant G82144A to R. Gacesa, D. Hranueli and P. F. Long), *the Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP grants 2010/50174-7 to A. C. Morandini and 2011/50242-5 to A. C. Marques), and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq grant 301039/2013-5 to A. C. Morandini). This is a contribution to the NP-BioMar program at the Universidade de São Paulo.

REFERENCES

Altenhoff, A.M. & Dessimoz, C. 2009. Phylogenetic and functional assessment of orthologs inference projects and methods. PLoS Computation Biology 5:e1000262.

Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403-10.

Balasubramanian, P.G.; Beckmann, A.; Warnken, U.; Schnölzer, M.; Schüler, A.; Bornberg-Bauer, E.; Holstein, T.W. & Özbek, S. 2012. Proteome of *Hydra* nematocyst. Journal of Biological Chemistry 287: 9672–9681.

Biegert, A.; Mayer, C.; Remmert, M.; Söding, J. & Lupas, A.N. 2006. The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic Acids Research 34:335-339.

Binford, G.J.; Bodner, M.R.; Cordes, M.H.J.; Baldwin, K.L.; Rynerson, M.R.; Burns, S.N. & Zobel-Thropp, P.A. 2009. Molecular evolution, functional variation, and proposed nomenclature of the gene family that includes sphingomyelinase D in sicariid spider venoms. Molecular Biology and Evolution 26:547-66.

Brinkman, D.L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D. & Mulvenna, J. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16: 1–15.

Carballo, J.L.; Bautista, E.; Nava, H.; Cruz-Barraza, A. & Chávez, J.A. 2013. Boring sponges, an increasing threat for coral reefs affected by bleaching events. Ecology and Evolution 3:872–86.

Casewell, N.R.; Huttley, G.A. & Wüster, W. 2012. Dynamic evolution of venom proteins in squamate reptiles. Nature Communications 3:1066.

Casewell, N.R.; Wagstaff, S.C.; Harrison, R.A.; Renjifo, C. & Wüster, W. 2011. Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. Molecular Biology and Evolution28:2637–49.

Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A. & Fry, B.G. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends in Ecology and Evolution 28: 219–229.

Chang, D. & Duda, T.F. 2010. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Molecular Biology and Evolution 29:2019–2029.

Chapman, J.A.; Kirkness, E.F.; Simakov, O.; Hampson, S.E.; Mitros, T.; Weinmaier, T.; Rattei, T.; Balasubramanian, P.G.; Borman, J.; Busam, D.; Disbennett, K.; Pfannkoch, C.; Sumin, N.; Sutton, G.G.; Viswanathan, L.D.; Walenz, B.; Goodstein, D.M.; Hellsten, U.; Kawashima, T.; Prochnik, S.E.; Putnam, N.H.; Shu, S.; Blumberg, B.; Dana, C.E.; Gee, L.; Kibler, D.F.; Law, L.; Lindgens, D.; Martinez, D.E.; Peng, J.; Wigge, P.A.; Bertulat, B.; Guder, C.; Nakamura, Y.; Ozbek, S.; Watanabe, H.; Khalturin, K.; Hemmrich, G.; Franke, A.; Augustin, R.; Fraune, S.; Hayakawa, E.; Hayakawa, S.; Hirose, M.; Hwang, J.S.; Ikeo, K.; Nishimiya-Fujisawa, C.; Ogura, A.; Takahashi, T.; Steinmetz, P.R.; Zhang, X.; Aufschnaiter, R.; Eder, M.K.; Gorny, A.K.; Salvenmoser, W.; Heimberg, A.M.; Wheeler, B.M.; Peterson, K.J.; Böttger, A.; Tischler, P.; Wolf, A.; Gojobori, T.; Remington, K.A.; Strausberg, R.L.; Venter, J.C.; Technau, U.; Hobmayer, B.; Bosch, T.C.; Holstein, T.W.; Fujisawa, T.; Bode, H.R.; David, C.N.; Rokhsar, D.S. & Steele R.E. 2010. The dynamic genome of *Hydra*. Nature 464:592-596.

Doley, R.; Mackessy, S.P. & Kini, R.M. 2009. Role of accelerated segment switch in exons to alter targeting (ASSET) in the molecular evolution of snake venom proteins. BMC Evolutionary Biology **9**:146.

Duda, T.F. & Palumbi, S.R. 1999. Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. Proceedings of the National Academy Sciences USA 96:6820-6823.

Dunlap, W.C.; Starcevic, A.; Baranasic, D.; Diminic, J.; Zucko, J.; Gacesa, R.; Van Oppen, M.; Hranueli, D.; Cullum, J. & Long, P.F. 2013. KEGG orthology-based

annotation of the predicted proteome of *Acropora digitifera*: ZoophyteBase - an open access and searchable database of a coral genome. BMC Genomics 14:509.

Dutertre, S.; Jin, A.H.; Kaas, Q.; Jones, A.; Alewood, P.F. & Lewis, R.J. 2013. Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. Molecular Cellular and Proteomics 12:312-29.

Fenner, P.J. & Harrisson, S.L. 2000. Irukandji and *Chironex fleckeri* jellyfish envenomation in tropical Australia. Wilderness & Environmental Medicine 11:233-40.

Force, A.; Lynch, M.; Pickett, F.B.; Amores, A.; Yan, Y.L. & Postlethwait, J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531-45.

Frazão, B.; Vasconcelos, V. & Antunes, A. 2012. Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: an overview. Marine Drugs 10:1812-51.

Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.; Lewis, R.J.; Norton, R.S.; Renjifo, C. & de la Vega, R.C.R. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10: 483–511.

Fry, B.G.; Wüster, W.; Kini, R.M.; Brusic, V.; Khan, A.; Venkataraman, D. & Rooney, A.P. 2003. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. Journal of Molecular Evolution 57:110-29.

Fry, B.G. 2005. From genome to "venome": Molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. Genome Research 15:403–20.

Gutiérrez, J.M. & Lomonte, B. 1995. Phospholipase A2 myotoxins from *Bothrops* snake venoms. Toxicon 33:1405-24.

Hargreaves, A.D.; Swain, M.T.; Hegarty, M.J.; Logan, D.W. & Mulley, J.F. 2014. Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. Genome Biology and Evolution 6:2088-95.

Hutchings, P.A. 1986. Biological destruction of coral reefs. Coral Reefs 4:239-52.

Jouiaei, M.; Sunagar, K.; Gross, A.F.; Scheib, H.; Alewood, P.F.; Moran, Y. & Fry, B.G. 2015. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Molecular Biology and Evolution 32: 1598–1610.

Jungo, F., L. Bougueleret, I. Xenarios, & S. Poux, 2012. The UniprotKB/Swiss-prot Tox-Prot program: a central hub of integrated venom protein data. Toxicon 60: 551–557.

Junqueira-de-Azevedo, I.L.; Val Bastos, C.M.; Ho, P.L.; Luna, M.S.; Yamanouye, N. & Casewell, N.R. 2015. Venom-related transcripts from *Bothrops jararaca* tissues provide novel molecular insights into the production and evolution of snake venom. Molecular Biology and Evolution 32:754-766.

Kini, R.M. 2005. Serine proteases affecting blood coagulation and fibrinolysis from snake venoms. Pathophysiology of Haemostasis and Thrombosis 34:200-204.

Kordiš, D. & Gubenšek, F. 2000. Adaptive evolution of animal toxin multigene families. Gene 261:43–52.

Lynch, M. & Conery, J.S. 2000. The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155.

Marques, A.C. & Collins, A.G. 2004. Cladistic analysis of Medusozoa and cnidarian evolution. Invertebrate Biology 123: 23–42.

Minh Le, T.N.; Reza, M.A.; Swarup, S. & Kini, R.M. 2005. Gene duplication of coagulation factor V and origin of venom prothrombin activator in *Pseudonaja textilis* snake. Journal of Thrombosis and Haemostasis 93:420-9.

Morandini, A.C; Marques, A.C. & Custódio, M.R. 2014. Phylum Porifera and Cnidaria. In: Gopalakrishnakone P, editor. Tonixology – Marine and freshwater toxins. Springer Science Business Media Dordrecht: Netherlands 1-24.

Moura-da-Silva, A.M.; Furlan, M.S.; Caporrino, M.C.; Grego, K.F.; Portes-Junior, J.A.; Clissa, P.B.; Valente, R.H. & Magalhães, G.S. 2011. Diversity of metalloproteinases in *Bothrops neuwiedi* snake venom transcripts: evidences for recombination between different classes of SVMPs. BMC Genetics 12:94.

Muscatine, L. 1990. The role of symbiotic algae in carbon and energy flux in reef corals. In: Dubinsky Z, editor. Ecosystems of the World: Coral Reefs. Elsevier, Amsterdam 75-84.

Nei, M.; Gu, X. & Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proceedings of the National Academy of Sciences USA 94:7799-7806.

Ogawa, T.; Chijiwa, T.; Oda-Ueda, N. & Ohno, M. 2005 Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom. Toxicon 45:1-14.

Oliveira, J.S.; Fuentes-Silva, D. & King, G.F. 2012. Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. Toxicon 60:539-550.

Putnam, N.H.; Srivastava, M.; Hellsten, U.; Dirks, B.; Chapman, J.; Salamov, A.; et al. 2007. Sea anemone genome reveals ancestral Eumetazoan gene repertoire and genomic organization. Science 317:86-94.

Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T. & Sher, D. 2014. The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. Molecular Biology and Evolution 32: 740–753.

Radwan, F.F. & Aboul-Dahab, H.M. 2004. Milleporin-1, a new phospholipase A2 active protein from the fire coral *Millepora platyphylla* nematocysts. Comparative Biochemistry and Physiology C Toxicology & Pharmacology 139:267-72.

Reyes-Velasco, J.; Card, D.C.; Andrew, A.L.; Shaney, K.J.; Adams, R.H.; Schield, D.R.; Casewell, N.R.; Mackessy, S.P. & Castoe, T.A. 2014. Expression of venom gene homologs in diverse python tissues suggests a new model for the evolution of snake venom. Molecular Biology and Evolution 32:173-83.

Reza, M.A.; Minh Le, T.N.; Swarup, S. & Kini, R.M. 2006. Molecular evolution caught in action: gene duplication and evolution of molecular isoforms of prothrombin activators in *Pseudonaja textilis* (brown snake). Journal of Thrombosis and Haemostasis 4:1346-53.

Sanz, L.; Harrison, R.A. & Calvete, J.J. 2012 First draft of the genomic organization of a PIII-SVMP gene. Toxicon 60:455-69.

Shinzato, C.; Shoguchi, E.; Kawashima, T.; Hamada, M.; Hisata, K.; Tanaka, M.; Manabu, F.; Fujiwara, M.; Koyanagi, R.; Ikuta, T.; Fujiyama, A.; Miller, D. & Satoh, N. 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature 476:320-3.

Shoguchi, E.; Shinzato, C.; Kawashima, T.; Gyoja, F.; Mungpakdee, S.; Koyanagi, R.; Takeuchi, T.; Hisata, K.; Tanaka, M.; Fujiwara, M.; Hamada, M.; Seidi, A.; Fujie, M.; Usami, T.; Goto, H.; Yamasaki, S.; Arakaki, N.; Suzuki, Y.; Sugano, S.; Toyoda, A.; Kuroki, Y.; Fujiyama, A.; Medina, M.; Coffroth, M.A.; Bhattacharya, D. & Satoh, N. 2013. Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. Current Biology 23:1399-408.

Siigur, E.; Aaspõllu, A. & Siigur, J. 2001. Sequence diversity of *Vipera lebetina* snake venom gland serine proteinase homologs - result of alternative splicing or genome alteration. Gene 263:199-203.

Stanley, G.D. 2006. Photosymbiosis and the evolution of modern coral reefs. Science 312: 857–8.

Starcevic, A. & Long, P.F. 2013. Diversification of animal venom peptides-were jellyfish amongst the first combinatorial chemists?. ChemBioChem 14: 1407–1409.

Tajima, F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A. & Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725-2729.

Temple-Smith, P. 1973. *Seasonal Breeding Biology of the Platypus, Ornithorhynchus anatinus with Special Reference to the Male.* Ph.D. thesis, Australian National University, Canberra.

Turk, T. & Kem, W.R. 2009. The phylum Cnidaria and investigations of its toxins and venoms until 1990. Toxicon 54: 1031–1037.

Van Iten, H; Leme, J.M.; Marques, A.C. & Simões, M.G. 2013. Alternative interpretations of some earliest Ediacaran fossils from China. Acta Palaeontologica Polonica 58:111-113.

Van Iten, H.; Marques, A.C.; Leme, J.D.M.; Pacheco, M.L.F. & Simões, M.G. 2014. Origin and early diversification of the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's Proterozoic-Cambrian history. Palaeontology 57: 677–690.

Vonk, F.J.; Casewell, N.R.; Henkel, C.V.; Heimberg, A.M.; Jansen, H.J.; McCleary, R.J.; *et al*. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proceedings of the National Academy of Sciences USA 110:20651-20656.

Weston, A.J.; Dunlap, W.C.; Shick, J.M.; Klueter, A.; Iglic, K.; Vukelic, A.; Starcevic, A.; Ward, M.; Wells, M.L.; Trick, C.G. & Long, P.F. 2012. A profile of an endosymbiont-enriched fraction of the coral *Stylophora pistillata* reveals proteins relevant to microbial-host interactions. Molecular & Cellular Proteomics 11. M111.015487.

Weston, A.J.; Chung, R.; Dunlap, W.C.; Morandini, A.C.; Marques, A.C.; Moura-da-Silva, A.M.; Ward, M.; Padilla, G.; da Silva, L.F.; Andreakis, N. & Long, P.F.

2013. Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish *Olindias sambaquiensis*. Toxicon 71: 11–17.

Wong, E.S.; Papenfuss, A.T.; Whittington, C.M; Warren, W.C. & Belov, K. 2012. A limited role for gene duplications in the evolution of *Platypus* venom. Molecular Biology and Evolution 29:167-77.

Zhijian, C.; Feng, L.; Yingliang, W.; Xin, M. & Wenxin, L. 2006. Genetic mechanisms of scorpion venom peptide diversification. Toxicon 47:348-55.

Zundelevich, A.; Lazar, B. & Ilan, M. 2007. Chemical versus mechanical bioerosion of coral reefs by boring sponges – lessons from *Pione cf. vastifica*. Journal of Experimental Biology 210:91-6.

FIGURES

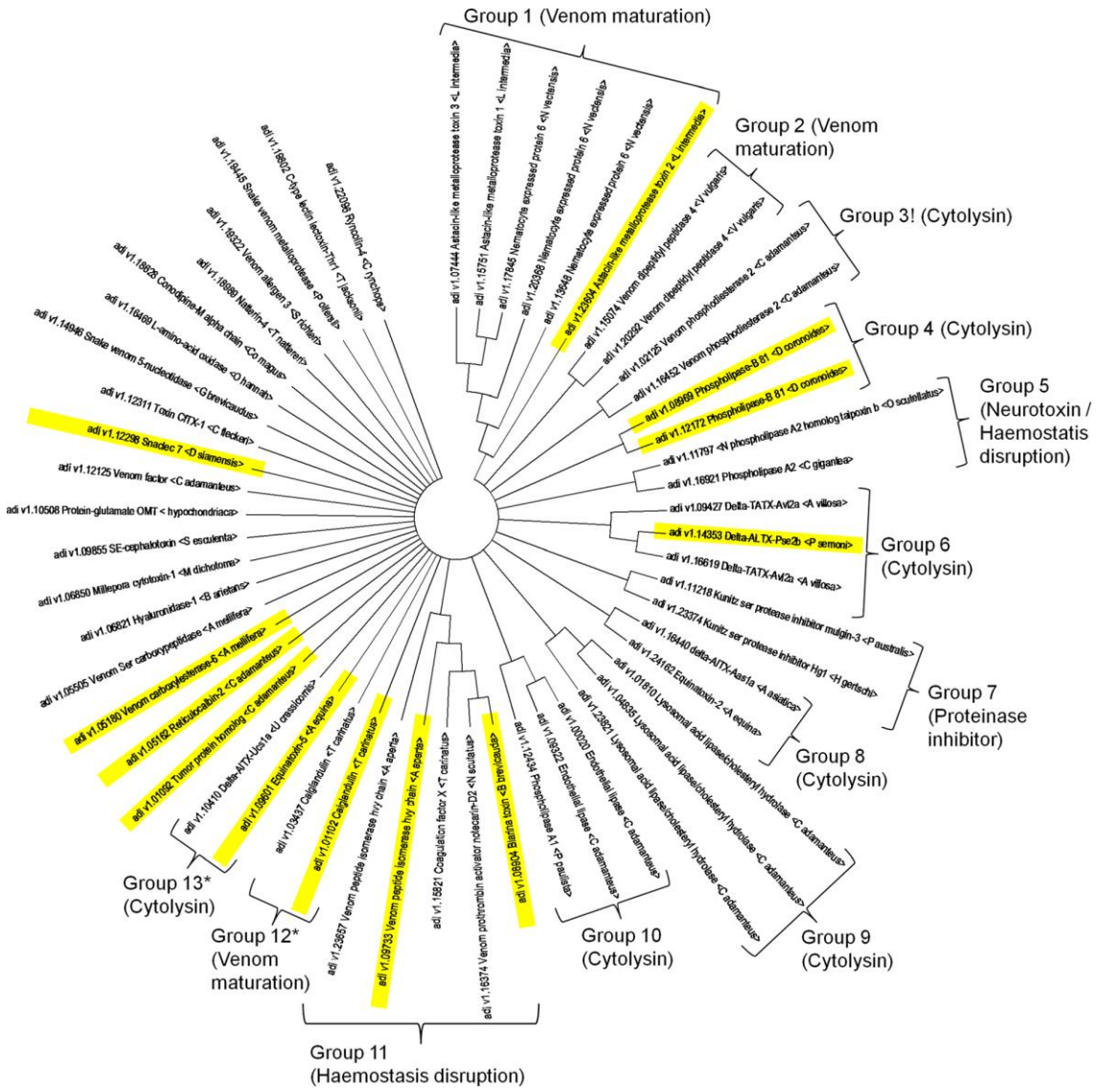

Figure 1. Gene duplication prediction by clustering of Acropora digitifera predicted toxins. Clustering was performed using the HHCompare tool and verified by Maximum Likelihood and Maximum Parsimony based methods. Groups marked with * are detected by Maximum Likelihood and Maximum Parsimony based clustering, while groups marked ! were not detected. Proteins highlighted in yellow were detected by high throughput mass spectrometric protein analysis of soluble proteins from discharged nematocysts.

TABLES

Table 1: Predicted venom proteome of potential toxins from the translated genome sequence of Acropora digitifera.

A. digitifera **protein accession number**	**Predicted biological effect**	**Sequence homology (e-value)**	**Uniprot accession number**	**Toxin with closest homology**	**Example of animal species with closest homology**
adi_v1.16452	Cytolysin	5.00E-55	J3SBP3	Phosphodiesterase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.02125	Cytolysin	2.0E-137	J3SBP3	Phosphodiesterase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.08969	Cytolysin	1.0E-76	F8J2D3	Phospholipase-B	*Drysdalia coronoides* (White-lipped snake)
adi_v1.12172	Cytolysin	4.0E-121	F8J2D3	Phospholipase-B	*Drysdalia coronoides* (White-lipped snake)
adi_v1.14353	Cytolysin	2.0E-83	[43]	Δ-ALTX-Pse	*Phyllodiscus semoni* (Night sea anemone)
adi_v1.09427	Cytolysin	7.0E-93	[43]	Δ-TATX-Avl2a	*Actineria villosa* (Okinawan sea anemone)
adi_v1.16619	Cytolysin	3.00E-88	[43]	Δ-TATX-Avl2a	*Actineria villosa* (Okinawan sea anemone)
adi_v1.16440	Cytolysin	1.00E-12	[43]	Δ-AITX-Aas1a (Bandaporin)	*Anthopleura asiatica* (Giant green sea anemone)
adi_v1.24162	Cytolysin	3.00E-11	P61914	Equinatoxin-2 (Actinoporin)	*Actinia equina* (Beadlet sea anemone)
adi_v1.04835	Cytolysin	2.0E-120	J3SDX8	Lipase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.23821	Cytolysin	3.00E-101	J3SDX8	Lipase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.01810	Cytolysin	7.0E-65	J3SDX8	Lipase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.09322	Cytolysin	4.0E-47	J3RZ81	Lipase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.00020	Cytolysin	1.0E-39	J3RZ81	Lipase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.12434	Cytolysin	1.00E-43	A2VBC4	Phospholipase A1	*Polybia paulista* (Neotropical social wasp)
adi_v1.09601	Cytolysin	2.0E-16	Q93109	Equinatoxin-5 (Actinoporin)	*Actinia equina* (Beadlet sea anemone)

adi_v1.10410	Cytolysin	4.0E-52	Reference [43]	Δ-AITX-Ucs1a (Urticinatoxin)	*Urticina crassicornis* (Christmas sea anemone)
adi_v1.12125	Cytolysin	2.00E-26	J3S836	Phosphodiesterase	*Crotalus adamanteus* (Eastern diamondback
adi_v1.06821	Cytolysin	2.0E-78	A3QVN9	Hyaluronidase	*Bitis arietans* (African puff adder)
adi_v1.16469	Cytolysin	5.00E-13	P81383	L-amino-acid oxidase	*Ophiophagus hannah* (King cobra)
adi_v1.12311	Cytolysin	2.0E-16	A7L035	Toxin CfTX-1	*Chironex fleckeri* (Box jellyfish)
adi_v1.16921	Disrupts	3.00E-28	D2X8K2	Phospholipase A2	*Condylactis gigantea* (Giant Caribbean sea anemone)
adi_v1.16374	Disrupts haemostasis	8.0E-32	Q58L94	Prothrombin activator (Notecarin)	*Notechis scutatus scutatus* (Tiger snake)
adi_v1.15821	Disrupts haemostasis	5.0E-41	Q4QXT9	Coagulation factor X	*Tropidechis carinatus* (Rough scaled snake)
adi_v1.08904	Disrupts haemostasis	5.0E-35	Q76B45	Serine-type endopeptidase (Blarina toxin)	*Blarina brevicauda* (Northern short-tailed shrew)
adi_v1.19445	Disrupts haemostasis	8.0E-112	C6JUN2	Metalloprotease (SVMP)	*Philodryas olfersii* (Green snake)
adi_v1.14946	Disrupts haemostasis	1.0E-30	B6EWW8	Snake venom 5'-nucleotidase	*Gloydius brevicaudus* (Korean slamosa snake)
adi_v1.12298	Disrupts	4.0E-15	Q4PRC6	Snaclec 7	*Daboia siamensis* (Eastern Russel's viper)
adi_v1.18989	Disrupts haemostasis	6.00E-08	Q66S13	Fish venom Natterin-4	*Thalassophryne nattereri* (Niquim)
adi_v1.19802	Disrupts haemostasis	1.0E-19	A7X3Z0	C-type lectin (Lectoxin-Thr1)	*Thrasops jacksonii* (Jackson's black tree snake)
adi_v1.06850	Disrupts	9.0E-16	A8QZJ5	Cytotoxin-1	*Millepora dichotoma* (Net fire coral)
adi_v1.22096	Disrupts	2.00E-68	D8VNT0	Ryncolin-4	*Cerberus rynchops* (Dog-faced water snake)
adi_v1.19322	Induces	4.0E-19	P35779	Venom allergen 3	*Solenopsis richteri* (Black imported fire ant)
adi_v1.01092	Induces immune	5.0E-36	J3SFJ3	Translationally-controlled tumor	*Crotalus adamanteus* (Eastern diamondback rattlesnake)

	response			protein homolog	
adi_v1.11797	Neurotoxin	3.0E-30	P00615	Phospholipase A2	*Oxyuranus scutellatus scutellatus* (Australian taipan)
adi_v1.18628	Neurotoxin	8.0E-15	Q9TWL9	Phospholipase A2 (Conodipine-M)	*Conus magus* (Magician's cone snail)
adi_v1.05505	Peptidase	1.0E-41	C9WMM5	Serine carboxypeptidase	*Apis mellifera* (Honeybee)
adi_v1.05180	Peptidase	7.0E-52	B2D0J5	Serine carboxypeptidase	*Apis mellifera* (Honeybee)
adi_v1.11218	Proteinase inhibitor	3.0E-17	Q6ITB9	Kunitz-type serine protease inhibitor (Mulgin-3)	*Pseudechis australis* (King brown snake)
adi_v1.23374	Proteinase inhibitor	8.0E-17	P0C8W3	Kunitz-type serine protease inhibitor (Hg1)	*Hadrurus gertschi* (Desert hairy scorpion)
adi_v1.09855	Unknown	3.0E-136	B2DCR8	SE-cephalotoxin	*Sepia esculenta* (Golden cuttlefish)
adi_v1.20368	Venom maturation	2.0E-51	K7Z9Q9	Metalloprotease (Nematocyte expressed protein 6)	*Nematostella vectensis* (Starlet sea anemone)
adi_v1.17845	Venom maturation	3.0E-44	K7Z9Q9	Metalloprotease (Nematocyte expressed protein 6)	*Nematostella vectensis* (Starlet sea anemone)
adi_v1.13648	Venom maturation	4.0E-38	K7Z9Q9	Metalloprotease (Nematocyte expressed protein 6)	*Nematostella vectensis* (Starlet sea anemone)
adi_v1.07444	Venom maturation	1.0E-28	C9D7R3	Metalloprotease (Astacin-like toxin)	*Loxosceles intermedia* (Brown spider)

adi_v1.23604	Venom maturation	3.0E-21	C9D7R2	Metalloprotease (Astacin-like toxin)	*Loxosceles intermedia* (Brown spider)
adi_v1.15751	Venom maturation	7.00E-37	A0FKN6	Metalloprotease (Astacin-like toxin)	*Loxosceles intermedia* (Brown spider)
adi_v1.20292	Venom maturation	5.0E-28	B1A4F7	Venom dipeptidyl peptidase 4	*Vespula vulgaris* (Yellow jacket wasp)
adi_v1.15074	Venom maturation	2.0E-28	B1A4F7	Venom dipeptidyl peptidase 4	*Vespula vulgaris* (Yellow jacket wasp)
adi_v1.09733	Venom maturation	9.0E-29	Q9TXD8	Serine type endopeptidase	*Agelenopsis aperta* (North American funnel-web spider)
adi_v1.03437	Venom	6.0E-13	Q3SB11	Calglandulin	*Tropidechis carinatus* (Australian rough-scaled snake)
adi_v1.01102	Venom	5.0E-38	Q3SB11	Calglandulin	*Tropidechis carinatus* (Australian rough-scaled snake)
adi_v1.05162	Venom maturation	1.0E-47	J3S9D9	Phospholipase A2 activation (Reticulocalbin-2)	*Crotalus adamanteus* (Eastern diamondback rattlesnake)
adi_v1.10508	Venom maturation	5.0E-53	Q8MMH3	Protein-glutamate O-methyltransferase	*Pimpla hypochondriaca* (Parasitoid wasp)

CAPÍTULO 5

Considerações Finais

Embora haja avanços no conhecimento, permanecemos ignorantes sobre muitas das facetas da história natural dos venenos e das interações entre esta história natural e processos e padrões de evolução do conteúdo desses arsenais químicos. A evolução molecular das distintas famílias de toxinas no veneno, e seus mecanismos subjacentes, também são pouco compreendidos (Casewell *et al*., 2013; Sunagar *et al*., 2015). Constatando-se esse desconhecimento em grupos amplamente estudados, pode-se imaginar as deficiências de conhecimento em relação a um grupo como os cnidários, em que esforços vêm recém tomando corpo. A identificação de peptídeos tóxicos específicos em cnidários está limitada a um pequeno número de toxinas e táxons (principalmente de anêmonas-do-mar), identificados por métodos analíticos tradicionais de proteínas, com uma amostragem sub-representada (Weston *et al*., 2013; Li *et al*., 2014; Rachamim *et al*., 2014; Brinkman *et al*., 2015; Gacesa *et al*., 2015; Jouiaei *et al*., 2015a, 2015b, 2015c).

Neste estudo descrevemos três novos proteomas de dois cubozoários e um cifozoário. Estes dados, somados aos 5 proteomas anteriormente descritos, permitiram comparações mais abrangentes da composição dos venenos dos diferentes grupos. Identificamos representantes de todas as superfamílias mais importantes de proteínas tóxicas, o que reflete a grande convergência das famílias de proteínas entre todos as linhagens venenosas, ao mesmo que demonstra a complexidade do veneno de Anthozoa e Medusozoa. Das 28 famílias de proteínas identificadas neste trabalho, proteínas similares a 13 famílias foram registradas pela primeira vez no proteoma de Cnidaria. Paralelamente, estes dados permitiram elaborar a mais completa hipótese sobre o desenvolvimento evolutivo do arsenal bioquímico em cnidários, em que vislumbra-se novas perspectivas sobre a origem do veneno no grupo. Há ao menos 15 famílias de toxinas recrutadas no proteoma de veneno de cnidários antes da diversificação dos grupos Anthozoa e Medusozoa, sugerindo que os ancestrais dos cnidários já possuíam intensa atividade tóxica, em um padrão similar aos venenos atuais.

Outras conclusões importantes deste estudo estão relacionadas à evolução molecular dos proteomas, demonstrando-se que famílias de toxinas nos cnidários se diversificam amplamente mediante à duplicação de genes, um mecanismo amplamente distribuído neste tipo de proteínas (Wong & Belov, 2012), mas não exclusivamente por esse mecanismo. A evolução do veneno de Cnidaria também contrasta com a evolução rápida e a seleção positiva das famílias de toxinas do veneno na maioria dos linhagens animais conhecidas até o momento (Kordis & Gubensek, 2000; Fry *et al*., 2009; Casewell *et al*., 2013). De fato, a família de toxinas específica de cnidários (jellyfish toxin, JFT) mostra um padrão diferente, com episódios de evolução diversificada que afeta poucos ramos e sítios individuais desse grupo de toxinas, o qual parece ser o padrão em linhagens mais antigas de animais venenosos (Jouiaei *et al*., 2015b; Sunagar & Moran, 2015).

Caracterizações genômicas, transcriptômicas e proteômicas de veneno, coletivamente conhecidas como "venômica", têm potencial para elucidar os mecanismos moleculares que operam na evolução das famílias de genes presentes no veneno, sobretudo nos elementos que controlam sua regulação e expressão. Além disso, esta maior disponibilidade de métodos permitirá que a investigação venômica no futuro próximo proporcione novas dimensões de compreensão para a evolução e ecologia dos venenos.

REFERÊNCIAS

Brinkman, D.L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D. & Mulvenna, J. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16: 1–15.

Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A. & Fry, B.G. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends in Ecology and Evolution 28: 219–229.

Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.; Lewis, R.J.; Norton, R.S.; Renjifo, C. & de la Vega, R.C.R. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10: 483–511.

Gacesa, R.; Chung, R.; Dunn, S.R.; Weston, A.; Jaimes-Becerra, A.; Marques, A.C.; Morandini, A.; Hranueli, D.; Starcevic, A.; Ward, M. & Long, P.F. 2015. Gene duplications are extensive and contribute significantly to the toxic proteome of

nematocysts isolated from Acropora digitifera (Cnidaria: Anthozoa: Scleractinia). BMC Genomics 16:774.

Jouiaei, M.; Casewell, N.; Yanagihara, A.; Nouwens, A.; Cribb, B.; Whitehead, D.; Jackson, T.; Ali, S.; Wagstaff, S.; Koludarov, I.; Alewood, P.; Hansen, J. & Fry, B. 2015a. Firing the sting: Chemically induced discharge of cnidae reveals novel proteins and peptides from box jellyfish (*Chironex fleckeri*) venom. Toxins 7: 936–950.

Jouiaei, M.; Sunagar, K.; Gross, A.F.; Scheib, H.; Alewood, P.F.; Moran, Y. & Fry, B.G. 2015b. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Molecular Biology and Evolution 32: 1598–1610.

Jouiaei, M.; Yanagihara, A.; Madio, B.; Nevalainen, T.; Alewood, P. & Fry, B. 2015c. Ancient venom systems: A review on Cnidaria toxins. Toxins 7: 2251–2271.

Kordis, D. & Gubensek, F. 2000. Adaptive evolution of animal toxin multigene families. Gene 261:43-52.

Li, R.; Yu, H.; Xue, W.; Yue, Y.; Liu, S.; Xing, R. & Li, P. 2014. Jellyfish venomics and venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. Journal of Proteomics 106: 17–29.

Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T. & Sher, D. 2014. The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. Molecular Biology and Evolution 32: 740–753.

Sunagar, K.; Morgenstern, D.; Reitzel, A.M. & Moran, Y. (in press). 2015. Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. Journal of proteomics.

Sunagar, K & Moran, Y. 2015. The Rise and Fall of an Evolutionary Innovation: Contrasting Strategies of Venom Evolution in Ancient and Young Animals. PLoS Genetics 11(10): e1005596.

Weston, A.J.; Chung, R.; Dunlap, W.C.; Morandini, A.C.; Marques, A.C.; Moura-da-Silva, A.M.; Ward, M.; Padilla, G.; da Silva, L.F.; Andreakis, N. & Long, P.F. 2013. Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish *Olindias sambaquiensis*. Toxicon 71: 11–17.

Wong, E.S.W. & Belov, K. 2012. Venom evolution through gene duplications. Gene 496: 1–7.

Resumo

A evolução do veneno, uma das misturas mais complexas da natureza, tem sustentado o sucesso da diversificação de inúmeras linhagens de animais. Serpentes deslizantes ou medusas flutuantes utilizam o veneno, um coquetel de peptídeos farmacologicamente ativos, sais e moléculas orgânicas. Esses animais surpreendentes têm provocado grande fascínio ao longo da história humana. Nesta dissertação propomos um estudo da evolução dos venenos no filo Cnidaria, englobando dados proteômicos e genômicos. Este projeto teve como objetivos: (1) caracterizar e elucidar a evolução da composição do veneno em Cnidaria por meio da comparação de listas de proteínas; (2) testar a hipótese de que a variação na família de toxinas específica de cnidários tem sido o resultado de um regime de seleção positiva; e (3) determinar a extensão em que a duplicação de genes pode ser considerada como a principal razão para a diversificação de toxinas em Cnidaria.

O capítulo "Comparative proteomics reveals common components of a powerful arsenal in the earliest animal venomous lineage, the cnidarians" propõe o estudo comparado mais completo sobre a composição do veneno de cnidários e uma hipótese sobre a montagem evolutiva do complexo arsenal bioquímico de cnidários e do veneno ancestral desse grupo basal. Vinte e oito famílias de proteínas foram identificadas. Destas, 13 famílias foram registradas pela primeira vez no proteoma de Cnidaria. Pelo menos 15 famílias de toxinas foram recrutadas no proteoma de veneno de cnidários antes da diversificação dos grupos Anthozoa e Medusozoa.

Nos capítulos "Evidence of episodic positive selection in the evolution of jellyfish toxins of the cnidarian venom" e "Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia)", nossas análises demonstram que as famílias de toxinas nos cnidários se diversificam amplamente mediante a duplicação de genes. Além disso, em contraste com as famílias de toxinas do veneno na maioria das linhagens animais; nós identificamos um padrão diferente na família de toxinas específica de cnidários, em que há uma seleção purificadora por longos períodos seguindo longos tempos de diversificação ou vice-versa.

Abstract

The evolution of venom, nature's most complex concoction, has underpinned the success and diversification of numerous animal lineages. Slithering serpents or buoyant jellyfishes employ venom, a cocktail of pharmacologically active peptides, salts, and organic molecules. These astonishing animals have generated a great fascination throughout human history. In this dissertation, we propose a study of the evolution of venoms in the phylum Cnidaria, encompassing proteomic and genomic data. This project aimed: (1) to characterize and elucidate the evolution of venom composition in Cnidaria by comparing protein lists; (2) to test the hypothesis that the variation in specific family of cnidarians toxins has been the result of a positive selection regime; and (3) to determine the extent to which the genes duplication may be regarded as the main reason for the diversity of toxins in Cnidaria.

The chapter "Comparative proteomics reveals common components of a powerful arsenal in the earliest animal venomous lineage, the cnidarians" presents the most comprehensive comparative study on the cnidarians venom composition and a hypothesis about the evolutionary assembly of the complex biochemical arsenal of cnidarians and of the ancestral venom of this basal group. Twenty eight protein families were identified. Of these, 13 families were described for the first time in the proteome of Cnidaria. At least 15 types of toxin families were recruited in cnidarians venom proteome before the diversification of Anthozoa and Medusozoa groups.

In the chapters "Evidence of episodic positive selection in the evolution of jellyfish toxins of the cnidarian venom" and "Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia)", our analyses indicate that the families of toxins in cnidarians diversify broadly through gene duplication. Besides, in contrast to the families of venom toxins in most animals lineages, we identified a different pattern in the specific family of cnidarians toxins, where there is a purifying selection for periods long, followed by long periods of diversification or vice versa.