Replacement and Parallel Simplification of Nonhomologous Proteinases Maintain Venom Phenotypes in Rear-Fanged Snakes

Juan David Bayona-Serrano,[1] Vincent Louis Viala,[1,2] Rhett M. Rautsaw,[3] Tristan D. Schramer,[3] Gesiele A. Barros-Carvalho,[1] Milton Yutaka Nishiyama Jr,[1,2] Luciana A. Freitas-de-Sousa,[4] Ana Maria Moura-da-Silva,[4,5] Christopher L. Parkinson ,[3,6] Felipe Gobbi Grazziotin,[7] and Inácio L.M. Junqueira-de-Azevedo *,[1,2]

[1]Laboratório Especial de Toxinologia Aplicada, Instituto Butantan, São Paulo, Brazil

[2]Center of Toxins, Immune-Response and Cell Signaling (CeTICS), São Paulo, Brazil

[3]Department of Biological Sciences, Clemson University, Clemson, SC

[4]Laboratório de Imunopatologia, Instituto Butantan, São Paulo, Brazil

[5]Instituto de Pesquisa Clínica Carlos Borborema, Fundação de Medicina Tropical Dr. Heitor Vieira Dourado, Manaus, Brazil

[6]Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC

[7]Laboratório de Coleções Zoológicas, Instituto Butantan, São Paulo, Brazil

***Corresponding author:** E-mail: inacio.azevedo@butantan.gov.br.

Associate editor: John True

Raw transcriptomic data are available at NCBI's GenBank under Bioproject accession numbers PRJNA625548 and PRJNA88989 (for *Conophis lineatus*).

Abstract

Novel phenotypes are commonly associated with gene duplications and neofunctionalization, less documented are the cases of phenotypic maintenance through the recruitment of novel genes. Proteolysis is the primary toxic character of many snake venoms, and ADAM metalloproteinases, named snake venom metalloproteinases (SVMPs), are largely recognized as the major effectors of this phenotype. However, by investigating original transcriptomes from 58 species of advanced snakes (Caenophidia) across their phylogeny, we discovered that a different enzyme, matrix metalloproteinase (MMP), is actually the dominant venom component in three tribes (Tachymenini, Xenodontini, and Conophiini) of rear-fanged snakes (Dipsadidae). Proteomic and functional analyses of these venoms further indicate that MMPs are likely playing an "SVMP-like" function in the proteolytic phenotype. A detailed look into the venom-specific sequences revealed a new highly expressed MMP subtype, named snake venom MMP (svMMP), which originated independently on at least three occasions from an endogenous MMP-9. We further show that by losing ancillary noncatalytic domains present in its ancestors, svMMPs followed an evolutionary path toward a simplified structure during their expansion in the genomes, thus paralleling what has been proposed for the evolution of their Viperidae counterparts, the SVMPs. Moreover, we inferred an inverse relationship between the expression of svMMPs and SVMPs along the evolutionary history of Xenodontinae, pointing out that one type of enzyme may be substituting for the other, whereas the general (metallo)proteolytic phenotype is maintained. These results provide rare evidence on how relevant phenotypic traits can be optimized via natural selection on nonhomologous genes, yielding alternate biochemical components.

Key words: **matrix metalloproteinase, protein family evolution, genotype–phenotype link, gene co-option, venom, snake.**

Introduction

Phenotypes are heavily influenced by the recruitment and modification of novel genes. At the initial moment of recruitment, genes may be selected over others with similar function due to subtle characteristics that make them prone to evolve a new function. Gene duplication facilitates the acquisition of new functions by either providing a new copy to freely neofunctionalize or by allowing genes with a latent new function to escape from the adaptive conflict that once prevented its emergence (Hughes 1994; True and Carroll 2002). In both cases, once the phenotype is established, it is counterintuitive to imagine a nonorthologous gene product assuming the function of another and driving this phenotype. Instances where similar biochemical phenotypes could be explicitly

associated with distinct genes rarely occur (Matsutani et al. 2014; Wheeler et al. 2015), and such situations generally involve distantly related organisms.

Snake venom has been an excellent model for exploring gene recruitment, duplication, and neofunctionalization (Casewell et al. 2011; Wong and Belov 2012; Vonk et al. 2013; Hargreaves et al. 2014; Junqueira-de-Azevedo et al. 2015). A small number of protein families are responsible for the major toxic activities (Tasoulis and Isbister 2017) and the quantitative balance between them and the modifications in their sequences dictate the differences in the venom phenotypes among snake taxa (Mackessy 2010). Proteinases are among the major constituents of viper venoms (Fox and Serrano 2008; Mackessy 2010), and only two types of them are recognized—snake venom metalloproteinases (SVMPs) and snake venom serine proteinases (SVSPs). Although SVSPs act at specific targets disrupting the regulation of the coagulation cascade and in the equilibrium of blood pressure (Serrano and Maroun 2005), SVMPs act more broadly through the destruction of the extracellular matrix (ECM), thus inducing local hemorrhage (Gutiérrez et al. 2009).

Despite the extensive knowledge regarding the composition and pharmacological properties of front-fanged snake venoms, the toxin repertoire of rear-fanged snakes remains largely neglected (Junqueira-de-Azevedo et al. 2016). Rear-fanged snakes commonly use their venom to predate (Mackessy 2010; Saviola et al. 2014) and sometimes cause mild envenomation in humans (Salomão et al. 2003; Medeiros et al. 2019; Menegucci et al. 2019). These snakes represent the most prominent parcel of snake biodiversity, with the family Dipsadidae being the most diverse in number of species (Uetz 2010). Similar to vipers, Dipsadidae venoms (specially within the Xenodontinae subfamily) have been reported to possess high proteolytic and hemorrhagic effects, and SVMPs have been suggested as the primary agent of their proteolytic activity (Gutiérrez and Sasa 2002; Acosta et al. 2003; Lemoine and Rodríguez-Acosta 2003; Ching et al. 2012). However, we previously described compositionally unique venoms in the family (Ching et al. 2012; Junqueira-de-Azevedo et al. 2016), in which the dominant component was an unknown type of matrix metalloproteinase (MMP)—termed snake venom MMPs (svMMPs). These proteins are related to the Matrixin family of Zn^{2+} proteinases rather than to the Adamalysins, where traditional SVMPs belong. Although both families share conserved catalytic features, each one represents an independent ancient protein lineage within the Metzincin clan of proteinases (Massova et al. 1998; Huxley-Jones et al. 2007; Stöcker et al. 2008).

In this study, we sought to investigate the prevalence of svMMPs across venomous snake lineages, their origin and evolution, and the impacts of their recruitment in the availability of functionally similar SVMPs. To do this, we performed a large-scale omics-screening and estimated expression levels across the advanced snake (Caenophidia) phylogeny. We discovered clades in which svMMPs are the dominant venom component; and we inferred an alternation between SVMP and svMMP expression in the venom glands during the evolution of snakes. Additionally, we recognized an evolutionary path toward a simplified molecular structure, paralleling that of SVMPs (Casewell et al. 2011; Sanz et al. 2012; Brust et al. 2013; Sanz and Calvete 2016). Herein, we discuss how both enzyme families may be functionally equivalent to maintain (or optimize) a similar phenotype (proteolysis). This peculiar race represents an opportunity to understand how genes can evolve and be replaced when a relevant trait is actively maintained.

Results

svMMPs Dominate the Venoms of Three Tribes within Xenodontinae

To investigate the occurrence of matrix metalloproteinases across the snake phylogeny, we sequenced the venom gland transcriptomes of 58 species, some with more than one individual, representing various families of advanced snakes. Based on previous research (Ching et al. 2012; Junqueira-de-Azevedo et al. 2016), we focused our investigation on the subfamily Xenodontinae (Dipsadidae). The transcriptomes of individual samples were de novo assembled and screened against a compiled list of MMP-9-related sequences.

We discovered MMP sequences in most of the transcriptomes, albeit at contrasting expression levels (fig. 1 and supplementary table S1, Supplementary Material online). MMP coding transcripts were highly expressed only in the venom glands from three tribes (groups of closely related genera) within Xenodontinae (Dipsadidae): Tachymenini, Xenodontini, and Conophiini. The first two tribes include the species in which svMMPs were first reported (i.e., *Thamnodynastes strigatus* and *Erythrolamprus miliaris*, respectively) as well as many additional taxa. The third tribe (Conophiini), whose venom has never been characterized, contains the species *Conophis lineatus* tested here.

The average expression level of MMPs in species from these three tribes varied from 5.8% to 72.1% of whole transcriptomes, which is comparable with SVMP expression in viperids (Junqueira-de-Azevedo et al. 2015). MMPs correspond to the predominant product coded in the transcriptomes of the genera *Thamnodynastes*, *Tomodon*, *Ptychophis*, and *Gomesophis* (Tachymenini), *Erythrolamprus* and *Lygophis* (Xenodontini), and *Conophis* (Conophiini), indicating that MMPs should be relevant components of their venoms. We tested the secreted venom of 17 representative species with a high abundance of MMP transcripts for the presence of the respective proteins by MS/MS spectrometry. The proteomic analysis unequivocally confirmed the occurrence of svMMPs as abundant proteins within all venoms tested (blue cross-marked tips on fig. 1), as demonstrated by the high proportion of the MS/MS spectral counts for peptides covering most of the mature enzymes (supplementary table S2 and fig. S1, Supplementary Material online).

Certain species within other Xenodontinae tribes exhibited much lower expression levels of MMP-9-like transcripts (fig. 1), and, in some of these venoms, endogenous MMP-9-derived peptides could be identified (yellow cross-marked tips on fig. 1). This low expression is in line with previous research,

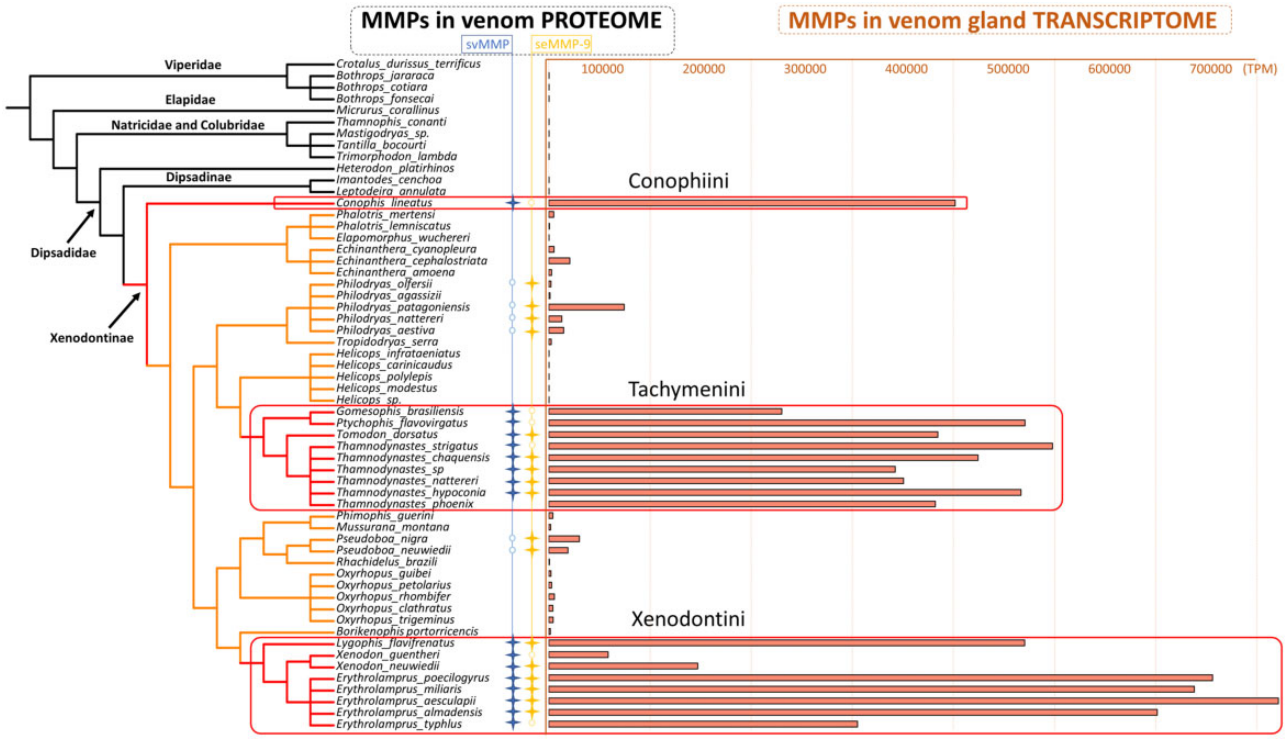

Fig. 1. Occurrence of MMP-9-like transcripts and proteins across snake phylogeny. In the right panel, the red bars represent the average transcripts per million (TPM) value for total MMPs (svMMPs and/or seMMP-9) for each sampled species. The three tribes with MMPs in high abundance are highlighted within a red box; in these cases, the expression values correspond to the total amount of svMMPs and seMMP-9s detected together in the initial screening. For the species outside the boxes, no svMMP sequence was identified and the expression values correspond to seMMP-9 solely. In the columns, blue and yellow crosses indicate the presence in the venom of svMMP and seMMP-9, respectively, detected by proteomics. An open circle indicates that the venom was tested but the protein was not identified (for all other species without marks, the venom was not tested).

which demonstrated that MMPs are minor components of other venoms (Ching et al. 2006; Komori et al. 2006; Campos et al. 2016). The mRNA expression levels in other snake families (outside Dipsadidae) showed even lower expression of MMP coding transcripts ($\leq$0.1%) (fig. 1 and supplementary table S1, Supplementary Material online). These results indicate that MMP-like proteins could be present at basal levels in various snake venoms, but they have a uniquely high expression in species of three tribes within the subfamily Xenodontinae.

Different Types of MMPs Exist in Snake Venoms

A thorough revision of the newly discovered sequences from all species (supplementary table S3, Supplementary Material online) indicates that there are two distinct cases of MMPs produced in snake venom glands and secreted in the venom (fig. 2A). The first case is a bona fide MMP-9 with the same domain organization of MMP-9s commonly found in vertebrates. This protein corresponds to the endogenous MMP-9 of the snakes, usually secreted in several other snake tissues (fig. 3A), and mainly used for their endophysiological functions. Hereafter, we refer to this protein as snake-endogenous MMP-9 (seMMP-9). The second case is an atypical MMP that lacks one or more domains present in MMP-9. We refer to these proteins as the true snake venom matrix metalloproteinases (svMMPs), since this configuration corresponds to the protein arrangement that we previously identified in two species (Ching et al. 2012; Junqueira-de-Azevedo et al. 2016). In svMMPs, a stop codon terminates a short and variable C-terminal tail after the catalytic domain (fig. 4B and supplementary figs. S1 and S4–S6, Supplementary Material online), leading to the absence of the hemopexin domain, a key structural difference from seMMP-9.

Further within the svMMPs, we recognize two protein subtypes: svMMP-A, characterized by the presence of three fibronectin type 2 (FN2) domain repeats placed before the Zn^{2+} binding site, as observed in MMP-9 (Gelatinase B) (Elkins et al. 2002); and svMMP-B, with an uninterrupted catalytic domain (without FN2). Both forms could be unequivocally distinguished in the venom proteome by detected peptides crossing the border between the catalytic and fibronectin domains (in svMMPA) and specific peptides of the catalytic domain that span over the region where the fibronectin domain is missing (in svMMP-B), as shown in supplementary fig. S1, Supplementary Material online. The proportions of svMMP-A and svMMP-B are different among the three tribes (fig. 2B and supplementary table S4, Supplementary Material online). Although xenodontins generally show variable amounts of svMMP-A and svMMP-B, tachymenins have more svMMP-B than svMMP-A, and conophiins possess only svMMP-A.

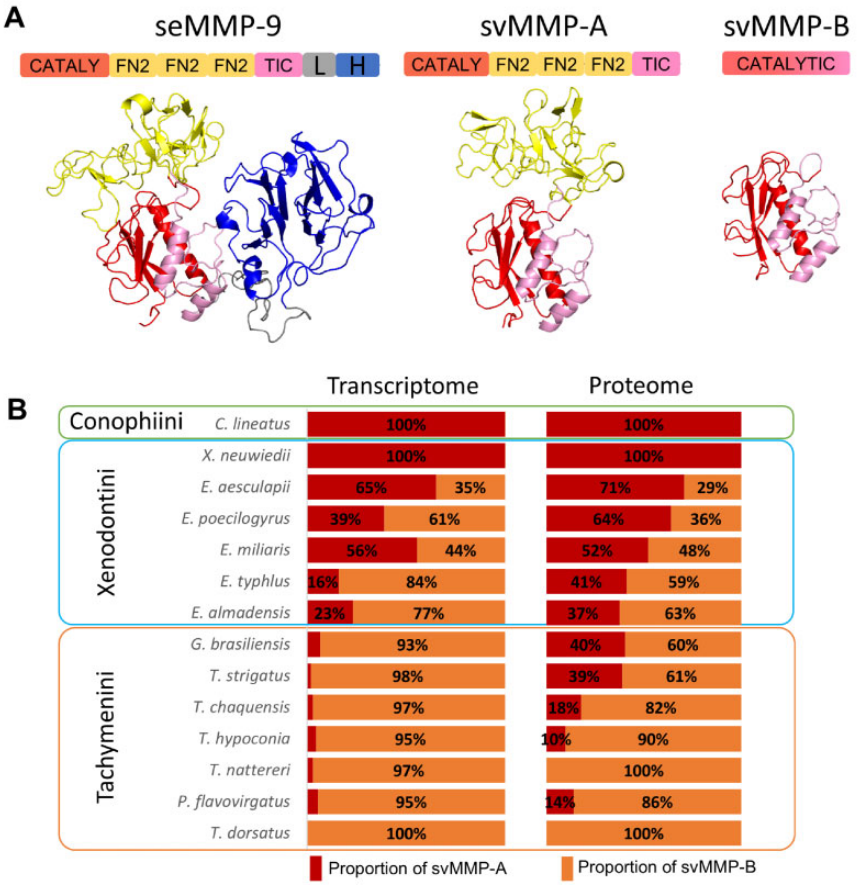

FIG. 2. The three types of MMPs found in snake venoms: one endogenous MMP-9 (seMMP-9) and two venom-specific MMPs (svMMP-A and svMMP-B). (*A*) Structural comparison between the three different types of MMPs identified in snake venom glands. Domains described in the linear model are: CATALY (first portion of catalytic), FN2 (fibronectin), TIC (second part of catalytic), L (linker), and H (hemopexin). 3D structures of seMMP-9, svMMP-A, and svMMP-B were modeled by homology using the SWISS-MODEL online platform on default parameters. The templates used were 1L6J (Elkins et al. 2002) for svMMP-A and B, and theoretical model 1LKG (https://www.modelarchive.org/doi/10.5452/ma-coftn, last accessed August 3, 2020) for seMMP-9. (*B*) Proportion of both types of svMMP, in both the venom gland transcriptome and the venom proteome, for one species of Conophiini, six species of Xenodontini, and seven species of Tachymenini.

In summary, seMMP-9 is a classic vertebrate MMP-9, svMMP-A is a seMMP-9 lacking the hemopexin domain, and svMMP-B has a more compact arrangement, similar to that observed in MMP-7 (Matrilysin) (fig. 2*A*). More importantly, although seMMP-9 is generally expressed at low levels in the venom glands and other tissues of many snake species across the phylogeny, svMMPs are highly expressed in the venom glands of three tribes within Xenodontinae (Tachymenini, Xenodontini, and Conophiini) (fig. 1). svMMPs are not detected, or detected only as a minor constituent, in other tissues than the venom gland from the same svMMP-positive individuals (fig. 3*A*); nor are they detected in the venom glands of species outside these three groups, even in those with high seMMP-9 content (fig. 3*B*).

svMMP Arose More Than Once from Endogenous MMP-9 in Xenodontinae

To trace the evolutionary origin of svMMPs and their relationships to seMMP-9 and other MMPs, a phylogenetic analysis of all curated sequences for each type of protein (svMMP-A, svMMP-B, seMMP-9, other MMP-9, and other MMP classes) was performed in two ways: 1) using DNA coding regions for complete proteins (supplementary fig. S2, Supplementary Material online); and 2) using the core catalytic domain alone (supplementary fig. S3, Supplementary Material online).

The results, summarized in the schematic tree of figure 4*A*, were consistent between analyses and revealed three key findings. 1) svMMPs and seMMP-9s form a strongly supported monophyletic group nested within the MMP-9s of other vertebrates, suggesting svMMPs originated from MMP-9 rather than another MMP family. 2) svMMPs evolved independently in the three tribes of Xenodontinae in which they were detected. That is, svMMPs and seMMP-9s from Xenodontini, Tachymennini, and Conophiini each form monophyletic groups, not related to each other. Although no seMMP-9s could be retrieved from *C. lineatus*, they are likely to be found with increased sampling or sequencing. Given these results, we conclude that there is not a single common ancestral gene for svMMPs, and instead, these proteins have evolved independently multiple times from seMMP-9. 3) svMMP-B is derived from a svMMP-A rather than from a seMMP-9 directly. This is observed in the two tribes where svMMP-Bs were found, suggesting two independent transitions from svMMP-A to svMMP-B, either by a

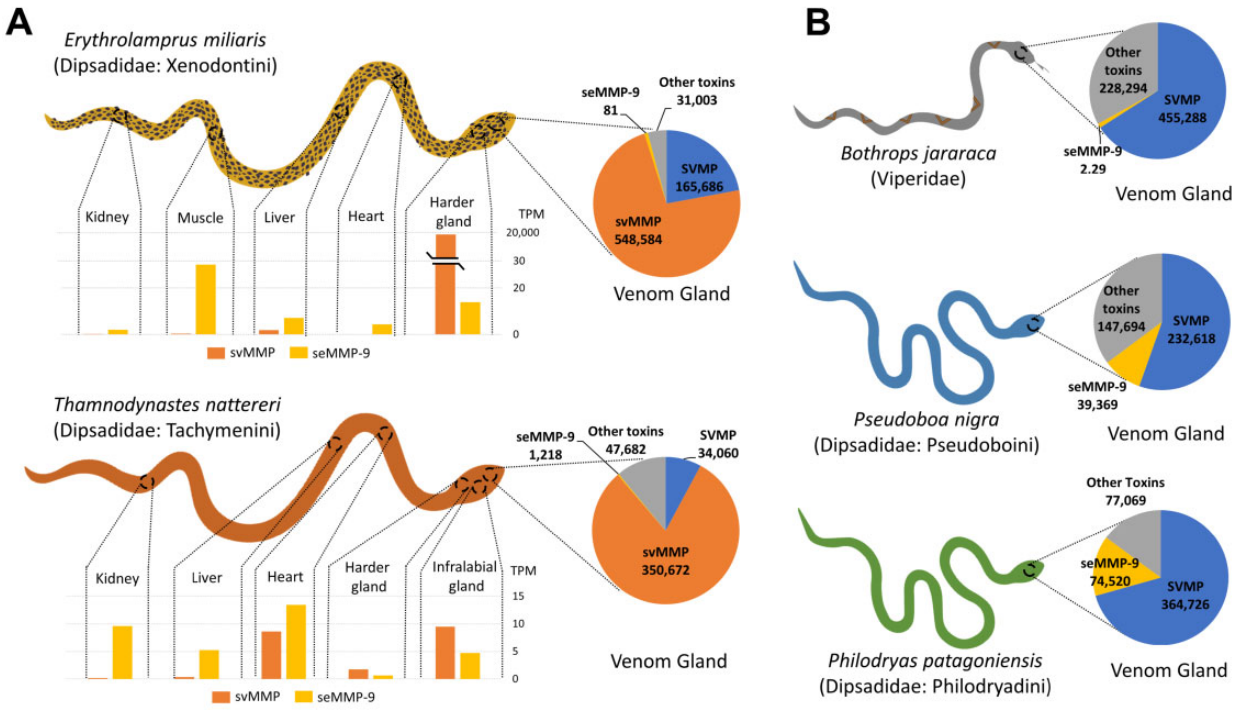

Fig. 3. Expression of putative toxins and MMPs in snake transcriptomes. (*A*) Expression levels of seMMP-9s and svMMPs in the transcriptomes of various tissues of representative species of svMMP-positive snakes. Note that svMMPs (orange bars and slices) may be detected in restricted amounts in other tissues but are only highly expressed in venom glands. (*B*) Comparison of the putative toxin proportions in the venom gland transcriptomes of representative species of svMMP-negative snakes. Note an increased proportion of seMMP-9 (yellow slice) in Dipsadidae. All expression values are expressed in TPM.

genomic deletion of the exons corresponding to the fibronectin domain (fig. 4C) or by a change in the splicing site leading to an intronization of the exons.

In addition to the phylogenetic evidence for independent recruitments of svMMPs from seMMP-9, we observed some traceable genetic differences between svMMPs from different tribes. Alignments of whole svMMP contigs (i.e., containing the 3′-UTR region) from the three tribes revealed different patterns in the positioning of the stop codon among Xenodontini, Tachymenini, and Conophiini (fig. 4B and supplementary figs. S4–S6, Supplementary Material online). In Xenodontini, a deletion from the end of the linker to the final quarter of the hemopexin domain resulted in the ablation of this domain. The original stop codon of seMMP-9 was used in most isoforms, whereas other isoforms evolved upstream stop codons likely by mutation. In Tachymenini, the deletion is longer and advances further into the 3′-UTR region. Thus, changing the open read frame, resulting in a downstream stop codon in all isoforms. Finally, in Conophiini, the region deleted is shorter than in the others and completely confined within the hemopexin domain. The stop codon in this case is uniquely positioned inside the region previously coding for the hemopexin domain. These differences are indicative of distinct deletional and mutational events in each tribe that led to a suppression of the hemopexin domain, further supporting the hypothesis of independent origins in the family.

Overall, our data indicate that svMMPs were recruited to venom three times during the evolution of xenodontines. If the evolution of svMMPs was due to a single recruitment event, we would expect all seMMP-9s to cluster together in the tree with one single-derived clade containing svMMPs from all three tribes, and all of them having similar C-terminals. Another explanation, however, could be a single early origin at the base of Xenodontinae, since Conophiini is the sister taxon to the remaining tribes (Zaher et al. 2018), followed by several independent losses of svMMPs. However, this explanation is less parsimonious and not supported by the genetic structure of the C-terminal region, which indicates independent solutions for abolishing the hemopexin domain in each tribe (fig. 4B).

A Parallel Trend in the Evolution of Metalloproteinases toward a Simplified Structure

The successive origins of svMMP-A from seMMP-9, and svMMP-B from svMMP-A (fig. 4C), involved the loss of the hemopexin and fibronectin domains, respectively. Three sequential exons code the three FN2 domains in seMMP-9, and the C-terminal hemopexin domain is encompassed in other 3′ exons exclusively coding for this domain (as observed in *Python bivittatus* gene XM_007436046.3). Therefore, the loss of these domains in the transitions from seMMP-9 to svMMP-A, and subsequently, to svMMP-B requires relatively simple gene deletions that would not compromise other parts of the protein. In fact, exon shuffling—along with gene duplication—is the proposed mechanism for the ancestral diversification of MMPs and the subsequent generation of shorter MMP types (Fanjul-Fernández et al. 2010).

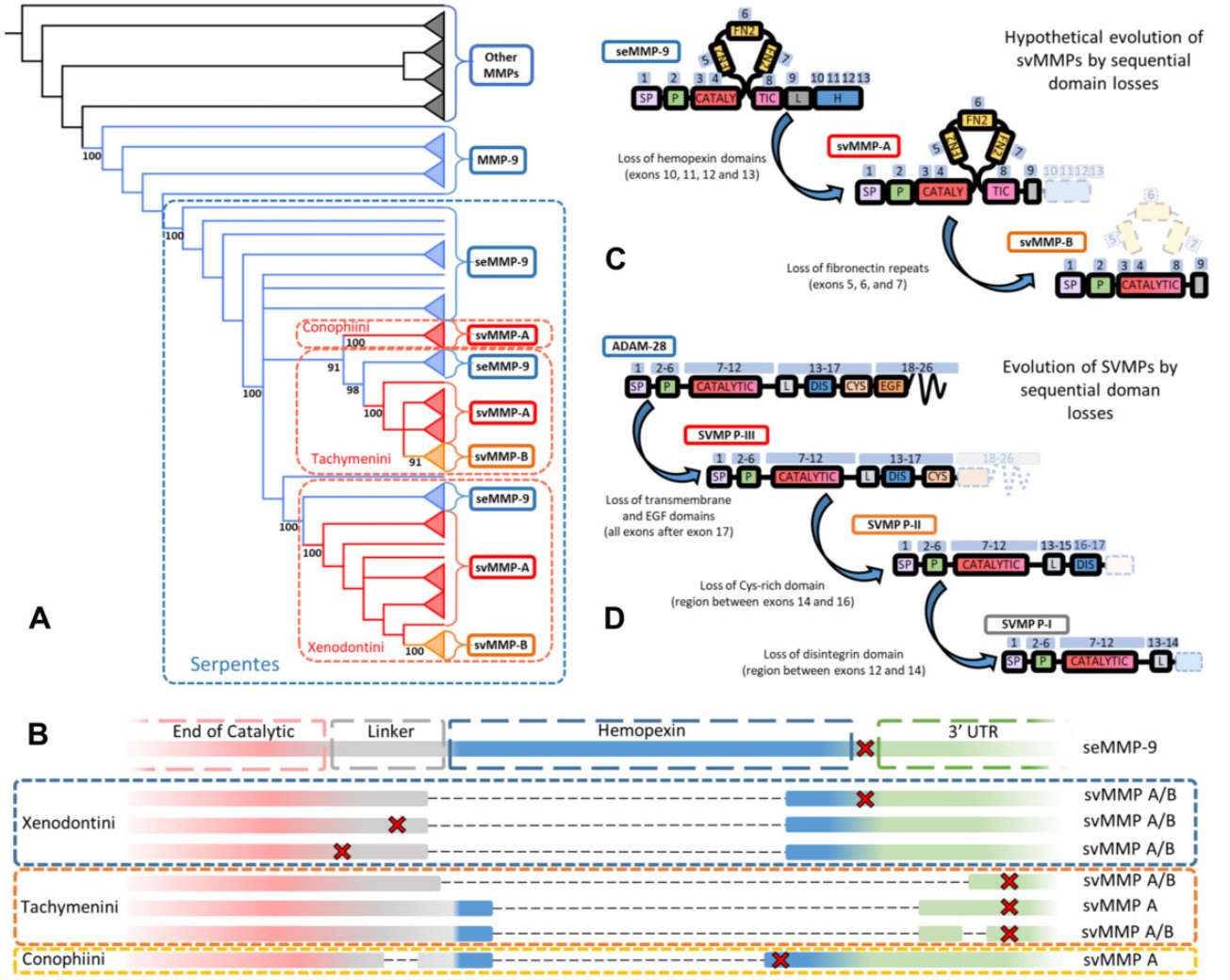

FIG. 4. Evolution of svMMPs. (*A*) Schematic ML tree of vertebrate MMPs, seMMP-9s, and svMMPs showing the phylogenetic relationships between the three types of proteins (the complete phylogenetic trees are supplied in supplementary figs. S2 and S3, Supplementary Material online). Dashed boxes indicate relevant taxa. Independent origins could be assumed at least for Xenodontini and Tachymenini, as svMMPs from these groups cluster with their endogenous seMMP-9. Bootstrap support values shown for relevant nodes were estimated from the full-length CDS analysis. (*B*) Schematic representation of the C-terminal part and 3′-UTRs of svMMPs from the three tribes aligned against their endogenous counterparts (the complete alignments are supplied in supplementary figs. S4–S6, Supplementary Material online). Red marks indicate the position of the stop codon in each protein. The svMMP type (*A* or *B*) in which each pattern was found is specified to the right. (*C*) Proposed evolution of svMMPs through sequential domain losses. Numbered boxes over the protein domains of seMMP-9 represent the exons coding them, based on the reported MMP-9 gene from *Python bivittatus* (Gene ID: 103064710). For svMMP-A and svMMP-B, the numbered boxes represent the exons expected to code for the domains, numbered according to the seMMP-9 gene. (*D*) Sequential domain losses via exon shuffling previously suggested to have occurred during SVMP evolution from ADAM28 (Sanz et al. 2012; Giorgianni et al. 2020), indicating a similar trend toward a simplified derived structure.

Interestingly, this pattern of structural evolution is similar to that documented for SVMPs (fig. 4*D*), where the ancestral endophysiological gene ADAM28 (Sanz et al. 2012; Sanz and Calvete 2016; Giorgianni et al. 2020) gave birth to a shorter form in venom, known as SVMP P-III precursor, by losing EGF-like and transmembrane regions (Sanz and Calvete 2016). Uniquely within the family Viperidae, an ancient gene duplication led to the appearance of P-II precursor lacking the cysteine-rich domain, and, ultimately, several lineage-specific duplications caused the appearance of SVMP P-I precursor that lost the disintegrin domain in different genera (Sanz et al. 2012; Brust et al. 2013). In the genome, this sequence of events was also associated with exon deletions resulting in a shortening of the proteins (Sanz et al. 2012; Brust et al. 2013; Sanz and Calvete 2016; Giorgianni et al. 2020). Moreover, expansion and retraction in SVMP gene numbers were demonstrated among related species in *Crotalus* (Viperidae), with evidence that recombination between genes of different classes evolved mixed isoforms (Giorgianni et al. 2020).

We thus recognize a parallel trend for the evolution of these two types of venom metalloproteinases (svMMPs and SVMPs) toward simpler functional structures. In both cases, ancestral endophysiological proteins with multiple domains were

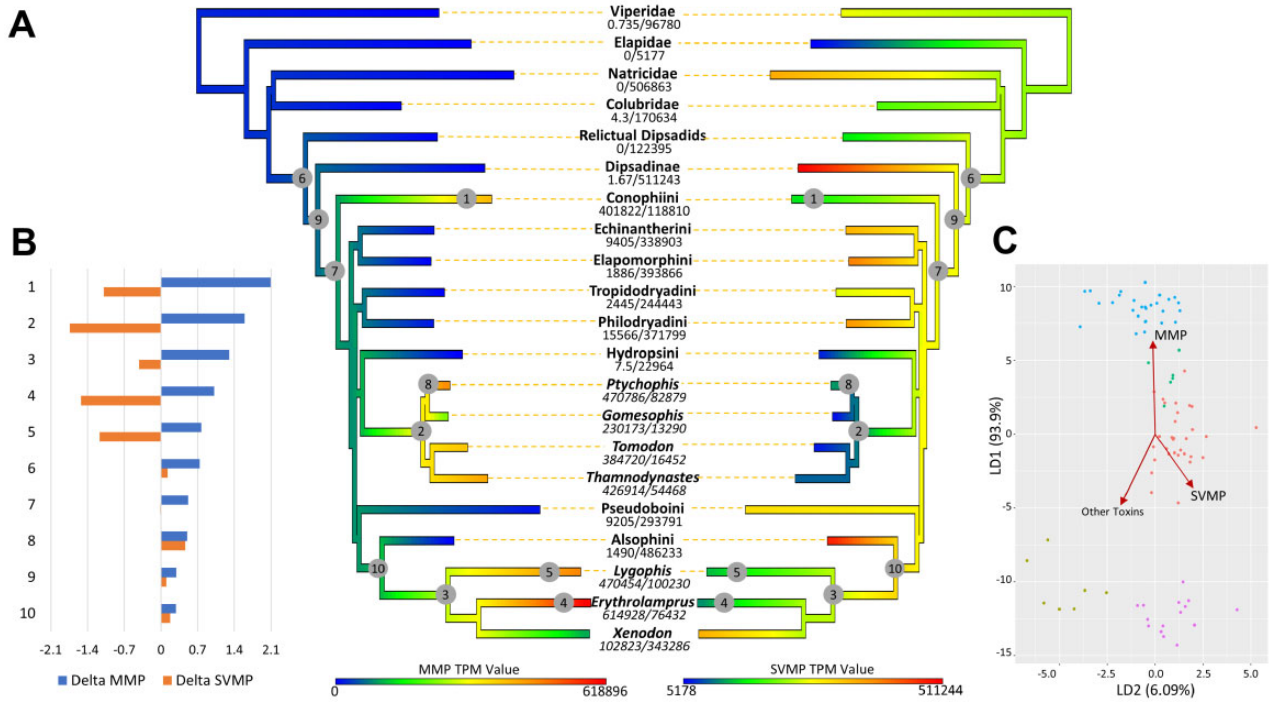

FIG. 5. The balance of venom metalloproteinases during evolution. (*A*) Ancestral state reconstruction of both MMP (left) and SVMP (right) across several snake groups. Blue branches indicate low expression levels, whereas branches tending to red indicate high expression levels. Mean expression values in TPM for MMPs/SVMPs are shown under each group name. Numbered lineages represent nodes or terminals with the highest MMP expression fold-change in comparison to the previous node. (*B*) The rank of top ten lineages according to the log2 fold-change of MMP expression levels (blue bars) in comparison to that of the preceding node; SVMP fold-change is also shown (orange bars). (*C*) LDA analysis showing how changes in MMP and SVMP expression levels are the major drivers in the differentiation of venom gland expression patterns. svMMP and SVMP-positive taxa group in five different clusters: 1) svMMP-rich venoms (blue): constituted by species from Tachymenini, most Xenodontini and Conophiini; 2) svMMP-rich venoms with high level of SVMP expression (green): species of Xenodon (Xenodontini); 3) SVMP-rich venoms with some marked expression levels of svMMP or seMMP-9 (red): species from other tribes of Xenodontinae; 4) SVMP-rich venoms (purple): species from other Dipsadidae subfamilies, Natricidae, some Colubridae and species of *Bothrops*; and 5) Venoms rich in other toxins (pistachio): species from Elapidae, Hydropsini, *Mastigodryas*, and *Crotalus*.

recruited and, via subsequent domain losses associated with exon shuffling, simplified until their mature products comprised only the catalytic domain (as seen in the most-derived venom forms). Nevertheless, all forms continued to exert the proteolytic effects in the venom, as further discussed.

A Substitution of SVMP per svMMP Maintains the Proteolytic Phenotype of Xenodontinae Venoms

Although the recruitment of SVMPs is thought to have happened at the base of advanced snakes, as these enzymes are found across snake families (Tasoulis and Isbister 2017), we demonstrate here that svMMP recruitment was a more derived event, occurring in the subfamily Xenodontinae. But, how did the appearance of this second type of proteinase impact the availability of SVMP in Xenodontinae venoms?

We observed that even in the species where svMMPs are highly abundant, SVMPs still coexist in the venom, but generally at reduced levels (fig. 3A). Nevertheless, using ancestral state reconstruction, we inferred changes in the expression of SVMPs and MMPs along the phylogeny (fig. 5A and supplementary table S5, Supplementary Material online). This analysis reveals that the ancestor of the Xendontinae, and possibly the ancestor of the whole family (Dipsadidae), may have had elevated expression of seMMP-9 in their venom glands (predicted TPM level ~70,000), that is, prior to the appearance of svMMPs. These hypothesized ancestors (lineages 6, 9, and 7 in fig. 5A left) show the first increases in predicted TPM values for MMPs, representing a time interval when this protein type, still in its endogenous configuration (seMMP-9), potentially became evolutionarily relevant to venom function. Accordingly, more substantial increases in the expression of MMP-like proteins are observed only in the three tribes where svMMP was detected (lineages 1, 2, and 3), indicating that the abundance of MMPs in these venoms is associated with the appearance of the derived form (svMMP).

Conversely, the ancestral state reconstruction of SVMP expression levels (fig. 5A right) demonstrated that noteworthy reductions in SVMP occurred within the three tribes where svMMPs evolved (as seen in the lineages 1, 2, and 3). In the tribe Tachymenini (node 2), all genera have low expression of SVMP (<10% of total transcriptome). In the genus *Tomodon*, for example, SVMP represents <5% of the transcriptome and the average proportion of MMP/SVMP is 23:1. Within the tribe Xenodontini, the genera *Erythrolamprus* and *Lygophis* have an 8:1 and 5:1 ratio of MMP/SVMP, respectively. In all other tribes of Xenodontinae (except Hydropsini), SVMP is highly expressed and corresponds to the dominant type of transcript

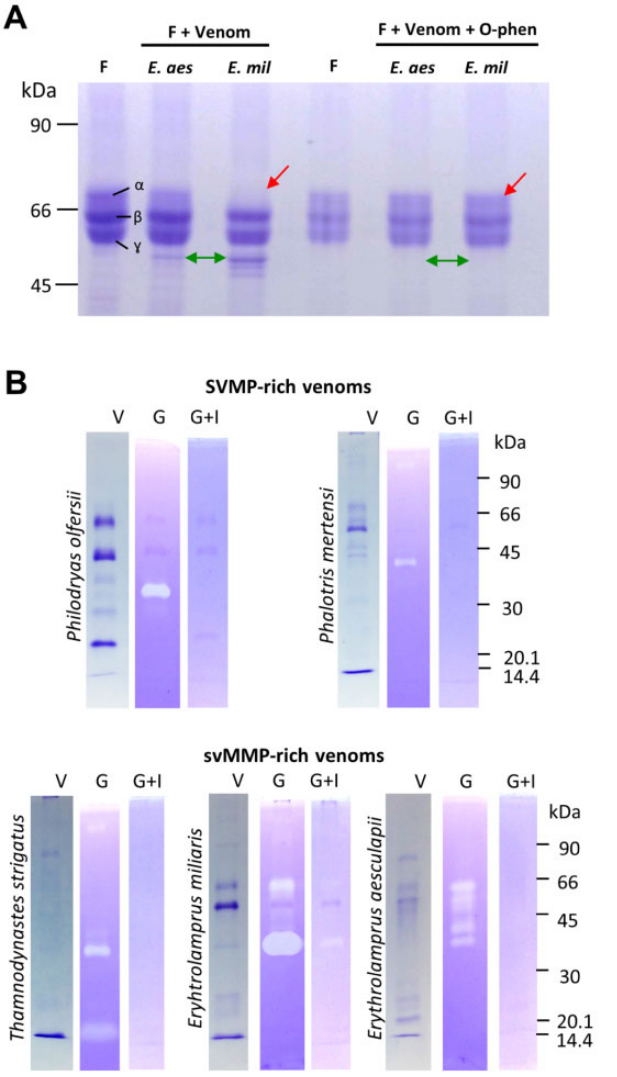

FIG. 6. Metalloproteolytic activities of Xenodontinae venoms. (*A*) Fibrinogenolytic activity of *E. aesculapii* (*E. aes*) and *E. miliaris* (*E. mil*) venoms over fibrinogen (F) in the absence or presence of the metalloproteinase inhibitor O-phenanthroline (O-phen). Fibrinogen alpha, beta, and gamma chains are indicated in the first lane. Single-headed arrows indicate bands of alpha chain that were visibly degraded when the inhibitor was not added. Double-headed arrows indicate degradation products, which are absent when the inhibitor was added. (*B*) Gelatinolytic activity assay of SVMP-rich venoms (upper panel) or svMMP-rich venoms (lower panel). V, Venoms from the species indicated on the side run in SDS–PAGE (without gelatin); G, Gelatinolytic activities of the same venoms run in SDS–PAGE polymerized with gelatin; G + I, Gelatinolytic activities of the same venoms run in SDS–PAGE polymerized with gelatin, and in the presence metalloproteinase inhibitor EDTA.

produced in the venom glands. Moreover, the deep nodes of Dipsadidae phylogeny (6, 7, and 9), which demonstrated an increase in MMP expression, did not show a reduction in SVMP levels, suggesting that SVMP decay succeeded, and not preceded, the appearance of high levels of MMPs in snake venoms.

Moreover, after correcting for the negative correlation bias of compositional data sets, the expression levels in Tachymenini and Xenodontini are still significantly different when compared with other tribes of Dipsadidae ($P < 1E\text{-}3$ in all pairwise comparisons) (supplementary table S6, Supplementary Material online). The post hoc analyses of expression levels indicate that the values of expression for SVMPs are significantly lower in these two tribes when compared with other dipsadids and are higher for MMPs ($P < 1E\text{-}3$ in all pairwise comparisons). Although Tachymenini and Xenodontini are significantly different from all other tribes, they do not differ statistically among themselves ($P = 0.15$ for SVMPs; $P = 0.194$ for MMP, and $P = 0.895$ for other toxins) (supplementary table S6, Supplementary Material online).

In summary, MMP expression levels negatively correlate to SVMP levels among the evaluated species (fig. 5*B*), even after the statistical transformations (supplementary fig. S7, Supplementary Material online). In addition, a linear discriminant analysis of transcriptomic venom composition (fig. 5*C*) corroborates the major role of metalloproteinases in differentiating the contrasting venom expression profiles found among snake taxa (LD1 represents 93.9% of observed variance). Besides, it indicates that svMMP-positive taxa cluster together and separate themselves from SVMP-positive taxa. These data also indicate that metalloproteinases in general (MMP or SVMP) are always the predominant venom toxin in most species of Dipsadidae, with the sum of MMP and SVMP kept consistently high (between 25% and 60% of total transcriptome). Accordingly, proteome data obtained here and elsewhere consistently reveal metalloproteinase as the major protein type secreted to the venoms of the group (Junqueira-de-Azevedo et al. 2016).

To functionally evaluate the contribution of metalloproteinases to the proteolytic phenotype of Xenodontinae venoms, we further examined metalloproteinase activity in representative species of svMMP-rich and SVMP-rich venoms (fig. 6). The high svMMP venoms of *E. miliaris* and *E. aesculapii* were able to cleave the fibrinogen alpha chain and generate degradation products (fig. 6*A*). This cleavage was completely inhibited by ortho-phenanthroline, a classic metalloproteinase inhibitor. In a comparative gelatin zymography (fig. 6*B*), all tested venoms with predicted metalloproteinases (svMMP or SVMP) showed marked proteolytic bands, and these bands were entirely absent when treating the venom with the metal chelating EDTA, a metalloproteinase inhibitor. It is then possible to conclude that the proteolytic character, and more specifically, the "metalloproteolytic" character, is maintained within Xenodontinae venoms, independent of which type of enzyme (svMMP or SVMP) is dominating those venoms. However, to obtain a complete understanding of the contribution of each enzyme type (SVMP or svMMP) in svMMP-rich venoms, further experiments would be required to correctly determine the exact enzymes responsible for the observed strong proteolysis.

Discussion

In this study, we demonstrate an instance of parallel evolution between two distantly related proteinases, leading to a possible functional replacement of one with the other in certain

groups of snakes. Considering the number of species in the three svMMP-positive tribes (Uetz 2010), we predict that this biochemical innovation could be present in the venoms of more than 100 species throughout the New World. Although the effects of Dipsadidae venoms are usually mild, the recognition of svMMPs as a major venom component may have treatment implications in the occasional human envenomation. For example, the antigenic determinants of svMMPs are certainly different from those of SVMPs present in the venoms used for antisera production. We can anticipate that antivenoms regularly used for treating envenomation caused by the proteolytic venoms of viperids, and occasionally applied for some SVMP-rich rear-fanged snakes (Prado-Franceschi and Hyslop 2002; Rocha et al. 2006), will likely not help the patients affected by svMMP abundant venoms.

We hypothesized here that svMMPs arose from an MMP-9 ancestor through independent gene duplications. Endogenous MMP-9s exert various functions in vertebrates through their proteolytic action on the ECM or their regulation of cytokines (Heikinheimo and Salo 1995; Chin and Werb 1997; Pilcher et al. 1999; Le et al. 2007). MMP-9 is typically expressed in a variety of tissues and cells and secreted into the extracellular space (Vandooren et al. 2013), which is markedly different from many other types of MMPs that remain attached to the cellular membrane through a C-terminal transmembrane domain (Fanjul-Fernández et al. 2010). The other ancillary noncatalytic domains present in MMP-9, and lost in svMMPs, are associated with the recognition of their natural substrates, guidance to the cleavage promoted by the catalytic domain, and interactions with inhibitors (Murphy et al. 1994; Shipley et al. 1996; Chung et al. 2004; Rosenblum et al. 2007; Vandooren et al. 2013). Thus, the co-option of a broadly expressed and secreted member of MMP family to venom is a reasonable evolutionary hypothesis. However, how the loss of ancillary domains impacted the function of svMMPs in vivo is a more open question.

A clue to understand these domain losses could be in the fact that most snake toxins, like three-finger toxins, C-type lectins, SVMPs, among others, evolved a more simplified structure than their physiological counterparts (Ogawa et al. 2005; Kini and Doley 2010; Casewell et al. 2011; Sanz and Calvete 2016; Giorgianni et al. 2020). In the latter example, all derived forms (SVMP PI, PII, PIII) exert the proteolytic effects in venom, whereas the ancestral ADAM28 is maintained in the genome playing the physiological function. Thus, the evolution of svMMPs from a multidomain ancestral MMP-9 through successive domain losses not only agrees with a common tendency toward a simpler functional structure in snake toxins but it also represents a parallel trend for these two types of metalloproteinases.

The parallel between SVMPs and svMMPs extends to the observation that svMMP appearance is accompanied by a reduction in all other components, including SVMPs. The resulting low amount of SVMPs might potentially compromise the proteolytic character of the venom in these three tribes. However, since both enzymes can promote the destruction of ECM, perhaps svMMPs have assumed a similar function in these cases. The effects of Xenodontinae venoms are not as well characterized as the ones from more venomous taxa, but the available studies are consistent in demonstrating that venoms of both svMMP-rich and SVMP-rich groups have high proteolytic action over gelatin, fibrinogen, and synthetic substrates, which was associated with metalloproteinases in a Ca^{2+} dependent manner (de Araújo and dos Santos 1997; Da Rocha and Furtado 2007; Zelanis et al. 2010; Ching et al. 2012; Sánchez et al. 2019). In addition to those previous evaluations, the results of in vitro tests performed here confirm the general proteolytic character of both venom types. The fact that proteolysis is equally present and relevant in SVMP-rich and svMMP-rich venoms indicates a certain functional redundancy of both enzymes, making the former, which came first, more easily replaceable if the right pressures favor the latter. Interestingly, it was recently shown that, within Viperidae, SVMP has also taken various evolutionary paths, from continuing to be the predominant component in the Viperinae subfamily to becoming less present in the neurotoxic venoms of some Crotalinae species (Barua and Mikheyev 2019).

MMPs and SVMPs are classified, respectively, in families M10 and M12 of the same MA clan in the Merops database (Rawlings et al. 2009). Each Merops clan is assumed to have a single evolutionary origin, but origins and time of divergence of families within the clan are more challenging to determine. Enzymes from the MA clan were not listed among typical nonhomologous isofunctional enzymes (NISEs) (Omelchenko et al. 2010), although this concept depends on the extent that conservation is recognized. Nevertheless, each family indeed represents long independent protein lineages within the MA clan, since they are present as different gene products at least from the base of Deuterostomia (Massova et al. 1998; Huxley-Jones et al. 2007; Stöcker et al. 2008; Fanjul-Fernández et al. 2010). Congregated as metalloenzymes, however, they share conserved Zn^{2+} binding pockets, likely with similar entatic states (Vallee and Williams 1968), which may facilitate an interchangeable function between them. Provided that ECM degradation capacity was maintained, or even optimized by the loss of ancillary domains in later events, the adaptive conflict (Hughes 1994; True and Carroll 2002), that is, the transformation of the original function of MMP-9 into a new venom function similar to that of SVMPs, should not be particularly constraining. Thus, the substitution suggested here may not be associated with a classic case of enzyme convergence, defined by Galperin and Koonin (2012) as involving members of distinct superfamilies being recruited to catalyze the same metabolic reaction; however, it could be interpreted as a particular case of phenotypic maintenance through the selection of different enzymes. This kind of interchange is more likely for recently diverging homologous enzymes with high similarity; for example, the substitution of ERK2 by ERK1 in Squamate reptiles (Buscà et al. 2015). However, it is less expected for cases involving distantly related sequences such as the two protein families discussed here.

Beyond the possibility of a functional replacement, the acquisition of svMMPs may have contributed novel activities to the venom of these tribes. The snakes possessing svMMPs

are very diverse, with species adapted to fossorial, aquatic, and arboreal habitats, living in environments as distinct as dry and wetlands over the New World, and feeding on different types of vertebrate and invertebrate prey. It is then challenging to imagine a unique role for svMMPs that attends the different demands they have. However, during snake evolution, proteinases have been selected at least three times to venom: 1) kallikrein-like SVSPs at a Toxicofera ancestral, 2) SVMPs at the base of advanced snakes, and 3) factor X/V-like serine proteinases in the Australian Elapidae lineage (Fry et al. 2008). Through gene duplications and selection, different paralogs of each of these proteinases, specially (1) and (2), specialized into a wide range of substrates and targeted multiple biological pathways. Similarly, a fourth family of venom-secreted proteinase (svMMPs) would represent a new asset that natural selection could optimize over multiple functions in different situations.

In conclusion, the biological contest represented by svMMPs and SVMPs in venoms across multiple clades illustrates how genes can be substituted or complemented in the coordination of a phenotype once their products have sufficient functional overlap and freedom to converge. Uncovering if a functional substitution in fact occurred, and why, will represent an important next step for understanding the processes underlying gene co-option and gene replacement. It further exemplifies the utility of studying snake venoms to explore gene evolution.

Materials and Methods

Collection and Storage of Samples

Specimens from 29 genera and 58 species across five different snake families were collected in different field trips in several localities (ICMBio permits 56576, 57585, and 66597) and deposited at the Herpetology Collection of Instituto Butantan. Their venom was extracted using pilocarpin on sedated individuals as described in previous works (Mackessy et al. 2006). Four days after extraction, their venom glands and other tissues were surgically collected and stored in RNAlater at −80 °C. All protocols were certified by the ethical committee of Instituto Butantan (CEUA No. 4479020217). For the species *Borikenophis portorricensis*, raw transcriptomic data were retrieved from the NCBI Sequence Read Archive (SRA) under the code SRP137035 (Modahl et al. 2018).

RNA Extraction and Analysis

Tissues were pulverized in a Precellys 24 homogenizer and RNA was extracted with TRIZOL (Thermo Fisher Scientific) following the modification of the method described by Chomczynski and Sacchi (1987). Total RNA was quantified by Quant-iTTMRiboGreen RNA reagent and Kit (Thermo Fisher Scientific). Quality control of the extracted RNA was then performed in an Agilent 2100 Bioanalyzer using Agilent RNA 6000 Nano Kit, to verify the integrity of total RNA, through band discrimination corresponding to fractions 18S and 28S of total RNA. All procedures with RNA were made with RNAse-free tubes and filter tips and using water treated with diethylpyrocarbonate (DEPC, Sigma).

Transcriptome Sequencing

cDNA libraries were prepared for each individual sample. About 1 μg of total RNA was used with Illumina TruSeq Stranded RNA HT Kit (Illumina). Fragment size distribution was evaluated by microfluidic gel electrophoresis in the Bioanalyzer device (Agilent 2100), using the Agilent DNA 1000 Kit, according to the manufacturer's protocol. Quantification of each library was then performed by Real-Time PCR using the KAPA SYBR FAST Universal qPCR Kit, according to the manufacturer's protocol, using the StepOnePlusTM Real-Time PCR System. Aliquots of each cDNA library were diluted to a concentration of 2 nM. Next, a pool of all samples, 5μl of each library, was prepared, and the concentration of the pool was again determined by real-time PCR. The cDNA libraries were sequenced on an Illumina HiSeq 1500 System on Rapid Run mode, using a paired-end flowcell for 300 cycles of 2×151 bp.

Screening of MMPs in the Transcriptomes

The screening process began with the transcriptome assembly of the samples in which we checked and removed cross-contamination and then trimmed the sequencing adaptors using TrimGalore (www.bioinformatics.babraham.ac.uk, last accessed August 3, 2020). We merged our reads using the PEAR software (Zhang et al. 2014), taking advantage of the common overlap on the 3′-ends that characterize paired-end short reads (Rokyta et al. 2012), and used those merged reads as an input for our assembly. We run all the assemblies in a standardized way using the Trinity v2.6.6 assembler with a k-mer value of 31, since it is optimized for Illumina paired-end read data (Grabherr et al. 2011; Bankar et al. 2015). We performed BlastN searches using the obtained contigs against a curated database of MMP sequences obtained from our previous works, from GenBank and it was interactively upgraded with newly retrieved sequences from key species. The curation process of MMPs sequences involved the use of various de novo assemblers (NGen, RNASpades, Bridger, Extender, and Trinity) (Rokyta et al. 2012; Chang et al. 2015; Bushmanova et al. 2019) to correctly reassemble full-length ORFs coding for these proteins and to distinguish the various paralogous forms in each species. Manual curation of complete ORFs was performed on the Geneious v.2020.0.5 software, and coverage analysis was done both by manually examining the read mapping profile and by using the ChimeraCheck software (Hofmann et al. 2018). Nonredundant sequences that passed all filters were incorporated into the screening database. The expression of each individual contig was estimated using RSEM (Li and Dewey 2011) by mapping reads from each sample using bowtie2. Expression was estimated in transcripts per million (TPM) (Wagner et al. 2012, 2013). Afterward, MMP-like contigs were identified and their expression values summed in order to obtain an approximate value for MMP participation in each individual transcriptome.

Proteomics: Liquid Chromatography–Tandem Mass Spectrometry

Shotgun liquid chromatography–tandem mass spectrometry analyses of the venoms from 11 species were performed by the Florida State University College of Medicine Translational lab and by the University of São Paulo BIOMASS—Core Facility for Scientific Research-USP (CEFAP-USP), as detailed in supplementary methods, Supplementary Material online. Protein identification of obtained spectra was performed using MASCOT (Matrix Science, London, UK; version 2.6.2) and X! Tandem (The GPM, thegpm.org, last accessed August 3, 2020; version X! Tandem Alanine [2017.2.1.4]) as the search engine with custom-generated FASTA databases containing curated sequences of MMPs for each specimen and translated protein sequences from the assembled transcriptome (Trinity contigs) for the species, as detailed in supplementary methods, Supplementary Material online.

MMPs and SVMPs Comparison

To compare the relationship between MMPs and SVMPs, we performed the same screening approach for SVMPs as we did for svMMPs (described above). With the obtained mean expression values in TPM for both types of proteinases, we performed an ancestral state reconstruction analysis using the default parameters of the *fastAnc* function implemented in the Phytools package (Revell 2012) on the R software using the phylogeny proposed by Zaher et al. (2018) to estimate branch lengths. Then we built a tree using the *contMap* function in Phytools to represent the continuous variations in MMP expression values (TPM) for each node using a color gradient. Afterward, using the TPM values of the screening data for MMPs, SVMPs, and other toxins, we performed a linear discriminant analysis (LDA) and a nonparametric or permutational multivariate analysis of variance (PerMANOVA) to check if variations in the expression of these proteinases and other toxins in the venom gland could cluster species following their venom gland expression patterns. To avoid spurious correlations due to the compositional nature of the transcriptomic data set, before comparing the expression levels among groups of snakes, the proportions of toxins sequenced were transformed by using the *clr* (centered logratio coefficients) and *ilr* (isometric logratio transformations) functions in R (see supplementary methods, Supplementary Material online). MMP and SVMP expression levels were further analyzed in species for which we managed to assemble a seMMP-9 full-length CDS transcript. In those species, contigs with >98% of identity toward their curated seMMP-9 sequence were assigned as seMMP-9-like contigs and thus distinguished from the svMMP contigs. TPM values for seMMP-9s, svMMPs, and SVMPs were then obtained from the RSEM data obtained for these contigs in the initial screening.

Phylogenetic and Evolutionary Analyses

We curated seMMP-9 and svMMP sequences from several snake groups and combined them with MMPs from other vertebrates publicly available. The final nucleotide data set was clustered using CD-HIT (Fu et al. 2012) with a 99% identity threshold and aligned through its corresponding translated amino acid sequences using the MUSCLE algorithm (Edgar 2004), with default parameters, on the Geneious v.2020.1 software. The nucleotide alignment was then manually edited due to the high structural diversity present in MMPs. Phylogenetic tree inference was then carried out using RAxML V.8.2.12 (Stamatakis 2014), ran on the CIPRES platform (http://www.phylo.org/, last accessed August 3, 2020) using the rapid bootstrap algorithm implemented in RAxML (-f a) with 1,000 bootstrap replicates on two sets of alignments, one containing the full-length CDS sequences for each protein and the other consisting of the isolated catalytic domain region. Both alignments were analyzed using GTRGAMMA as evolution model (Stamatakis 2014). Trees were visualized and edited using the iTol online platform (Letunic and Bork 2016). We aligned the original whole contigs of assembled svMMP-9s and seMMPs from all three tribes containing a portion of the 3′-UTR region to look for unique structural patterns on their C-terminals. Alignments were performed using the MUSCLE algorithm (Edgar 2004), with default parameters, using Geneious v.2020.1. Image files for each alignment were plotted (Geneious version 2020.1 created by Biomatters. Available from https://www.geneious.com, last accessed August 3, 2020).

Functional Analyses

The gelatinolytic activities of 0.5μg of each venom in the absence of reducing agent were evaluated on 12.5% SDS–PAGE gel with 2 mg/ml gelatin, in the presence and absence of 20 mM EDTA, as described in supplementary methods, Supplementary Material online. The fibrinogenolytic activities of 1μg of each venom were evaluated over 100μg of bovine fibrinogen in 12.5% SDS–PAGE gel, in the presence or absence of 62.5 mM ortho-phenantroline, as described in supplementary methods, Supplementary Material online.

Data Availability

Curated sequences (CDS and translated proteins) of all seMMP-9, svMMP-A, and svMMP-B generated in this work are available in the supplementary table S3, Supplementary Material online, organized per species.

Supplementary Material

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

Acknowledgments

We thank Giuseppe Puorto (Reception of Venomous Animals of Instituto Butantan) and the team of students and technicians associated to the Laboratório de Coleções Zoológicas (LCZ—Instituto Butantan) for the assistance in the field and in the processing of the biological material used in this study. This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (2016/50127-5, 2013/07467-1, 2017/24498-9, and 2017/24546-3), the National Science Foundation (DEB 1638879

and DEB 1822417), Conselho Nacional de Desenvolvimento Cient¡fico e Tecnológico (309791/2017-0 and 303958/2018-9), and Fundação de Amparo à Pesquisa do Estado do Amazonas-FAPEAM (PRO-ESTADO).

References

Acosta O, Leiva LC, Peichoto ME, Maruñak S, Teibler P, Rey L. 2003. Hemorrhagic activity of the Duvernoy's gland secretion of the xenodontine colubrid *Philodryas patagoniensis* from the north-east region of Argentina. *Toxicon* 41(8):1007–1012.

Bankar KG, Todur VN, Shukla RN, Vasudevan M. 2015. Ameliorated de novo transcriptome assembly using Illumina paired end sequence data with Trinity Assembler. *Genomics Data* 5:352–359.

Barua A, Mikheyev AS. 2019. Many options, few solutions: over 60 My snakes converged on a few optimal venom formulations. *Mol Biol Evol*. 36(9):1964–1974.

Brust A, Sunagar K, Undheim EAB, Vetter I, Yang DC, Casewell NR, Jackson TNW, Koludarov I, Alewood PF, Hodgson WC, et al. 2013. Differential evolution and neofunctionalization of snake venom metalloprotease domains. *Mol Cell Proteomics*. 12(3):651–663.

Buscà R, Christen R, Lovern M, Clifford AM, Yue JX, Goss GG, Pouysségur J, Lenormand P. 2015. ERK1 and ERK2 present functional redundancy in tetrapods despite higher evolution rate of ERK1. *BMC Evol Biol*. 15:179.

Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. 2019. RnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *Gigascience* 8(9):1–13.

Campos PF, Andrade-Silva D, Zelanis A, Leme AFP, Rocha MMT, Menezes MC, Serrano SMT, Junqueira-De-Azevedo IDLM. 2016. Trends in the evolution of snake toxins underscored by an integrative omics approach to profile the venom of the colubrid *Phalotris mertensi*. *Genome Biol Evol*. 8(8):2266–2287.

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(9):2637–2649.

Chang Z, Li G, Liu J, Zhang Y, Ashby C, Liu D, Cramer CL, Huang X. 2015. Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biol*. 16(1):30.

Chin JR, Werb Z. 1997. Matrix metalloproteinases regulate morphogenesis, migration and remodeling of epithelium, tongue skeletal muscle and cartilage in the mandibular arch. *Development* 124(8):1519–1530.

Ching ATC, Paes Leme AF, Zelanis A, Rocha MMT, Furtado MDFD, Silva DA, Trugilho MRO, Da Rocha SLG, Perales J, Ho PL, et al. 2012. Venomics profiling of *Thamnodynastes strigatus* unveils matrix metalloproteinases and other novel proteins recruited to the toxin arsenal of rear-fanged snakes. *J Proteome Res*. 11(2):1152–1162.

Ching ATC, Rocha MMT, Paes Leme AF, Pimenta DC, de Fátima D. Furtado M, Serrano SMT, Ho PL, Junqueira-de-Azevedo ILM. 2006. Some aspects of the venom proteome of the Colubridae snake *Philodryas olfersii* revealed from a Duvernoy's (venom) gland transcriptome. *FEBS Lett*. 580(18):4417–4422.

Chomzynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal Biochem*. 162(1):156–159.

Chung L, Dinakarpandian D, Yoshida N, Lauer-Fields JL, Fields GB, Visse R, Nagase H. 2004. Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J*. 23(15):3020–3030.

Da Rocha MMT, Furtado MDFD. 2007. Analysis of biological activities from *Philodryas olfersii* (Lichtenstein) and *P. patagoniensis* (Girard) venoms (Serpentes, Colubridae). *Rev Bras Zool*. 24(2):410–418.

de Araújo ME, dos Santos AC. 1997. Cases of human envenoming caused by *Philodryas olfersii* and *Philodryas patagoniensis* (Serpentes: Colubridae). *Rev Soc Bras Med Trop*. 30(6):517–519.

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.

Elkins PA, Ho YS, Smith WW, Janson CA, D'Alessio KJ, McQueney MS, Cummings MD, Romanic AM. 2002. Structure of the C-terminally truncated human ProMMP9, a gelatin-binding matrix metalloproteinase. *Acta Crystallogr D Biol Crystallogr*. 58(7):1182–1192.

Fanjul-Fernández M, Folgueras AR, Cabrera S, López-Otín C. 2010. Matrix metalloproteinases: evolution, gene regulation and functional analysis in mouse models. *Biochim Biophys Acta Mol Cell Res*. 1803(1):3–19.

Fox JW, Serrano SMT. 2008. Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures. *Proteomics* 8(4):909–920.

Fry BG, Scheib H, van der Weerd L, Young B, McNaughtan J, Ramjan SR, Vidal N, Poelmann RE, Norman JA. 2008. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol Cell Proteomics*. 7(2):215–246.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23):3150–3152.

Galperin MY, Koonin EV. 2012. Divergence and convergence in enzyme evolution. *J Biol Chem*. 287(1):21–28.

Giorgianni MW, Dowell NL, Griffin S, Kassner VA, Selegue JE, Carroll SB. 2020. The origin and diversification of a novel protein family in venomous snakes. *Proc Natl Acad Sci U S A*. 117(20):10911–10920.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 29(7):644–652.

Gutiérrez JM, Rucavado A, Chaves F, Díaz C, Escalante T. 2009. Experimental pathology of local tissue damage induced by *Bothrops asper* snake venom. *Toxicon* 54(7):958–975.

Gutiérrez JM, Sasa M. 2002. Bites and envenomations by colubrid snakes in Mexico and Central America. *J Toxicol Toxin Rev*. 21(1-2):105–115.

Hargreaves AD, Swain MT, Hegarty MJ, Logan DW, Mulley JF. 2014. Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. *Genome Biol Evol*. 6(8):2088–2095.

Heikinheimo K, Salo T. 1995. Expression of basement membrane type IV collagen and type IV collagenases (MMP-2 and MMP-9) in human fetal teeth. *J Dent Res*. 74(5):1226–1234.

Hofmann EP, Rautsaw RM, Strickland JL, Holding ML, Hogan MP, Mason AJ, Rokyta DR, Parkinson CL. 2018. Comparative venom-gland transcriptomics and venom proteomics of four Sidewinder rattlesnake (*Crotalus cerastes*) lineages reveal little differential expression despite individual variation. *Sci Rep*. 8(1):15534.

Hughes AL. 1994. The evolution of functionally novel proteins after gene duplication. *Proc Biol Sci*. 256(1346):119–124.

Huxley-Jones J, Clarke TK, Beck C, Toubaris G, Robertson DL, Boot-Handford RP. 2007. The evolution of the vertebrate metzincins; Insights from *Ciona intestinalis* and *Danio rerio*. *BMC Evol Biol*. 7(1):63.

Junqueira-de-Azevedo ILM, 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(3):754–766.

Junqueira-de-Azevedo ILM, Campos PF, Ching ATC, Mackessy SP. 2016. Colubrid venom composition: an -omics perspective. *Toxins (Basel)* 8(8):230.

Kini RM, Doley R. 2010. Structure, function and evolution of three-finger toxins: mini proteins with multiple targets. *Toxicon* 56(6):855–867.

Komori K, Konishi M, Maruta Y, Toriba M, Sakai A, Matsuda A, Hori T, Nakatani M, Minamino N, Akizawa T. 2006. Characterization of a novel metalloproteinase in Duvernoy's gland of *Rhabdophis tigrinus tigrinus*. *J Toxicol Sci*. 31(2):157–168.

Le NTV, Xue M, Castelnoble LA, Jackson CJ. 2007. The dual personalities of matrix metalloproteinases in inflammation. *Front Biosci*. 12(1):1475–1487.

Lemoine K, Rodríguez-Acosta A. 2003. Haemorrhagic, proteolytic and neurotoxic activities produced by Duvernoy's gland secretion from the false coral snake (*Erythrolamprus bizona* Jan 1863) (Serpentes: Colubridae). *Rev Cient la Fac Ciencias Vet la Univ del Zulia*. 13:371–377.

Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 44(W1):W242–W245.
Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12(1):323.
Mackessy SP. 2010. Handbook of venoms and toxins of reptiles. 1st ed. Boca Raton: CRC Press.
Mackessy SP, Sixberry NM, Heyborne WH, Fritts T. 2006. Venom of the brown treesnake, *Boiga irregularis*: ontogenetic shifts and taxa-specific toxicity. *Toxicon* 47(5):537–548.
Massova I, Kotra LP, Fridman R, Mobashery S. 1998. Matrix metalloproteinases: structures, evolution, and diversification. *FASEB J*. 12(12):1075–1095.
Matsutani M, Fukushima K, Kayama C, Arimitsu M, Hirakawa H, Toyama H, Adachi O, Yakushi T, Matsushita K. 2014. Replacement of a terminal cytochrome c oxidase by ubiquinol oxidase during the evolution of acetic acid bacteria. *Biochim BiophysActa Bioenerg*. 1837(10):1810–1820.
Medeiros C. D, Souza S. D, Silva M. D, Ventura J. D S, Piorelli R. D O, Puorto G. 2019. Bites by *Tomodon dorsatus* (Serpentes, Dipsadidae): clinical and epidemiological study of 86 cases. *Toxicon* 162:40–45.
Menegucci RC, Bernarde PS, Monteiro WM, Ferreira Bisneto P, Ferreira Neto P, Martins M. 2019. Envenomation by an opisthoglyphous snake, Erythrolamprus aesculapii (Dipsadidae), in *southeastern Brazil*. Rev Soc Bras Med Trop. 52:1–3.
Modahl CM, Frietze S, Mackessy SP. 2018. Transcriptome-facilitated proteomic characterization of rear-fanged snake venoms reveal abundant metalloproteinases with enhanced activity. *J Proteomics*. 187:223–234.
Murphy G, Nguyen Q, Cockett MI, Atkinson SJ, Allan JA, Knight CG, Willenbrock F, Docherty AJP. 1994. Assessment of the role of the fibronectin-like domain of gelatinase A by analysis of a deletion mutant. *J Biol Chem*. 269(9):6632–6636.
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):1–14.
Omelchenko MV, Galperin MY, Wolf YI, Koonin EV. 2010. Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. *Biol Direct*. 5(1):31.
Pilcher BK, Wang M, Qin XJ, Parks WC, Senior RM, Welgus HG. 1999. Role of matrix metalloproteinases and their inhibition in cutaneous wound healing and allergic contact hypersensitivity. *Ann NY Acad Sci*. 878:12–24.
Prado-Franceschi J, Hyslop S. 2002. South American colubrid envenomations. *J Toxicol Toxin Rev*. 21(1-2):117–158.
Rawlings ND. 2006. MEROPS: the peptidase database. Nucleic Acids Res. 34(90001):D270–D272.
Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 3(2):217–223.
Rocha MMT, Paixão-Cavalcante D, Tambourgi DV, Furtado MDFD. 2006. Duvernoy's gland secretion of *Philodryas olfersii* and *Philodryas patagoniensis* (Colubridae): neutralization of local and systemic effects by commercial bothropic antivenom (*Bothrops* genus). *Toxicon* 47(1):95–103.
Rokyta DR, Lemmon AR, Margres MJ, Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics* 13(1):312.
Rosenblum G, Van den Steen PE, Cohen SR, Grossmann JG, Frenkel J, Sertchook R, Slack N, Strange RW, Opdenakker G, Sagi I. 2007. Insights into the structure and domain flexibility of full-length pro-matrix metalloproteinase-9/gelatinase B. *Structure* 15(10):1227–1236.
Salomão MDG, Albolea ABP, Almeida Santos SM. 2003. Colubrid snakebite: a public health problem in Brazil. *Herpetol Rev*. 34:307–312.
Sánchez MN, Teibler GP, Sciani JM, Casafús MG, Maruñak SL, Mackessy SP, Peichoto ME. 2019. Unveiling toxicological aspects of venom from the Aesculapian false coral snake *Erythrolamprus aesculapii*. *Toxicon* 164:71–81.
Sanz L, Calvete JJ. 2016. Insights into the evolution of a snake venom multi-gene family from the genomic organization of *Echis ocellatus* SVMP genes. *Toxins (Basel)* 8(7):216.
Sanz L, Harrison RA, Calvete JJ. 2012. First draft of the genomic organization of a PIII-SVMP gene. *Toxicon* 60(4):455–469.
Saviola AJ, Peichoto ME, Mackessy SP. 2014. Rear-fanged snake venoms: an untapped source of novel compounds and potential drug leads. *Toxin Rev*. 33(4):185–201.
Serrano SMT, Maroun RC. 2005. Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon* 45(8):1115–1132.
Shipley JM, Doyle GAR, Fliszar CJ, Ye QZ, Johnson LL, Shapiro SD, Welgus HG, Senior RM. 1996. The structural basis for the elastolytic activity of the 92-kDa and 72-kDa gelatinases: role of the fibronectin type II-like repeats. *J Biol Chem*. 271(8):4335–4341.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
Stöcker W, Grams F, Reinemer P, Bode W, Baumann U, Gomis-Rüth F-X, Mckay DB. 2008. The metzincins—topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a super family of zinc peptidases. *Protein Sci*. 4(5):823–840.
Tasoulis T, Isbister GK. 2017. A review and database of snake venom proteomes. *Toxins (Basel)* 9(9):290.
True JR, Carroll SB. 2002. Gene co-option in physiological and morphological evolution. *Annu Rev Cell Dev Biol*. 18(1):53–80.
Uetz P. 2010. The original descriptions of reptiles. *Zootaxa* 2334(1):59–68.
Vallee BL, Williams RJ. 1968. Metalloenzymes: the entatic nature of their active sites. *Proc Natl Acad Sci U S A*. 59(2):498–505.
Vandooren J, Van Den Steen PE, Opdenakker G. 2013. Biochemistry and molecular biology of gelatinase B or matrix metalloproteinase-9 (MMP-9): *the next decade. Crit Rev Biochem Mol Biol*. 48(3):222–272.
Vonk FJ, Casewell NR, Henkel CV, Heimberg AM, Jansen HJ, McCleary RJR, Kerkkamp HME, 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 U S A*. 110(51):20651–20656.
Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci*. 131(4):281–285.
Wagner GP, Kin K, Lynch VJ. 2013. A model based criterion for gene expression calls using RNA-seq data. *Theory Biosci*. 132(3):159–164.
Wheeler G, Ishikawa T, Pornsaksit V, Smirnoff N. 2015. Evolution of alternative biosynthetic pathways for vitamin C following plastid acquisition in photosynthetic eukaryotes. Elife 4:25.
Wong ESW, Belov K. 2012. Venom evolution through gene duplications. *Gene* 496(1):1–7.
Zaher H, Yánez-Muñoz MH, Rodrigues MT, Graboski R, Machado FA, Altamirano-Benavides M, Bonatto SL, Grazziotin FG. 2018. Origin and hidden diversity within the poorly known Galápagos snake radiation (Serpentes: Dipsadidae). *Syst Biodivers*. 16(7):614–642.
Zelanis A, Teixeira da Rocha MM, de Fátima Domingues Furtado M. 2010. Preliminary biochemical characterization of the venoms of five Colubridae species from Brazil. *Toxicon* 55(2-3):666–669.
Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30(5):614–620.