Review

Venom systems as models for studying the origin and regulation of evolutionary novelties

Giulia Zancolli[1,2,*] and Nicholas R. Casewell[3]

[1]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; giulia.zancolli@gmail.com

[2]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[3]Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, United Kingdom; Nicholas.Casewell@lstmed.ac.uk

[*]Correspondence: giulia.zancolli@gmail.com

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Abstract

A central goal in biology is to determine the ways in which evolution repeats itself. One of the most remarkable examples in nature of convergent evolutionary novelty is animal venom. Across diverse animal phyla, various specialized organs and anatomical structures have evolved from disparate developmental tissues to perform the same function, i.e. produce and deliver a cocktail of potent molecules to subdue prey or predators. Venomous organisms therefore offer unique opportunities to investigate the evolutionary processes of convergence of key adaptive traits, and the molecular mechanisms underlying the emergence of novel genes, cells, and tissues. Indeed, some venomous species have already proven to be highly amenable as models for developmental studies, and recent work with venom gland organoids provides manipulatable systems for directly testing important evolutionary questions. Here, we provide a synthesis of the current knowledge that could serve as a starting point for the establishment of venom systems as new models for evolutionary and molecular biology. In particular, we highlight the potential of various venomous species for the study of cell differentiation and cell identity, and the regulatory dynamics of rapidly-evolving, highly expressed, tissue-specific, gene paralogs. We hope that this review will encourage researchers to look beyond traditional study organisms and consider venom systems as useful tools to explore evolutionary novelties.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Introduction

Across the tree of life, species are often found to evolve analogous novel traits when independently presented with similar environmental or biological challenges, a process known as convergent evolution (Stern 2013). One of nature's most remarkable examples of such repeated evolution of a key adaptive trait is animal venom. On more than one hundred occasions, animal lineages, including extinct dinosaurs (Gong et al. 2010) and mammals (Fox and Scott 2005), have independently evolved the ability to produce and deliver cocktails of bioactive molecules into other animals (Schendel et al. 2019). These biochemical weapons have typically evolved as an adaptation for foraging, defence or intraspecific competition, and their evolutionary acquisition remodels the predator-prey interaction from a physical to a biochemical battle, often enabling small venomous organisms to subdue or escape from much larger animals (Casewell et al. 2013). There are more than 200,000 venomous animal species known to science and, unsurprisingly, they have been the subject of public fascination throughout human history – often small and fragile-looking animals are capable of injecting complex secretions that can result in devastating damage, including death (Figure 1).

Acquiring a venom system is not a simple process, and it involves several steps including: i) turning normal, physiological proteins into toxic molecules and, at the same time, developing mechanisms of auto-resistance against them; ii) evolving novel tissues to secrete and transport venom such as exocrine glands, ducts, and muscular bulbs; and iii) developing anatomical structures for efficient venom delivery, such as fangs, stingers, harpoons, forcipules, barbs among others. Despite often having the same overarching function (e.g. prey capture or predator deterrence), the anatomy and organization of venom systems differ dramatically between lineages, providing insights into the constraints (e.g. morphology, ecology, genetics) that shape organismal complexity (Schendel et al. 2019). Venom systems therefore represent exceptional opportunities to investigate the mechanisms underlying the evolution and development of novel proteins, tissues and morphological structures, and ultimately to understand the basis of how similar solutions are used repeatedly in nature - in turn contributing to a more general and predictive formulation of evolutionary theories.

Venom systems have already proven to be unique models for the study of regulatory evolution of novel cells. For instance, in the last 15 years, cnidarians (mainly *Nematostella vectensis* and *Hydra magnipapillata*) have been established as laboratory models for this emerging field of research (Darling et al. 2005; Galliot 2012; Babonis and Martindale 2014; Layden et al. 2016; Sachkova and Burkhardt 2019), providing important insights into the relative contribution of conserved and derived genes in regulatory networks (GRNs), and the role of various

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

transcription factors and signalling pathways in cell identity (see below). While Cnidaria are amenable models for a detailed resolution at the cellular level, other venomous animals offer the opportunity to study the molecular basis of the repeated evolution of novel anatomical structures and organs. However, only a small number of studies, mainly focused on snakes, have investigated and identified regulatory elements and their potential roles within the venom production cycle (e.g. Ma et al. 2001; Nakamura et al. 2014; Han et al. 2016; Schield et al. 2019) (Figure 2). This is surprising given the extensive research effort that has focused on characterizing the molecular and biochemical composition of animal venom systems. While we generally lack a comprehensive understanding of the precise cellular and regulatory mechanisms responsible for the origin, development and functioning of venom systems, work to this point provides a promising foundation to build a unified understanding of the processes underlying the evolution of these key evolutionary innovations. Furthermore, the recent development of new models, such as venom gland organoids (Post et al. 2020), enable in depth investigations of the regulatory mechanisms governing venom production in homologous study systems.

Animal toxins are also remarkable models for studying evolutionary processes relating to the acquisition of novel gene function or protein neofunctionalization, and the regulatory dynamics of convergently recruited proteins. The classic model for the origin of venom toxins involves the duplication of genes encoding for physiological body proteins, often followed by neofunctionalization (Kordiš and Gubenšek 2000; Casewell et al. 2011; Chang and Duda 2012), additional duplications, and selective increased expression in the venom gland (Whittington et al. 2008; Reyes-Velasco et al. 2014; Junquiera-de-Azevedo et al. 2015; Margres et al. 2017). However, other mechanisms, including the co-option of single-copy genes (Wong et al. 2012; Martinson et al. 2017; Drukewitz et al. 2019), horizontal gene transfer (Moran et al. 2012a; Martinson et al. 2016), alternative splicing (Cousin et al. 1998; Zeng et al. 2012), and exon shuffling (Wang et al. 2017) may also contribute to the suite of mechanisms that underpin the acquisition of venom genes. Several of the most abundant and potent toxins found in animal venoms belong to large, highly diverse protein families characterized by accelerated sequence evolution (Casewell et al. 2011; Sunagar et al. 2015). Indeed, certain toxin families represent some of the fastest diversifying proteins identified to date (Chang and Duda 2012). Ultimately, these variable processes can result in a single venom containing several hundred different bioactive molecules. Despite this extraordinary complexity and variability, there is a remarkable degree of convergence in both the physiological targets of toxins and the molecules selected for use as toxins. Indeed, of the plethora of available building

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

blocks, a relatively limited number of protein groups (e.g. cysteine rich secretory proteins, hyaluronidase, kunitz, phospholipase A_2 [PLA_2], serine proteases) have been convergently recruited as venom components across multiple metazoan lineages, suggesting that relatively few protein families are amenable for a toxic role (Fry et al. 2009; Casewell et al. 2013, 2019). The reasons for this remain unclear; the non-toxin 'physiological' roles of these protein types are diverse, and many of these toxin paralogs are themselves members of large multi-locus gene families with, in snakes at least, broad low-level expressional profiles across diverse tissue types (Reyes-Velasco et al. 2014; Junquiera-de-Azevedo et al. 2015). These features make venom toxins highly amenable to use as models for investigating the regulatory dynamics of novel, highly-expressed and tissue-specific genes that have evolved rapidly under strong natural selection from paralogous genes with distinct expression backgrounds.

Thus, venom systems provide unique opportunities to investigate the: i) evolutionary processes relating to the convergence of key adaptive traits; ii) molecular mechanisms underlying the emergence and development of novel cell types and organs, and iii) regulatory dynamics of novel functional genes. It is well known from developmental studies that GRNs control cell differentiation and might therefore be considered to define cell types (Davidson 2010). Understanding the evolution of novel cells requires first identifying the components of the GRN, because a well characterized GRN is a critical tool for testing hypotheses on the conservation of molecular mechanisms and regulatory relationships across different cell types and taxa (Babonis and Martindale 2014). The goal of this review is thus to provide a synthesis of the current knowledge that could serve as starting point for the establishment of the venom system as a novel model in evolutionary and molecular biology. First, we provide an overview of the different origins of venom-secreting cells and tissues; then we discuss the regulation of venom in a hierarchical fashion. More specifically, we summarize current knowledge on the regulatory mechanisms involved in: i) differentiation and identity of venom secreting cells; ii) overall functioning of the venom gland; iii) transcription of toxin genes, and iv) selective expression of venom components. While post-translational modifications contribute substantially to the diversity of toxins found in many venom systems (Buczek et al., 2005; Moura-da-Silva et al. 2016; Delgado-Prudencio et al. 2019) we choose not to discuss this in depth here, and instead focus on the processes directly affecting the regulation of venom-secreting cells and their gene expression profiles.

The origin of venom-secreting cells

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Cells that produce venom are all secretory cells of the epithelium. The most striking feature of these cells, either as a unicellular system or organized into tissues and glands, is the variety of ways in which they have evolved. Indeed, even if the venom apparatus appears outwardly similar in terms of structure and function between independent venomous lineages, the mode of development may vary. The question of great interest is whether the molecular basis of different venom systems are similar, i.e. whether distinct lineages adopted the same regulatory strategies for the evolution of the venom system. Here, we provide an overview of what is currently known about the origins of animal venom systems.

Cnidaria

Cnidarians (e.g. jellyfish, sea anemones, etc) represent the oldest extant group of venomous animals. They produce venom in specialized cells called cnidocytes, which contain miniature stinging devices, cnidocysts; additionally, toxins are also produced in ectodermal gland cells (Moran et al. 2012b). A close evolutionary relationship has been proposed between cnidocytes and the elaborate extrusive cells of some protists (Shostak 1993; Denker et al. 2008; Ozbek et al. 2009), although no molecular evidence of lateral gene transfer was found to support this hypothesis (Technau et al. 2005). Additionally, structural, functional and phylogenomic data support the independent origin of cnidocytes in cnidarians and the homologous harpoon-like secretory organelles in dinoflagellate protists (Gavelis et al. 2017). Because of the ability of cnidocytes to sense and respond to the surrounding environment, it was hypothesized that they are an unusual type of neuron (Pantin 1942). The expression of several common regulatory genes in the precursor of both neurons and cnidocytes (e.g. *cnox-2*, *prdl-b*, and *COUP-TF*) (Galliot and Quiquand 2011) indeed support the hypothesis of a common neuron-cnidocyte progenitor (Miljkovic-Licina et al. 2007). However, single-cell RNA-seq analysis of *Hydra* has shown that cnidocytes emerge from a common progenitor before this differentiates into neurons or gland cells (Siebert et al. 2019). Cnidocytes differentiate synchronously from the centralised population of progenitor cells, and then migrate to the tissue where they will be deployed (Khalturin et al. 2007). While in Hydrozoa cnidocytes originate from multipotent interstitial stem cells (David 2012), in other cnidarians, such as Anthozoa (e.g. *Nematostella*) and Scyphozoa (e.g. *Aurelia*), they originate from epitheliomuscular cells (Gold and Jacobs 2013; Sebé-Pedrós et al. 2018; Gold et al. 2019). In *Nematostella*, cnidocytes develop early, at the onset of gastrulation (Zenker et al. 2011), and throughout all life stages they develop individually and asynchronously in the epithelium (Babonis and Martindale 2014, 2017). An interesting research direction that would contribute to our understanding of cell-type

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

differentiation would be to investigate how interstitial cells came to regulate cnidogenesis in the most derived lineage of cnidarians, the Hydrozoa (Gold et al. 2019).

Snakes

In snakes, the venom glands and fangs arise from a joint invagination of the oral epithelium in the ocular region, separately from other supralabial glands (Kochva 1963; Shayer-Wollberg and Kochva 1967). Interestingly, despite this development from oral tissue and a potential origin of the venom gland from a labial gland, it was hypothesized by Kochva (1987) and further supported by Vonk et al. (2013), that a developmental link also exists between the venom gland and the pancreas. Some pancreatic enzymes (e.g. PLA_2) have counterparts that are detectable in the salivary and venom glands, the latter of which are utilized as toxins (Fry 2005). The observation of abundant venom gland expression of miR-375, which is a canonical microRNA in the vertebrate pancreas, also suggests that regulatory components of the venom secretory system may have a pancreatic origin (Vonk et al. 2013). These data hint that the evolution of the venom system in snakes may have involved the co-option and orchestration of components from multiple anatomical origins, although further research is required to fully elucidate this process.

Cone snails

Cone snails are marine predatory Neogastropods that possess a sophisticated venom apparatus consisting of: i) a long, convoluted duct where toxins are secreted, ii) a muscular bulb that contracts to expel the venom through the duct, and iii) a harpoon-like structure for the injection of venom into prey or predator. The unusual morphology of this venom gland has generated controversy about its evolutionary derivation which has been interpreted as either a highly derived salivary gland (Alpers 1931; Fretter and Graham 1962), a completely new structure (Smith 1967), or homologous to the mid-oesophageal glands of other gastropods (Ponder 1973; Ponder and Lindberg 1997; Modica and Holdford 2010). Support for the latter hypothesis has been provided from developmental studies showing that the venom gland arises during metamorphosis from hypertrophy of epithelial cells of the ventral zone within the mid-foregut region, followed by rapid pinching-off of the newly formed duct from the main dorsal channel (Ball 2002; Page 2011). Although the precise homology between the various mid-oesophageal structures is still controversial (Ponder et al. 2019) and currently under investigation, it is clear that the epithelial tubes and granules of the venom duct originate from a pre-existing epithelial sheet via remodelling (Page 2011).

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Other lineages

For most other animal lineages, data on the origin of their venom secreting cells remains scarce. In spiders, the venom system has evolved from small 'primitive' glands in the basal segment of the chelicera of the oldest taxon, the Mesothelae (Foelix and Erb 2010), to well-developed glands located in the prosoma of the advanced araneomorphs (true spiders) (Langenegger et al. 2019). Venom glands are already present in the first prelarval stage of *Phoneutria nigriventer* (Silva et al. 2011) and are derived from ectodermal invagination (Simpson and Casas 2011). Contrastingly, in centipedes, several recent studies have suggested that the venom glands originated as patches of glandular epidermal epithelium and its adjoining cuticle, which became increasingly invaginated into the interior of the forcipules (summarized in von Reumont et al. 2014). Among vertebrates, Siluriformes (catfishes) possess relatively simple venom glands composed of aggregations of glandular cells enclosed within an integumentary sheath of epithelial cells and associated with bony spines in dorsal and pectoral fins (Wright 2017). Multiple evidences suggest an origin from epidermal secretory cells, possibly from thickening of glandular tissue surrounding the spines followed by subsequent evolutionary changes (summarized in Wright 2017). However, whether the venom glands derive from club cells secreting defensive crinotoxins (Cameron and Endean 1973) or cells producing healing and antimicrobial secretions (Wright 2009) is still debated and further research is needed. The venom glands of the platypus, one of very few venomous mammals, are derived from modified apocrine sweat glands, which migrate during development from the inner surface of each thigh to their final position on the dorsocaudal surface of the pelvis (Krause 2010). The earliest stages of gland ontogeny examined are from nestling and recently emerged platypus juveniles (six months old).

Regulation of venom-secreting cells and tissues

To the best of our knowledge, the GRN of venom-secreting cells have been poorly explored; however, a relatively small number of studies performed over the past twenty years have started to shed some light on the molecular basis of the venom system (Figure 2). The regulatory dynamics of cell identity and differentiation has only been studied in Cnidaria; by contrast, DNA analysis of toxin genes and the potential elements regulating their expression have been mainly performed on snakes and a few other organisms. Below, we present the main findings and discuss the potential of these results as candidates to test in future studies by means of advanced molecular techniques and/or comparative genomic and transcriptomic approaches.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Regulatory elements involved in the differentiation of venom-secreting cells

The processes upstream of the formation of venom-secreting cells have only been studied in cnidarians. Research on this topic has benefited from a long history of pioneering studies (summarized in Babonis and Martindale 2014), and the relative ease of applying genetic manipulations that are challenging to perform in other venomous organisms.

Transgenic reporter lines and knockdown experiments were used to identify various transcription factors crucial for cnidocyte differentiation from multipotent cells. SoxB2 and SoxB3, orthologs of bilaterian SoxB genes, were found to regulate the differentiation of the progenitor cells into neurons and cnidocytes in two distantly related cnidarian species, the anthozoan *Nematostella vectensis* (Richards and Rentzsch 2014) and the hydrozoan *Hydractinia echinata* (Flici et al. 2017). Indeed, the Sox family of conserved transcription factors is known to be involved in the development and maintenance of cell identity (Bowles et al. 2000). Downstream of SoxB2, PaxA plays a critical role in the development of cnidocytes in *Nematostella* (Babonis and Martindale 2017). Paired box (Pax) genes are a family of conserved transcription factors involved in the development of sensory structures in diverse metazoan tissues (Blake and Ziman 2014), and cnidarians possess orthologs of Pax genes as well as lineage-specific ones, namely PaxA, PaxC and PaxD (Matus et al. 2007; Suga et al. 2010). Knockdown of PaxA, but not PaxC, results in a loss of developing cnidocytes, and causes downregulation of myocyte enhancer factor 2 (Mefv2), another transcription factor found to be important for the normal expression of cnidocyst-specific genes (Genikhovich and Technau 2011). Recent single-cell transcriptomic analyses of *Nematostella* confirmed these previous findings, with SoxB2 expressed in the early stage of cnidocyte differentiation, followed by the expression of PaxA (Sebé-Pedrós et al. 2018). A similar study in *Hydra* (Siebert et al. 2019) found Pax binding motifs significantly enriched in the regulatory regions of a set of genes ("metagene") that are expressed during early and mid-cnidogenesis; furthermore, PaxA itself was expressed, confirming its role in cnidogenesis also in *Hydra*, and suggesting some degree of generality. Other motifs were also found enriched during cnidogenesis, including forkhead (Fox) which was enriched at mid- and late stages, and Pou at the late stage. Similarly, and also in *Nematostella*, FoxL2 and Pou4 were expressed at mid- and late stages (Sebé-Pedrós et al. 2018).

While it is less challenging to study the regulatory elements involved in cell differentiation relating to unicellular venom systems, particularly since these cells regenerate throughout the lifetime of cnidarians, it is far more problematic to perform such research for centralized,

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

tissue-based systems like venom glands. However, with ever increasing advances in DNA sequencing and genomic engineering (e.g. CRISPR/Cas9), coupled with recently-described, amenable and manipulatable model systems such as venom gland-derived organoids (Post et al. 2020), we foresee the possibility to test whether elements involved in cell differentiation are shared across distinct venomous lineages, while simultaneously establishing representatives of other venomous lineages as models for developmental biology studies.

Regulatory elements expressed in venom-secreting cells and tissues

While SoxB2 and PaxA are involved early on in the development of cnidocytes associate with cnidarian venom, a small number of other regulatory elements have been found implicated downstream of this cell identity pathway. For example, by means of CRISPR/Cas9 and FACS sorting followed by RNA-seq, Sunagar et al. (2018) found transcription factors differentially expressed in cnidocytes. Two genes paralogous to proto-oncogenic transcription factors c-Jun and c-Fos, named by the authors cnido-Jun and cnido-Fos1, were strongly upregulated, and knockdown experiments of cnido-Jun resulted in reduced development and impairment of cnidocytes, which displayed atypical shape. c-Jun and c-Fos are known to form the activation protein 1 (AP1) complex which mediates gene regulation in response to stress, infection and other stimuli (Hess et al. 2004), and c-Jun in particular has previously been found involved in various signalling pathways including tissue development (Meng and Xia 2011). Phylogenetic analysis showed that these two transcription factors originated by gene duplication around 500 million years ago in the common ancestor of hexacorallians. The paralogous c-Jun and c-Fos were not upregulated in cnidocytes, therefore it appears likely that cnido-Jun and cnido-Fos1 evolved a different function to that related to stress-response. Furthermore, a Jun factor was also found expressed in the late differentiation stages of cnidocyte development in *Nematostella*, and motif-enrichment analysis on potential enhancer elements found Jun associated with a cnidocyte-specific set of genes (Sebé-Pedrós et al. 2018).

Intriguingly, AP1 and another transcription factor, NFkB, have been shown to increase their activity in the venom glands of the snake *Bothrops jararaca* after venom gland stimulation (Luna et al. 2009). In other systems, NFkB is present in cells in an inactivated state, enabling a rapid-acting response to stimuli (Gilmore 2006); thus, it appears that stress-related transcription factors are also part of the GRN of venom-secreting cells in snakes. AP1 and NFkB are members of the extracellular signal-related kinases 1 and 2 (ERK 1/2) signalling pathway, and other members of this pathway were found involved in toxin gene regulation. Several full-length sequences encoding proteins with homology to epithelium-specific

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

transcription factors (ETS) were identified in a cDNA library of the venom gland of the habu snake, *Protobothrops flavoviridis* (Nakamura et al. 2014). ETS are also activated by ERK, and neighbouring ETS and AP1 binding sites can act as response elements for transcriptional activation by this pathway (Hollenhorst 2012). Within the ETS family, proteins that are specifically expressed in epithelial-rich tissues are classified into the ESE subfamily. Among these, ESE-1 is broadly expressed in multiple tissues, while ESE-2 and ESE-3 are expressed in tissues enriched with glandular epithelia, such as salivary and mammary glands (Kas et al. 2000). All three ESEs were found expressed in the *P. flavoviridis* venom gland, with ESE-3 showing higher expression levels and specificity for the venom gland compared to other tissues. ESE-3 was also the only factor able to activate the promoter of PLA_2 toxin encoding genes (Nakamura et al. 2014). More recently, Schield and colleagues (2019) reported a set of 12 transcription factors significantly upregulated in the venom gland of the prairie rattlesnake, *Crotalus viridis*, compared to other body tissues. Of these 12 transcription factors, grainyhead like transcription factor 1 (GRHL1), which is known to function in epidermal barrier formation and repair (Ting et al. 2005; Kim and McGinnis 2011), is also directly regulated by ERK. Outside the ERK pathway, they detected transcription factors linked to the repair of the glandular epithelium (e.g. ELF5) and secretory functions (e.g. ATF6 and CREB3L2). Additionally, transcript sequences belonging to the nuclear factor I (NFI) family were also found to be upregulated; NFI members are known to drive tissue-specific expression (Gronostajski 2000) and function in chromatin remodelling and transactivation (Fane et al. 2017). While investigations of the transcription factors expressed in venom glands can provide us with exploratory evidence of which regulatory elements are expressed and thus might play important regulatory roles, they do not provide direct evidence of the interactions with DNA. Advanced genomics technologies such as chromatin immunoprecipitation sequencing (ChIP-seq), which directly analyses proteins-DNA interactions, and assay for transposase-accessible chromatin sequencing (ATAC-seq), which assesses genome-wide chromatin accessibility, could be readily implemented to identify the transcription factor binding sites of toxin genes and their associated proteins (Geertz and Maerkl 2010).

Transcriptional regulation of toxin genes

While differential expression analysis of transcription factors provides insights into the GRN of venom-secreting cells, it does not reveal which elements directly control for the expression of their main products, i.e. toxins. Studies performing comparative DNA analyses of gene structure, reporter assays, motif binding testing, among others, have started to shed some light

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

on the processes controlling for individual expression of venom genes. A full list of the transcription binding sites collected from the literature is provided in Table S1 and S2.

Cis-regulatory elements in promoter regions

A good starting point for gaining insights into transcriptional regulation is to analyse the region ~1000 bp upstream of the transcription start site of toxin genes because it contains specific DNA sequences ("*cis*-regulatory elements") that provide binding sites for transcription factors, enhancers, silencers and other elements that up- or down-regulate gene expression. Studies in *Nematostella* have indeed demonstrated that the promoter region and first part of the genes encoding toxins contain all of the necessary elements, not only for expression, but also for tissue-specificity (Colombus-Shenkar et al. 2018; Sachkova et al. 2019). Unfortunately, the transcription factors directly controlling for toxin gene expression were not investigated in these studies; nonetheless, the large amount of genomic data currently available offers the opportunity to identify such motifs and binding sites for further exploration of this topic.

DNA sequence analysis of promoter regions have been mainly performed in snakes. A study on a cardiotoxin gene belonging to the three-finger toxin (3FTX) family in the spitting cobra *Naja sputatrix* revealed the presence of two putative glucocorticoid receptor binding sites, and chloramphenicol acetyltransferase (CAT) reporter assay experiments showed an inhibitory effect on gene expression when glucocorticoid receptors were present (Ma et al. 2001). Activated glucocorticoid receptors can complex with NFkB or AP1 - both found in venom gland nuclear extracts (Luna et al. 2009) - and prevent them from binding their target genes, hence indirectly repressing their expression (Ray and Prefontaine 1994). The glucocorticoid receptor has also been shown to be important in salivary gland development and secretion (Jaskoll et al. 1994), and multiple binding sites have previously been found in the promoter region of the salivary-gland specific protein cystatin S (Shaw and Chaparro 1999). Therefore, glucocorticoid receptors are good candidates for playing a role in the higher machinery regulating toxin transcription; however, further analysis on the presence of glucocorticoid receptors in venom gland cells is required to understand the potential generality of the role these play in different venom systems.

Insertions in promotor regions have also been implicated in influencing the expression of venom genes. For example, nucleotide sequence comparison of Trocarin D (TroD), a venom prothrombin activator (vPA) from *Tropidechis carinatus*, and its paralogous physiological counterpart, the blood coagulation factor X (TrFX), reveals the presence of an insertion in the promoter region of TroD, termed *VERSE*. A similar insertion is also found in the promoter of

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

the catalytic subunit of Pseutarin C (PCCS), another vPA from *Pseudonaja textilis*, although these two insertions are thought to originate from independent events (Reza et al. 2007). Such insertions have been postulated to be important elements in the recruitment process responsible for the switch from low level expression of factor X in the liver, to high expression of TroD and PCCS in the venom gland (Reza et al. 2007, Kwong et al. 2009). In TroD, *VERSE* harbours *cis*-elements that upregulate gene expression by 19- to 49-fold but do not control for venom-gland specific expression (Kwong et al. 2009). Analogously, an insertion was identified in the promoter region of group IA PLA_2 toxin genes in the sea snake *Laticauda semifasciata* (Fujimi et al. 2002a, 2004; Han et al. 2016). Comparative sequence analysis and reporter gene assays showed that this insertion significantly increased gene expression compared to the related group IB PLA_2 gene where the promoter insertion was absent. Group IA PLA_2 are highly expressed in the venom of this snake, whereas group IB PLA_2 – which contain a pancreatic loop in their amino acid sequences and are more similar to the digestive PLA_2 secreted in the pancreas of mammals – are generally lowly expressed and undetectable in the venom. The promoter insertion in group IA PLA_2 contains two E boxes and one GC box, and is present in other group IA PLA_2 venom genes from several snake species, suggesting that this sequence likely pre-dates the elapid snake radiation (Fujimi et al. 2002b, 2004). Insertions in promoter regions therefore seem to play an important role in increasing the expression levels of toxin genes, particularly compared to their physiological, non-toxin counterparts. However, these experiments were all performed in mammalian cell cultures, hence the specific function of these insertions and the role of the relevant binding transcription factors remain to be systematically tested in homologous systems.

Early experiments in *Naja sputatrix* revealed that the transcription binding sites of a 3FTX cardiotoxin were largely distinct to those found in the promoter region of PLA_2 genes (Jeyaseelan et al. 2001; Ma et al. 2001), though some were present in both toxin-encoding gene families, such as SP-1 with activator/enhancer activities. Transcription binding sites were also found to be conserved within members of the same toxin gene family, but not between different toxin families, in *L. semifasciata* (Fujimi et al. 2002a, 2004) and *T. carinatus* (Han et al. 2016). Since venom gene paralogs originate from duplication events, we can thus infer that each new copy likely inherited the promoter region, and hence the regulatory machinery, from the paralog from which it was derived, and this would explain the similarity in DNA sequences in the promoter regions within but not between gene families. Indeed, such venom toxin paralogs are often found arrayed in tandem in snake genomes, but with different toxin family gene clusters placed at distinct chromosomal locations (Vonk et al. 2013; Shibata et al. 2018; Schield

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

et al. 2019; Suryamohan et al. 2020). Subsequent paralog-specific mutations in the promoter region of individual genes might be at least partially responsible for differential expression within any particular gene family, and thus likely contribute towards the 'fine tuning' of gene expression. Since some of the toxin families discussed here are found in numerous animal venom systems (e.g. PLA_2 toxins are found in reptiles, arachnids, cephalopods, cnidarians, hymenopterans, among others) (Fry et al. 2009), it would be fascinating to compare the promoter regions across these diverse animal lineages to explore whether the regulatory elements associated with the convergent molecular evolution of these toxins have also evolved in a repeatable manner.

Promoter regions are also known to interact with elements located at long distances from target genes (Schoenfelder and Fraiser 2019). These long-range enhancers, silencers and insulators are found in close spatial proximity to the promoters of their target genes, and deletion experiments have shown that distal enhancers are essential for the regulation of spatiotemporal expression of target genes (e.g. Sagai et al. 2005). Recently developed sequencing techniques such as Promoter Capture Hi-C (PCHi-C) enable the genome-wide detection of these distal promoter-interacting regions for all promoters in a single experiment (Schoenfelder et al. 2018). The application of PCHi-C to venom-secreting cells can therefore facilitate the pull-down of promoter sequences and the identification of their frequent long-range interaction partners. Subsequent validation of the identified promotors via genome manipulation experiments in homologous systems, such as venom gland organoids (Post et al. 2020), would in turn enable robust assessments of promoter function, while comparative analyses of these regions across different venomous lineages can provide extensive insights into the large-scale regulatory architecture of this key evolutionary trait.

Cis-regulatory elements in introns

Besides promoter regions, introns also play functional roles, with the first mostly involved in transcriptional regulation, while downstream introns control gene splicing (Storbeck et al. 1998; Levy et al. 2001; Majewski and Ott 2002). In the Chinese scorpion (*Mesobuthus martensii*), genes encoding for neurotoxins have a conserved structure with two exons and only one intron which may control both splicing and transcription levels (Cao et al. 2013). Indeed, when replacing the intron of a potassium-channel toxin gene with that of another toxin in cultured human cells, the expression of the artificial construct increased (Zhijian et al. 2006). A similar study in cone snails showed that *E. coli* competent cells transfected with a vector containing an intron-free A-conotoxin construct had lower expression levels than the vector

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

containing the intron (Wu et al. 2013). Gene analysis of α-latrotoxin and its related paralog in the black widow (*Latrodectus hesperus*) revealed several transcription binding sites in the introns, although those elements seemed unrelated to venom production (Bhere et al. 2014). Studies on snakes went one step further and identified potential regulatory elements within introns of venom genes. For instance, luciferase assays were employed to show that venom gland-specific expression of the TroD toxin is regulated by elements located within three insertions of the first intron (Fujima et al. 2004; Han et al. 2016). Among these regulatory elements, AG-rich motifs were identified to have tissue-specific silencing function, with the most potent capable of switching off >95% of gene expression (Han et al. 2016). Transcription factors including YY1, Sp3 and HMGB2 were identified to bind to the AG-rich motifs and to silence gene expression in mammalian cells. Similar AG-rich motifs were observed in venom genes belonging to other toxin families, such as PLA_2 in *L. semifasciata*, but not in their physiological counterparts, nor in housekeeping genes (SDHA, GAPDH, TBP, RPL13A and ACTG1) (Han et al. 2016). Interestingly, the insertions harbouring the AG-rich motifs have low sequence similarity between distinct toxin genes; bearing in mind that venom gene families have different chromosomal locations and recruitment times, it is plausible that the evolutionary histories and mechanisms involved are distinct across the various gene families that encode venom toxins (Fujimi et al. 2002a, Fujimi 2004; Han et al. 2016).

Some of the AG-rich motifs identified in TroD are similar to polycomb response elements. These elements recruit polycomb group proteins and transcription factors to repress transcription of particular target genes (Han et al. 2016; Lanzuolo and Orlando 2012). In *Drosophila*, transcription factors binding to polycomb response element sites include Pho, Dsp1 and Spps, which are homologous to YY1, HMGB2 and Sp-1/3, respectively, in snakes (Lanzuolo and Orlando 2012; Kassis and Brown 2013). Since these transcription factors and their binding sites are conserved through *Drosophila* to mammals, it is conceivable to hypothesize that they may play a similar regulatory role in venom gland-specific expression. Investigations to test whether polycomb repressive complexes play crucial roles in venom gland-specific expression will thus likely be of fundamental importance to also expand our understanding of polycomb group proteins in non-model organisms, and therefore of gene regulation in general.

Modes and mechanisms of selective toxin expression

Another layer of complexity relating to the regulation of venom is those factors that control which toxins are produced, when they are produced, and how much of them are produced. Of

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

the numerous toxin genes found in the genome of venomous animals, it is apparent that not all of them make peptides/proteins that are detectable in secreted venoms - some are not transcribed, whereas others are modified or not translated. The selective expression of toxins has important implications: it can increase venom diversity and complexity which, in turn, might be advantageous, for example, to maximize the chance of targeting a crucial receptor across a diverse range of prey, or for compensating against the evolution of toxin resistance (Arbuckle et al. 2017). While tissue-specific expression of toxin genes seems to be mostly controlled by transcriptional regulation, e.g. gene silencing, variation in venom proteins appear to be regulated both by pre- and post-transcriptional processes (Figure 3).

Spatiotemporal variation in venom composition

In recent years, an increasing number of venomous animals have been found to produce toxins in a heterogeneous way across venom-secreting tissue. Cone snails are capable of producing distinct venoms for use either as paralytic toxins against predators or to elicit a state of sensory deprivation and hypoactivity for prey immobilization. The distinct sets of toxins associated with these predator- or prey-specific venoms are differentially expressed and synthesized along the venom duct, with defence-evoked toxins secreted in the proximal region and predation-evoked venom produced in the distal region (Dutertre et al. 2014). A similar compartmentalization has also been observed in centipedes (Undheim et al. 2015) and assassin bugs (Walker et al. 2018), with both cases similar to cone snails, where venom with different functions are produced in different parts of the venom gland. In all cases, transcriptomic and proteomic levels correlate, suggesting a regulatory machinery for differential transcription, either through silencing or enhancing, or a combination of both. Although no evidence exists to suggest that snakes are capable of producing analogous, functionally distinct, venoms from the same gland, recent work suggests some elements of toxin localization may exist in the venom gland. Single cell RNA-seq revealed that different cell populations of venom gland-derived organoids differentially express distinct toxin classes (Post et al. 2020), while mass spectrometry imaging showed that different toxins are spatially distributed in a highly variable manner in snake venom glands (Hamilton et al. 2020).

Alternative splicing

A common post-transcriptional regulatory mechanism is alternative splicing, which is often generated by exon skipping or intron retention. The latter is most common in plant, fungi, and unicellular eukaryotes; it is generally associated with down-regulation of gene expression via

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

nonsense mediated decay (NMD), whereby the retained intron sequences interrupt the main open reading frame (ORF) of the mRNA leading to the introduction of premature stop-codons or secondary start-codons (Jacob and Smith 2017). Both mechanisms occur in the habu snake (*P. flavoviridis*), where extensive alternative splicing was observed in three protein families, namely metalloproteinase (SVMP), serine protease (SVSP) and vascular endothelial growth factor (VEGF); however, no other venom genes, nor their physiological counterparts, exhibited evidence of alternative splicing (Ogawa et al. 2019). Besides *cis*-spliced transcripts, *trans*-spliced variants were also found to originate from multiple, tandemly clustered genes (Ogawa et al. 2019). Chimeric mRNA sequences from PLA_2 toxin-encoding genes have been described from the prairie rattlesnake (*Crotalus viridis*) (Tsai et al. 2003); however, these chimeric PLA_2s were absent in the venom, suggesting a mechanism (e.g. NMD) preventing their translation or secretion. A *trans*-spliced toxin transcript was also found in the Chinese scorpion (*Mesobuthus martensii*); this was characterized by an early stop codon and multiple ORFs suggesting rapid degradation by the NMD pathway (Zhu et al. 2001). The common house spider (*Parasteatoda tepidariorum*) also appears to show complex patterns of transcription of venom genes, produced by alternative 5' and 3' splice sites, exon skipping, mutually exclusive exons and alternative first and last exons, but not via intron retention (Haney et al. 2019). Interestingly, Bhere et al. (2014) found the presence of an intron in the 3' UTR of a α-latrotoxin paralog of the black widow spider, which may result in the transcript having reduced translation via the NMD pathway. Indeed, the α-latrotoxin paralog was not detected by mass spectrometry in the venom, whereas the main α-latrotoxin, which does not possess the intron in the 3' UTR, was secreted.

Compared to other venomous animals, variation in venom toxins in sea anemones appears to be controlled by both transcriptional and post-transcriptional processes. Several toxins display variation in expression patterns across developmental stages and body tissues, and their transcriptomic and proteomic abundances correlate well, suggesting that transcriptional mechanisms control for spatiotemporal dynamics (Moran et al. 2012b; Macrander et al. 2016; Colombus-Shenkar et al. 2018; Sachkova et al. 2019; Surm et al. 2019). However, the neurotoxin Nv1 is transcribed throughout the whole life cycle of *Nematostella*, but the embryos and planulas lack the mature toxin due to intron retention in the Nv1 transcript (Moran et al. 2008).

MicroRNAs

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Besides the regulatory processes discussed above, another post-transcriptional mechanism identified as being responsible for, or contributing to, variation in venom composition, are microRNAs (miRNA). miRNA represents a class of small (~22 nt) noncoding RNA that regulate gene expression by degrading their target mRNA and/or inhibiting their translation (Ambros 2004). Durban et al. (2013) demonstrated that age-dependent changes in the concentration of miRNAs that target SVMP and crotoxin (a neurotoxic PLA_2) toxin-encoding mRNAs modulate the transition from juvenile crotoxin-rich venom to adult SVMP-rich venom in the rattlesnake species *Crotalus simus*. The same mechanism was then described in three other congeneric rattlesnake species (Durban et al. 2017), suggesting miRNAs may play an important role in modulating ontogenetic shifts in venom, in snakes at least. Interestingly, miRNAs appear to have a dual action in this case, as they not only silence the translation of crotoxin, but seemingly simultaneously up-regulate SVMP-targeting mRNAs. Surprisingly, the role of miRNA in venom regulation has only been investigated in depth in these few species of rattlesnakes, and thus it would be of great interest to extend such investigations of small RNAs to other venomous lineages.

Concluding remarks and future directions

In summary, venomous organisms offer exciting models for addressing key question relating to evolutionary and molecular biology (Figure 4), and with ever increasing advances in -omics technologies, the appreciable value of venom research is gaining momentum among the wider scientific community (Holford et al. 2018; Post et al. 2020; Suryamohan et al. 2020). However, much more research effort is required before a broad understanding of the molecular mechanisms underlying the origin, evolution and regulation of diverse animal venom systems can be realized. In particular, we encourage researchers to:

i. establish representatives of different venomous lineages as models that can be applied to expand our knowledge on the origin of venom-secreting cells and tissues. In turn, these models will allow to robustly explore one of the fundamental questions in evolutionary biology: how did distinct animal lineages evolve similar traits?
ii. Expand existing genomic data from venomous animals by improving the assemblies of currently available genomes, and using new sequencing technologies to generate new genomic data. These data resources are required to underpin the adoption of advanced genomics techniques (e.g. PCHi-C, Chip-Seq) to shed light onto the GRN of venom-secreting cells.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

iii. Embark on comparative genomics studies of regulatory regions of toxin genes across different protein families and animal lineages to explore whether the regulatory elements associated with the convergent molecular evolution of toxin families have also evolved in a repeatable manner.

The considerable amount of sequencing data currently available for a diverse array of venomous organisms already provides a valuable initial resource for such comparative analyses to identify elements involved in the GRN. The role of *cis*- and *trans*-regulatory elements resulting from these comparisons, together with those already reported in literature, can be tested by means of genomic manipulations such as CRISPR/Cas9 - already successfully employed in *Nematostella* - and RNA interference (RNAi) - which has been established in spiders to study embryonic development and segmentation (e.g. Schönauer et al. 2016). These data, coupled with gene expression levels, proteomic data, mass spectrometry imaging and single cell sequencing approaches, will continue to advance the field in helping to identify the regulatory processes and elements involved in the selective expression of novel adaptive genes and members of the GRN associated with the emergence, retention and diversification of evolutionary novelties. Ultimately, we believe that the convergent evolution of animal venoms, which have emerged frequently across the breadth of the animal tree of life, can be harnessed as valuable models to provide powerful insights in to the generalities of the mechanisms that underpin the evolution of novel adaptations.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank M. Robinson-Rechavi for fruitful insights, and Yehu Moran and two anonymous reviewers for their feedback and advice that helped broaden the scope of a first draft of this manuscript.

References

Alpers F. 1931. Zur Kenntnis der Anatomie von *Conus lividus* BRUG., besonders des Darmkanals. *Jenaische Z. Nat.* 65:587-658.

Ambros V. 2004. The functions of animal microRNAs. *Nature* 431:350-355.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Arbuckle K, Rodríguez de la Vega RC, Casewell NR. 2017. Coevolution takes the sting out of it: Evolutionary biology and mechanisms of toxin resistance in animals. *Toxicon* 140:118-131.

Babonis LS, Martindale MQ. 2014. Old cell, new trick? Cnidocytes as a model for the evolution of novelty. *Integr Comp Biol*. 54:714-722.

Babonis LS, Martindale MQ. 2017. PaxA, but not PaxC , is required for cnidocyte development in the sea anemone *Nematostella vectensis. EvoDevo* 8:1-20.

Ball AD. 2002. Foregut ontogeny of the Neogastropoda: comparison of development in Nucella lapillus and Conus anemone. *Boll. Malacologico Supplemento* 4:51-78.

Bhere KV, Haney RA, Ayoub NA, Garb JE. 2014. Gene structure, regulatory control, and evolution of black widow venom latrotoxins. *FEBS Letters* 588:3891-3897.

Blake JA, Ziman MR. 2014. Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development* 141:737-751.

Bowles J, Schepers G, Koopman P. 2000. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol*. 227:239-255.

Buczek O, Bulaj G, Olivera BM. 2005. Conotoxins and the posttranslational modification of secreted gene products. *Cell Mol Life Sci*. 62:3067-3079.

Cameron AM, Endean R. 1973. Epidermal secretions and the evolution of venom glands in fishes. *Toxicon* 11:401-410.

Cao Z, Yu Y, Wu Y, Hao P, Di Z, He Y, Chen Z, Yang W, Shen Z, He X, et al. 2013. The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods. *Nat Commun.* 4:1–10.

Casewell NR, Petras D, Card DC, Suranse V, Mychajliw AM, Richards D, Koludarov I, Albulescu LO, Slagboom J, Hempel BF, et al. 2019. Solenodon genome reveals convergent evolution of venom in eulipotyphlan mammals. *PNAS* 116:25745–25755.

Casewell NR, Wagstaff SC, Harrison RA, Renjifo C, Wüster W. 2011. Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Mol Biol Evol*. 28:2637-2649.

Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. 2013. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol.* 28:219-229.

Chang D, Duda Jr TF. 2012. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol Biol Evol.* 29:2019-2029.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Columbus-Shenkar YY, Sachkova MY, Macrander J, Fridrich A, Modepalli V, Reitzel AM, Sunagar K, Moran Y. 2018. Dynamics of venom composition across a complex life cycle. *eLife* 7:e35014.

Cousin X, Bon S, Massoulié J, Bon C. 1998. Identification of a novel type of alternatively spliced exon from the acetylcholinesterase gene of *Bungarus fasciatus*. Molecular forms of acetylcholinesterase in the snake liver and muscle. *J Biol Chem.* 73:9812-9820.

Darling JA, Reitzel AR, Burton PM, Mazza ME, Ryan JF, Sullivan JC, Finnerty JR. 2005. Rising starlet: the starlet sea anemone, *Nematostella vectensis. Bioessays* 27:211–221.

David CN. 2012. Interstitial stem cells in *Hydra*: multipotency and decision-making. *Int J Dev Biol.* 56:489–497.

Davidson EH. 2010. Emerging properties of animal gene regulatory networks. *Nature* 468:911–920.

Delgado-Prudencio G, Possani LD, Becerril B, Ortiz E. 2019. The ual α-amidation system in scorpion venom glands. *Toxins* 11:425.

Denker E, Bapteste E, Le Guyader H, Manuel M, Rabet N. 2008. Horizontal gene transfer and the evolution of cnidarian stinging cells. *Curr Biol.* 18:R858–859.

Drukewitz SH, Bokelmann L, Undheim EAB, von Reumont BM. 2019. Toxins from scratch? Diverse, multimodal gene origins in the predatory robber fly *Dasypogon diadema* indicate a dynamic venom evolution in dipteran insects. *GigaScience* 8:giz081.

Durban J, Pérez A, Sanz L, Gómez A, Bonilla F, Rodríguez S, Chacón D, Sasa M, Angulo Y, Gutiérrez JM, Calvete JJ. 2013. Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake, *Crotalus simus simus. BMC Genomics* 14:234.

Durban J, Sanz L, Trevisan-Silva D, Neri-Castro E, Alagoón A, Calvete JJ. 2017. Integrated venomics and venom gland transcriptome analysis of juvenile and adult Mexican rattlesnakes *Crotalus simus*, *C. tzabcan* and *C. culminatus* revealed miRNA-modulated ontogenetic shifts. *J Proteome Res*. 16:3370-3390.

Dutertre S, Jin A-H, Vetter I, Hamilton B, Sunagar K, Lavergne V, Dutertre V, Fry BG, Antunes A, Venter DJ, et al. 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nat Commun.* 5:1–9.

Fane M, Harris L, Smith AG, Piper M. 2017. Nuclear factor one transcription factors as epigenetic regulators in cancer. *Int J Cancer* 140:2634-2641.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Flici H, Schnitzler CE, Millane RC, Govinden G, Houlihan A, Boomkamp SD, Shen SB, Baxevanis AD, Frank U. 2017. An evolutionarily conserved SoxB-Hdac2 crosstalk regulates neurogenesis in a cnidarian. *Cell Rep*. 18:1395–1409.

Foelix R, Erb B. 2010. Mesothelae have venom glands. *J Arachnol.* 38:596–598.

Fox RC, Scott CS. 2005. First evidence of a venom delivery apparatus in extinct mammals. *Nature* 435:1091.

Fretter V, Graham A. 1962. British prosobranch gastropods. London: Ray Society.

Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JDA, King GF, Nevalainen TJ, Norman JA, Lewis RJ, Norton RS, et al. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genom Hum G.* 10:483–511.

Fry BG. 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 Res*. 15:403–420.

Fujimi TJ, Kariya Y, Tsuchiya T, Tamiya T. 2002a. Nucleotide sequence of phospholipase A2 gene expressed in snake pancreas reveals the molecular evolution of toxic phospholipase A2 genes. *Gene* 292:225-231.

Fujimi TJ, Tsuchiya T, Tamiya T. 2002b. A comparative analysis of invaded sequences from group IA phospholipase A2 genes provides evidence about the divergence period of genes groups and snake families. *Toxicon* 40:873-884.

Fujimi TJ, Yasuoka S, Ogura E, Tsuchiya T, Tamiya T. 2004. Comparative analysis of gene expression mechanisms between group IA and IB phospholipase A2 genes from sea snake *Laticauda semifasciata*. *Gene* 332:179-190.

Galliot B, Quiquand M. 2011. A two-step process in the emergence of neurogenesis. *Eur J Neurosci.* 34:847–862.

Galliot B. 2012. *Hydra*, a fruitful model system for 270 years. *Int J Dev Biol.* 56:411–423.

Gavelis GS, Wakeman KC, Tillmann U, Ripken C, Mitarai S, Herranz M, Özbek S, Holstein T, Keeling PJ, Leander BS. 2017. Microbial arms race: Ballistic "nematocysts" in dinoflagellates represent a new extreme in organelle complexity. *Sci Adv.* 3:e1602552.

Geertz M, Maerkl SJ. 2010. Experimental strategies for studying transcription facto-DNA binding specificities. *Brief Funct Genomics* 9:362-373.

Genikhovich G, Technau U. 2011. Complex functions of Mef2 splice variants in the differentiation of endoderm and of a neuronal cell type in a sea anemone. *Development* 138:4911–4919.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Gilmore TD. 2006. Introduction to NF-κB: players, pathways, perspectives. *Oncogene* 25:6680-6684.

Gold DA, Jacobs DK. 2013. Stem cell dynamics in Cnidaria: are there unifying principles? *Dev Genes Evol.* 223:53–66.

Gold DA, Lau CLF, Fuong H, Kao G, Hartenstein V, Jacobs DK. 2019. Mechanisms of cnidocyte development in the moon jellyfish *Aurelia*. *Evol Dev.* 21:72–81.

Gong E, Martin LD, Burnham DA, Falk AR. 2010. The birdlike raptor *Sinornithosaurus* was venomous. *Proc Natl Acad Sci USA.* 107:766-768.

Gronostajski RM. 2000. Roles of the NFI/CTF gene family in transcription and development. *Gene* 249:31-45.

Hamilton BR, Marshall DL, Casewell NR, Harrison RA, Blanksby SJ, Undheim EAB. 2020. Mapping enzyme activity on tissue by functional mass spectrometry imaging. *Angew Chem*. 132:1-5.

Han SX, Kwong S, Ge R, Kolatkar PR, Woods AE, Blanchet G, Kini RM. 2016. Regulation of expression of venom toxins: silencing of prothrombin activator trocarin D by AG-rich motifs. *FASEB J* 30: 2411-2425.

Haney RA, Matte T, Forsyth FS, Garb JE. 2019. Alternative transcription at venom genes and its role as a complementary mechanism for the generation of venom complexity in the common house spider. *Front Ecol Evol.* 7:85.

Hess J, Angel P, Schorpp-Kistner M. 2004. AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci.* 117:5965–5973.

Holford M, Daly M, King GF, Norton RS. 2018. Venoms to the rescue. *Science* 361:842-844.

Hollenhorst PC. 2012. RAS/ERK pathway transcriptional regulation through ETS/AP-1 binding sites. *Small GTPases* 3:154-158.

Jacob AG, Smith CWJ. 2017. Intron retention as a component of regulated gene expression programs. *Hum Genet.* 136:1043–1057.

Jaskoll T, Choy, HA, Melnick, M. 1994. Glucocorticoids, TGF-beta, and embryonic mouse salivary gland morphogenesis. *J Craniofac Genet Dev Biol.* 14:217-230.

Jeyaseelan K, Ma D, Armugam A. 2001. Real-time detection of gene promoter activity: quantitation of toxin gene transcription. *Nucleic Acids Res.* 29:E58.

Junqueira-de-Azevedo IL, Bastos CMV, Ho PL, Luna MS, Yamanouye N, Casewell NR. 2015. Venom-related transcripts from *Bothrops jararaca* tissues provide novel molecular insights into the production and evolution of snake venom. *Mol Biol Evol.* 32:754-766.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Kas K, Finger E, Grall F, Gu X, Akbarali Y, Boltax, J, Weiss A, Oettgen P, Kapeller R, Libermann TA. 2000. ESE-3, a novel member of an epithelium-specific ETS transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J Biol Chem.* 275:2986-2998.

Kassis JA, Brown JL. 2013. Polycomb group response elements in *Drosophila* and vertebrates. *Adv Genet.* 81:83-118.

Khalturin K, Anton-Erxleben F, Milde S, Plotz C, Wittlieb J, Hemmrich G, Bosch TC. 2007. Transgenic stem cells in *Hydra* reveal an early evolutionary origin for key elements controlling self-renewal and differentiation. *Dev Biol.* 309: 32-44.

Kim M, McGinnis W. 2011. Phosphorylation of grainy head by ERK is essential for wound-dependent regeneration but not for development of an epidermal barrier. *Proc Natl Acad Sci USA* 108:650-655.

Kochva E. 1963. Development of the venom gland and the trigeminal muscles in *Vipera palaestinae*. *Acta Anat*. 52:49-89.

Kochva E. 1987. The origin of snakes and evolution of the venom apparatus. *Toxicon* 25:65–106.

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

Krause WJ. 2010. Morphological and histochemical observations on the crural gland-spur apparatus of the echidna (*Tachyglossus aculeatus*) together with comparative observations on the femoral gland-spur apparatus of the duckbilled platypus (*Ornithorhyncus anatinus*). *CTO* 191:336–354.

Kwong S, Woods AE, Mirtschin PJ, Ge R, Kini RM. 2009. The recruitment of blood coagulation factor X into snake venom gland as a toxin: the role of promoter *cis*-elements in its expression. *Thromb Haemost.* 102:469-478.

Langenegger N, Nentwig W, Kuhn-Nentwig L. 2019. Spider venom: components, modes of action, and novel strategies in transcriptomic and proteomic analyses. *Toxins* 11:611.

Lanzuolo C, Orlando V. 2012. Memories from the polycomb group proteins. *Annu Rev Genet.* 46:561-589.

Layden MJ, Rentzsch F, Röttinger E. 2016. The rise of the starlet sea anemone *Nematostella vectensis* as a model system to investigate development and regeneration. *WIREs Dev Biol.* 5:408–428.

Levy S, Hannenhalli S, Workman C. 2001. Enrichment of regulatory signals in conserved non-coding genomic sequence. *Bioinformatics* 17:871-877.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Luna MS, Hortencio TM, Ferreira ZS, Yamanouye N. 2009. Sympathetic outflow activates the venom gland of the snake *Bothrops jararaca* by regulating the activation of transcription factors and the synthesis of venom gland proteins. *J Exp Biol.* 212:1535-1543.

Ma D, Armugam A, Jeyaseelan K. 2001. Expression of cardiotoxin-2 gene: cloning, characterization and deletion analysis of the promoter. *Eur J Biochem.* 268:1844-1850.

Macrander J, Broe M, Daly M. 2016. Tissue-Specific Venom composition and differential gene expression in sea anemones. *Genome Biol Evol.* 8:2358–2375.

Majewski J, Ott J. 2002. Distribution and characterization of regulatory elements in the human genome. *Genome Res.* 12:1827-1836.

Margres MJ, Bigelow AT, Lemmon EM, Lemmon AR, Rokyta DR. 2017. Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. *Genetics* 206:1569-1580.

Martinson EO, Martinson VG, Edwards R, Mrinalini, Werren JH. 2016. Laterally transferred gene recruited as a venom in parasitoid wasps. *Mol Biol Evol.* 33:1042–1052.

Martinson EO, Mrinalini, Kelkar YD, Chang C-H, Werren JH. 2017. The evolution of venom by co-option of single-copy genes. *Curr Biol.* 27:2007-2013.

Matus DQ, Pang K, Daly M, Martindale MQ. 2007. Expression of Pax gene family members in the anthozoan cnidarian, *Nematostella vectensis*. *Evol Dev.* 9:25–38.

Meng Q, Xia Y. 2011. c-Jun, at the crossroad of the signaling network. *Protein Cell* 2:889–898.

Miljkovic-Licina M, Chera S, Ghila L, Galliot B. 2007. Head regeneration in wild-type *Hydra* requires de novo neurogenesis. *Development* 134:1191–1201.

Modica MV, Holford M. 2010. The Neogastropoda: evolutionary innovations of predatory marine snails with remarkable pharmacological potential. In: Pontarotti P, editor. Evolutionary Biology – Concepts, Molecular and Morphological Evolution: 13th Meeting 2009. Berlin, Heidelberg: Springer. p. 249–270.

Moran Y, Fredman D, Szczesny P, Grynberg M, Technau U. 2012a. Recurrent horizontal transfer of bacterial toxin genes to Eukaryotes. *Mol Biol Evol.* 29:2223–2230.

Moran Y, Genikhovich G, Gordon D, Wienkoop S, Zenkert C, Özbek S, Technau U, Gurevitz M. 2012b. Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proc R Soc B.* 279:1351–1358.

Moran Y, Weinberger H, Reitzel AM, Sullivan JC, Kahn R, Gordon D, Finnerty JR, Gurevitz M. 2008. Intron retention as a posttranscriptional regulatory mechanism of neurotoxin

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

expression at early life stages of the starlet anemone *Nematostella vectensis*. *J Mol Biol.* 380:437–443.

Moura-da-Silva AM, Almeida MT, Portes-Junior JA, Nicolau CA, Gomes-Neto F, Valente RH. 2016. Processing of snake venom metalloproteinases: generation of toxin diversity and enzyme inactivation. *Toxins* 8:183.

Nakamura H, Murakami T, Hattori S, Sakaki Y, Ohkuri T, Chijiwa T, Ohno M, Oda-Ueda N. 2014. Epithelium specific ETS transcription factor, ESE-3, of *Protobothrops flavoviridis* snake venom gland transactivates the promoters of venom phospholipase A2 isozyme genes. *Toxicon.* 92:133-139.

Ogawa T, Oda-Ueda N, Hisata K, Nakamura H, Chijiwa T, Hattori S, Isomoto A, Yugeta H, Yamasaki S, Fukumaki Y, et al. 2019. Alternative mRNA splicing in three Venom families underlying a possible production of divergent venom proteins of the habu snake, *Protobothrops flavoviridis. Toxins* 11:581.

Ozbek S, Balasubramanian PG, Holstein TW. 2009. Cnidocyst structure and the biomechanics of discharge. *Toxicon* 54:1038–1045.

Page LR. 2011. Metamorphic remodeling of a planktotrophic larva to produce the predatory feeding system of a cone snail (Mollusca, Neogastropoda). *Biol Bull.* 221:176–188.

Pantin CFA. 1942. Excitation of nematocysts. *Nature* 149:109.

Ponder WF, Lindberg DR, Ponder JM. 2019. The oesophagus. In: Biology and evolution of the Mollusca. Boca Raton: CRC Press.

Ponder WF, Lindberg DR. 1997. Towards a phylogeny of gastropod mollusks: an analysis using morphological characters. *Zool J Linn Soc.* 119:83–265.

Ponder WF. 1973. The origin and evolution of the Neogastropoda. *Malacologia* 12:295-338.

Post Y, Puschhof J, Beumer J, Kerkkamp HM, de Bakker MAG, Slagboom J, de Barbanson B, Wevers NR, Spijkers XM, Olivier T, et al. 2020. Snake venom gland organoids. *Cell* 180:233-247.

Ray A, Prefontaine KE. 1994. Physical association and functional antagonism between the p65 subunit of transcription factor NF-kappa B and the glucocorticoid receptor. *Proc Natl Acad Sci US*A 91:752-756.

Reyes-Velasco J, Card DC, Andrew AL, Shaney KJ, Adams RH, Schield DR, Casewell NR, Mackessy SP, Castoe TA. 2014. Expression of venom gene homologs in diverse python tissues suggests a new model for the evolution of snake venom. *Mol Biol Evol.* 32:173-183.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Reza MA, Swarup S, Kini RM. 2007. Structure of two genes encoding parallel prothrombin activators in *Tropidechis carinatus* snake: gene duplication and recruitment of factor X gene to the venom gland. *J Thromb Haemost.* 5:117-126.

Richards GS, Rentzsch F. 2014. Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian *Nematostella vectensis. Development 141:*4681–4689.

Sachkova M, Burkhardt P. 2019. Exciting times to study the identity and evolution of cell types. *Development* 146:dev178996.

Sachkova MY, Singer SA, Macrander J, Reitzel AM, Peigneur S, Tytgat J, Moran Y. 2019. The birth and death of toxins with distinct functions: a ase study in the sea anemone *Nematostella. Mol Biol Evol.* 36:2001–2012.

Sagai T, Hosoya M, Mizushina Y, Tamura M, Shiroishi T. 2005. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132:797–803.

Schendel V, Rash LD, Jenner RA, Undheim EAB. 2019. The diversity of venom: the importance of behavior and venom system morphology in understanding its ecology and evolution. *Toxins.* 11:666.

Schield DR, Card DC, Hales NR, Perry BW, Pasquesi GM, Blackmon H, Adams RH, Corbin AB, Smith CF, Ramesh B, et al. 2019. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. *Genome Res.* 29:590-601.

Schönauer A, Paese CLB, Hilbrant M, Leite DJ, Schwager EE, Feitosa NM, Eibner C, Damen WGM, McGregor AP. 2016. The Wnt and Delta-Notch signalling pathways interact to direct pair-rule gene expression via caudal during segment addition in the spider *Parasteatoda tepidariorum. Development* 143:2455–2463.

Schoenfelder S, Fraser P. 2019. Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet*. 20:437–455.

Schoenfelder S, Javierre B-M, Furlan-Magaril M, Wingett SW, Fraser P. 2018. Promoter Capture Hi-C: High-resolution, Genome-wide Profiling of Promoter Interactions. *J Vis Exp*. 136:e57320.

Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé M-P, Renno J, Loe-Mie Y, Lifshitz A, Mukamel Z, Schmutz S, et al. 2018. Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-Seq. *Cell* 173:1520-1534.

Shaw PA, Chaparro O. 1999. The 5'-flanking sequence and regulatory elements of the cystatin S gene. *Biochem Biophys Res Commun.* 261:705-711.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Shayer-Wollberg M, Kochva E. 1967. Embryonic development of the venom apparatus in *Causus rhombeatus* (Viperidae, Ophidia). *Herpetologica* 23:249–259.

Shibata H, Chijiwa T, Oda-Ueda N, Nakamura H, Yamaguchi K, Hattori S, Matsubara K, Matsuda Y, Yamashita A, Isomoto A, et al. 2018. The habu genome reveals accelerated evolution of venom protein genes. *Sci Rep.* 8:1–11.

Shostak S. 1993. A symbiogenetic theory for the origins of cnidocysts in cnidaria. *Biosystems* 29:49–58.

Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, Juliano CE. 2019. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science* 365:eaav9314.

Silva LM, Fortes-Dias CL, Schaffert PP, Carvalho Botelho AC, Nacif-Pimenta R, Estevão-Costa MI, Cordeiro M do N, Paolucci Pimenta PF. 2011. Developmental biology of the Brazilian 'armed' spider *Phoneutria nigriventer* (Keyserling, 1891): microanatomical and molecular analysis of the embryonic stages. *Toxicon* 57:19–27.

Simposon SJ, Casas J. 2011. Spider physiology and behaviour: physiology vol. 41. In: Casas J editor. Advances in insect physiology. London: Academic Press.

Smith, E. H. 1967. The proboscis and oesophagus of some British turrids. *Trans R Soc Edin*. 67:1–22.

Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Gen.* 14:751-764.

Storbeck CJ, Sabourin LA, Waring JD, Korneluk RG. 1998. Definition of regulatory sequence elements in the promoter region and the first intron of the myotonic dystrophy protein kinase gene. *J Biol Chem.* 273:9139-9147.

Suga H, Tschopp P, Graziussi DF, Stierwald M, Schmid V, Gehring WJ. 2010. Flexibly deployed Pax genes in eye development at the early evolution of animals demonstrated by studies on a hydrozoan jellyfish. *Proc Natl Acad Sci USA* 107:14263–14268.

Sunagar K, Columbus-Shenkar YY, Fridrich A, Gutkovich N, Aharoni R, Moran Y. 2018. Cell type-specific expression profiling unravels the development and evolution of stinging cells in sea anemone. *BMC Biology* 16:108.

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:e1005596.

Surm JM, Smith HL, Madio B, Undheim EAB, King GF, Hamilton BR, Burg CA van der, Pavasovic A, Prentis PJ. 2019. A process of convergent amplification and tissue-specific

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

expression dominates the evolution of toxin and toxin-like genes in sea anemones. *Mol Ecol.* 28:2272–2289.

Suryamohan K, Krishnankutty SP, Guillory J, Jevit M, Schröder MS, Wu M, Kuriakose B, Mathew OK, Perumal RC, Koludarov I, et al. 2020. The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. *Nat Genet.* 52:106–117.

Technau U, Rudd S, Maxwell P, Gordon P, Saina M, Grasso L, Hayward D, Sensen C, Saint R, Holstein T. 2005. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends Genet.* 21:633–639.

Ting SB, Caddy J, Hislop N, Wilanowski T, Auden A, Zhao L-l, Ellis S, Kaur P, Uchida Y, Holleran WM. 2005. A homolog of *Drosophila* grainy head is essential for epidermal integrity in mice. *Science* 308:411-413.

Tsai I-H, Wang Y-M, Chen Y-H, Tu AT. 2003. Geographic variations, cloning, and functional analyses of the venom acidic phospholipases A2 of *Crotalus viridis viridis. Arch Biochem Biophys.* 411:289–296.

Undheim EAB, Hamilton BR, Kurniawan ND, Bowlay G, Cribb BW, Merritt DJ, Fry BG, King GF, Venter DJ. 2015. Production and packaging of a biological arsenal: evolution of centipede venoms under morphological constraint. *Proc Natl Acad Sci USA* 112:4026–4031.

von Reumont BM, Campbell LI, Jenner RA. 2014. Quo vadis venomics? A roadmap to neglected venomous invertebrates. *Toxins* 6:3488–3551.

Vonk FJ, Casewell NR, Henkel CV, Heimberg AM, Jansen HJ, McCleary RJ, Kerkkamp HM, Vos RA, Guerreiro I, Calvete, JJ, et al. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc Natl Acad Sci USA* 110:20651-20656.

Walker AA, Mayhew ML, Jin J, Herzig V, Undheim EAB, Sombke A, Fry BG, Meritt DJ, King GF. 2018. The assassin bug *Pristhesancus plagipennis* produces two distinct venoms in separate gland lumens. *Nat Commun.* 9:1–10.

Wang X, Gao B, Zhu S. 2017. Exon shuffling and origin of scorpion venom biodiversity. *Toxins* 9:10.

Whittington CM, Papenfuss AT, Bansal P, Torres AM, Wong ES, Deakin JE, Graves T, Alsop A, Schatzkamer K, Kremitzki C et al. 2008. Defensins and the convergent evolution of platypus and reptile venom genes. *Genome Res.* 18:986-994.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Wong ESW, Papenfuss AT, Whittington CM, Warren WC, Belov K. 2012. A limited role for gene duplications in the evolution of platypus venom. *Mol Biol Evol.* 29:167–177.

Wright JJ. 2009. Diversity, phylogenetic distribution, and origins of venomous catfishes. *BMC Evol Biol.* 9:282.

Wright JJ. 2017. Evolutionary history of venom glands in the siluriformes. In: Malhotra A, editor. Evolution of Venomous Animals and Their Toxins. Toxinology. Dordrecht: Springer Netherlands. p. 279–301.

Wu Y, Wang L, Zhou M, You Y, Zhu X, Qiang Y, Qin M, Luo S, Ren Z, Xu A. 2013. Molecular evolution and diversity of *Conus* peptide toxins, as revealed by gene structure and intron sequence analyses. *PLoS ONE* 8:e82495.

Zeng X-C, Zhang L, Nie Y, Luo X. 2012. Identification and molecular characterization of three new K+-channel specific toxins from the Chinese scorpion *Mesobuthus martensii* Karsch revealing intronic number polymorphism and alternative splicing in duplicated genes. *Peptides* 34:311-323.

Zenkert C, Takahashi T, Diesner MO, Ozbek S. 2011. Morphological and molecular analysis of the *Nematostella vectensis* cnidom. *PLoS ONE* 6:e22725.

Zhijian C, Chao D, Shijin Y, Yingliang W, Jiqun S, Yonggang S, Wenxin L. 2006. Splicing of scorpion toxin gene BmKK2 in HEK 293T cells. J Biochem Mol Toxicol. 20:1–6.

Zhu S, Li W, Cao Z. 2001. A naturally occurring non-coding fusion transcript derived from scorpion venom gland: implication for the regulation of scorpion toxin gene expression. *FEBS Letters* 508:241–244.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Figure Legends

Figure 1. Diversity of venomous animals and their venom apparatuses. (A) The starlet sea anemone, *Nematostella vectensis*; (B) cone snail, *Conus legatus*, with extended proboscis; (C) African rain spider, *Palystes* sp., showing chelicerae; (D) scorpion *Parabuthus* sp. with stinger; (E) platypus, *Ornithorhynchus anatinus*; (F) the Mediterranean banded centipede, *Scolopendra cingulata*, with forcipules; (G) bluespotted stingray, *Taeniura lymma*, with spines on the tail; (H) spitting cobra, *Naja mossambica*, spitting venom from its fangs. Photo credits: Wikimedia Commons contributors (A, E, F), D. Massemin (B), W. Wüster (C, D, H), G. Zancolli (G).

Figure 2. Schematic illustration of our current knowledge of the regulatory elements and mechanisms detected in different animal lineages that influence transcriptional and post-transcriptional regulation of venom genes.

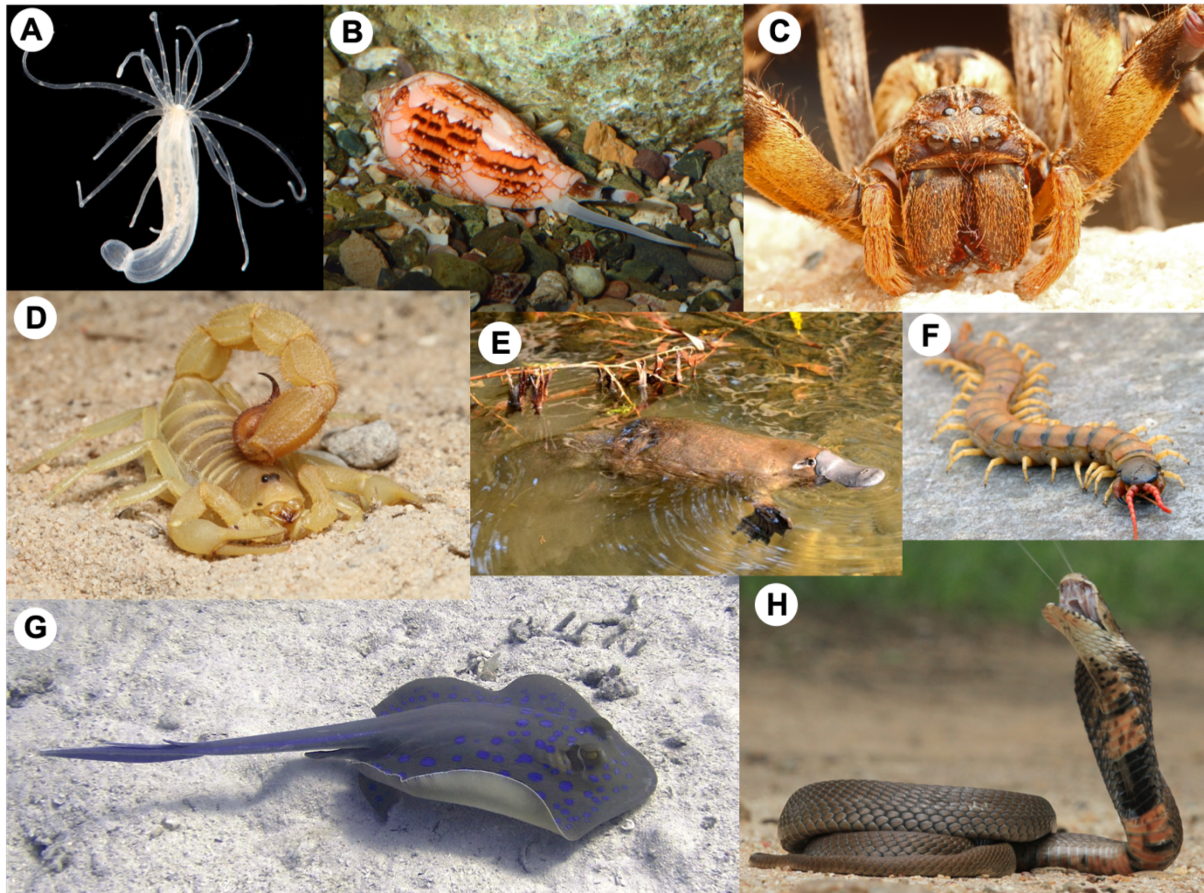
Figure 3. Schematic of the mechanisms that generate venom toxin diversity. The complexity and variability of the toxins detected in venoms are the result of a variety of processes acting at different levels of the genome to proteome continuum. First, at the genome level, novel toxins can arise by gene duplication and subsequent evolution of the resulting paralogs. During transcription, multiple isoforms can be produced by alternative splicing. During translation of the resulting mRNAs, certain transcripts can be silenced by the action of microRNAs (miRNA) and other mechanisms. Finally, toxin proteins are often further modified by means of post-translational (PT) modifications, which can include proteolytic cleavage, protein splicing and the formation of multimeric structures.

Figure 4. Summary of outstanding questions relating to the origin and regulation of evolutionary novelties that can be addressed by using venomous animals as model systems.

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

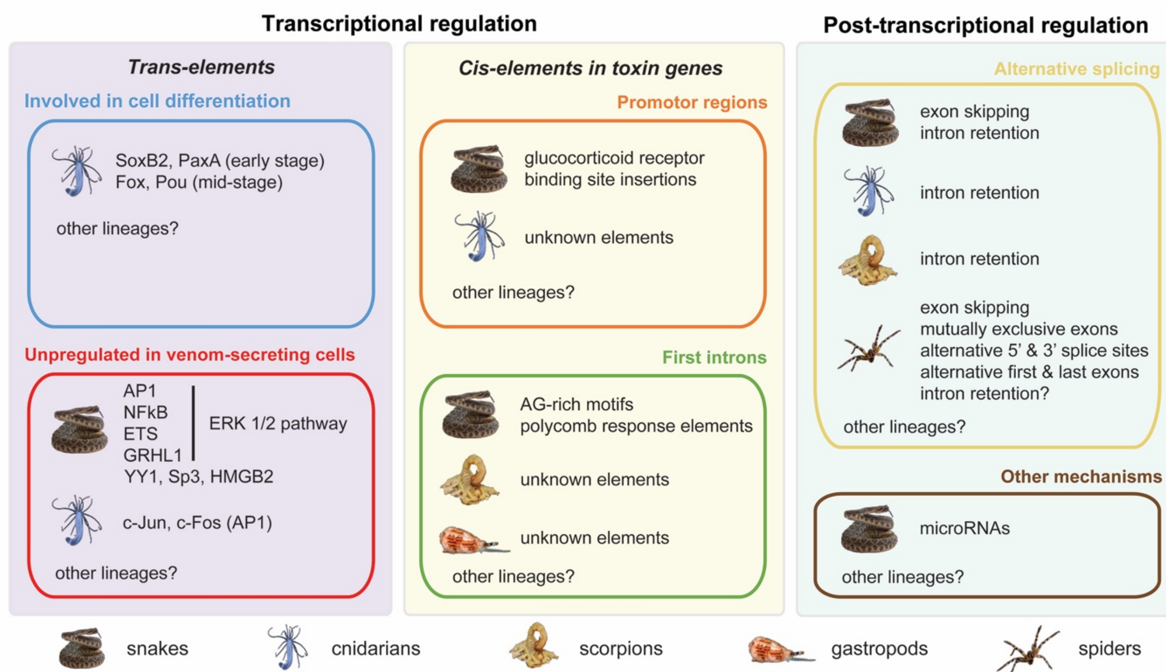
Figures

Figure 1



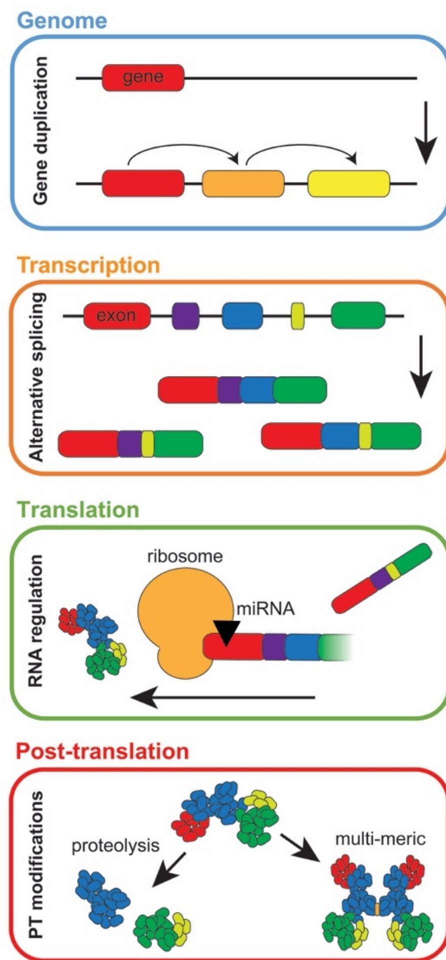
Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Figure 2



Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Figure 3.



Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020

Figure 4.

OUTSTANDING QUESTIONS

Convergence in animal venoms

- How have animals independently evolved different organs and anatomical structures that perform the same function?
- To what extent is this independent evolution of novel adaptive traits associated with convergence in the GRN?
- Have different animal lineages adopted similar molecular mechanisms to evolve similar traits?

Origins of venom-secreting cells

- To what extent do novel cell types inherit the GRN and patterns of expression from their tissue of origin?
- What is the relative contribution of conserved and derived genes in the GRN of novel cell types?
- Do novel genes or the differential orchestration of pre-existing genes contribute to the evolution of novel traits?

Evolution of venom gene families

- How many times have certain orthologs been recruited for the same purpose across the tree of life?
- What makes certain gene families amenable for expansion by gene duplication?
- Do certain chromosomal locations play important roles in underpinnning gene family expansion?

Toxin gene expression patterns

- Are the expression patterns of novel genes inherited from their closely related paralogs?
- How does the novel regulatory architecture associated with new paralogs develop? Are existing regulatory elements modified, or do transposable elements create new transcription factor binding sites?
- Which transcriptional or post-transcriptional mechanisms control for adaptive differential expression?

Downloaded from https://academic.oup.com/mbe/advance-article-abstract/doi/10.1093/molbev/msaa133/5848011 by guest on 02 June 2020