Assembling an Arsenal: Origin and Evolution of the Snake Venom Proteome Inferred from Phylogenetic Analysis of Toxin Sequences

B. G. Fry* and W. Wüster†

*Australian Venom Research Unit, Department of Pharmacology, University of Melbourne, Parkville, Australia; and
†School of Biological Sciences, University of Wales, Bangor, Wales, UK

We analyzed the origin and evolution of snake venom toxin families represented in both viperid and elapid snakes by means of phylogenetic analysis of the amino acid sequences of the toxins and related nonvenom proteins. Out of eight toxin families analyzed, five provided clear evidence of recruitment into the snake venom proteome before the diversification of the advanced snakes (Kunitz-type protease inhibitors, CRISP toxins, galactose-binding lectins, M12B peptidases, nerve growth factor toxins), and one was equivocal (cystatin toxins). In two others (phospholipase A$_2$ and natriuretic toxins), the nonmonophyly of venom toxins demonstrates that presence of these proteins in elapids and viperids results from independent recruitment events. The ANP/BNP natriuretic toxins are likely to be basal, whereas the CNP/BPP toxins are Viperidae only. Similarly, the lectins were recruited twice. In contrast to the basal recruitment of the galactose-binding lectins, the C-type lectins were shown to be Viperidae only, with the α-chains and β-chains resulting from an early duplication event. These results provide strong additional evidence that venom evolved once, at the base of the advanced snake radiation, rather than multiple times in different lineages, with these toxins also present in the venoms of the "colubrid" snake families. Moreover, they provide a first insight into the composition of the earliest ophidian venoms and point the way toward a research program that could elucidate the functional context of the evolution of the snake venom proteome.

Introduction

The evolution of the venomous function of snakes and the diversification of their toxins has been of tremendous research interest and considerable debate. The advanced snakes (superfamily Colubroidea) make up over 80% of the approximately 2,900 species of snake currently described and contain all the known venomous forms (Greene 1997; Vidal 2002). The origin and evolution of venom-secreting glands and venom toxins has been a subject of much speculation. At present, the evidence-based majority view is that venom-secreting glands evolved at the base of the colubroid radiation, with extensive subsequent "evolutionary tinkering" (Vidal 2002), including the multiple evolution of front-fanged venom delivery systems in the families Viperidae, Elapidae, and Atractaspididae (Underwood 1967; Underwood and Kochva 1993; Vidal 2002) and secondary loss in some other lineages. The remaining majority of the Colubroidea lack front fangs, but most lineages have a venom gland (formerly termed Duvernoy's gland, but see Fry et al. [2003c] for a discussion of why this term has been abandoned) and may or may not have differentiated posterior maxillary teeth to facilitate venom inoculation, including advanced, highly mobile and efficient rear fangs evolving at least once. These snakes have traditionally been lumped into the family Colubridae, but multiple studies of Colubroid phylogeny have shown this family to be paraphyletic, at least with respect to the Atractaspididae and Elapidae (fig. 1) (Underwood 1967; Slowinski and Lawson 2002; Vidal 2002; Vidal and Hedges 2002; Kelly, Barker, and Villet 2003). However, for convenience, we retain the term "colubrid" as an informal designation for the colubroid snakes lacking front-fanged venom delivery systems.

Evidence for the origin of venom-secreting glands at the base of the colubroid radiation comes from comparative morphology and the demonstrated homology of venom-secreting glands of different colubroid lineages (Kochva 1963, 1965, 1978; Underwood and Kochva 1993; Underwood 1997; Jackson 2003) as well as the distribution of these glands across the full spectrum of "colubrid" lineages (Vidal 2002). Nonetheless, some authors maintain the view that the venom-secreting glands of different lineages of "colubrid" snakes may have evolved independently on multiple occasions (Chiszar and Smith 2002).

One source of evidence that has been largely neglected in discussions of the evolution of venom and venom delivery systems in snakes are the venoms themselves (Jackson 2003). The toxins found in snake venoms evolve from recruitment events by which a body protein is recruited into the chemical arsenal of the snake. The toxins often undergo significant variations in sequence and structure, yet typically retain the molecular scaffold of the ancestral protein (Fry et al. 2003b). Phylogenetic analyses of toxin sequences can reconstruct the evolutionary history of toxin gene families, and, in conjunction with an organismal phylogeny, this allows the recruitment and diversification of the toxins to be related to the phylogeny of the snakes.

The previous lack of use of toxin data may be partly because of the fact that, until very recently (Yamazaki et al. 2002; Fry et al. 2003a), the only toxins that have been sequenced were from the front-fanged families of medical importance (Elapidae, Viperidae, and Atractaspididae), leaving the majority of colubroid lineages unstudied. However, the first full sequences of "colubrid" toxins proved revealing. We previously isolated and characterized a neurotoxic 3FTx (three-finger toxin), a toxin family previously believed to be unique to the Elapidae (Fry et al. 2003b), from the colubrine *Coelognathus radiatus*, and demonstrated that it was phylogenetically rooted within the elapid 3FTx family (Fry et al. 2003a). This finding suggested that the 3FTx family was recruited into the chemical arsenal of snakes before the split between the
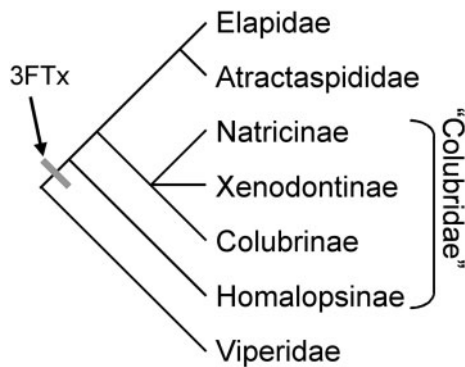
FIG. 1.—Phylogeny of the major families of advanced snakes (Colubroidea) discussed in this study (Vidal and Hedges 2002). Only major lineages and those mentioned in text are shown. The timing of the recruitment of 3FTx into the snake venom proteome, based on the analyses of Fry et al. (2003*a*, 2003*c*), is also shown.

elapid and colubrine lineages (fig. 1). However, the 3FTx family probably does not represent a basal recruitment, as these toxins are lacking in the venoms of the Viperidae, the most basal colubroid lineage (fig. 1) (Fry et al. 2003*c*). The recruitment events leading to PLA$_2$ toxins (phospholipase A$_2$) being present in both Elapidae and Viperidae venoms are clearly independent, with the Elapidae toxins belong to the "pancreatic-type" (group I) PLA$_2$ toxins, whereas the Viperidae toxins belong to the "synovial-type" (group II) PLA$_2$ toxins (Heinrikson, Krueger, and Keim 1977). In view of the fact that the PLA$_2$ of viperids and elapids result from separate recruitment events, this leaves open the possibility that vipers and elapids plus "colubrids" may have evolved venom independently.

The aim of this paper is to use the phylogenetic analysis of toxin amino acid sequences to understand the pattern of recruitment of toxin families into the snake venom proteome. In particular, we aim to test whether toxin families represented in both vipers and elapids go back to a single recruitment event before the radiation of the Colubroidea or to independent recruitment events after the lineage split between the two families.

Our approach is based on the fundamental prediction that the genes of a given toxin family should only be monophyletic in the clade in which the family was first recruited into the venom proteome but nonmonophyletic in any subclade of that clade, as their coalescence time would predate the diversification of the subclade. The interpretations that can be derived from different patterns of toxin gene phylogeny are shown in figure 2. Because vipers and elapids are representatives of the two most distantly related lineages among the Colubroidea (fig. 1) (Slowinski and Lawson 2002; Vidal and Hedges 2002; Kelly, Barker, and Villet 2003), with the possible exception of the Pareatinae or Xenodermatinae, our results have implications for the pattern of presence/absence of toxins in other advanced snake lineages. Where the presence of a given toxin family in the elapids and vipers goes back to a single recruitment event, the presence of that toxin family in other colubroid lineages is likely. However, this is not necessarily the case if the presence of the toxin family in vipers and elapids is the result of independent recruitment events.

We examined the molecular phylogeny and evolution of toxin classes shared between the Viperidae, the most basal group of colubroid snakes, and the Elapidae, one of the most derived groups of colubroid snakes, which are rooted among other colubroids in the other basal branch of the phylogeny of the Colubroidea (Vidal and Hedges 2002). Molecular scaffold types examined were BPTI/Kunitz-type proteinase inhibitor toxins (BPTI = bovine pancreatic trypsin inhibitor), CRISP toxins (cysteine-rich secretory proteins), cystatin toxins, C-type lectin toxins and GBL toxins (galactose-binding lectins), M12B peptidases, natriuretic toxins, NGF (nerve growth factor) toxins, and PLA$_2$ (phospholipase A$_2$) toxins. We did not examine toxin types for which sequences have been reported to date from only one family.

Materials and Methods

To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein). Sequences were aligned using the program ClustalX (Thompson et al. 1997), followed by visual inspection for errors. In the case of the BPTI/Kunitz protease inhibitors, as these protein domains have been incorporated into longer preproteins, alignments were trimmed on either side of the domain. Because of the large number of sequences in each data set, we conducted our phylogenetic analyses in two steps. For each data set, phylogenetic trees containing all sequenced proteins were initially reconstructed using the maximum-parsimony (MP) and Neighbor-Joining (Saitou and Nei 1987) methods. MP heuristic searches were conducted using the program PAUP* version 4.0b10 (Swofford 2002) and random stepwise taxon addition with TBR branch swapping and the PROTPARS weighting scheme (Felsenstein 2001), which takes into account the number of changes required at the nucleotide level to substitute one amino acid for another. Number of sequences, alignment length (including gaps), and parsimony-informative sites were as follows: BPTI/Kunitz —198 sequences, 182 characters, 142 of which were parsimony informative; CRISP—57 sequences, 336 characters, 292 of which were parsimony informative; Lectins—104 sequences, 267 characters, 192 of which were parsimony informative; Cystatin—30 sequences, 174 characters, 143 of which were parsimony informative; M12B—238 sequences, 897 characters, 792 of which were parsimony informative; nerve growth factor—174 sequences, 352 characters, 242 of which were parsimony—informative; PLA$_2$—127 sequences, 458 characters, 265 of which were parsimony informative; natriuretic peptides—83 sequences, 348 characters, 193 of which were parsimony informative.

Neighbor-Joining analysis was carried out using the program MEGA version 2.1 (Kumar et al. 2001), using Poisson-corrected distances. In this manner, we identified the clades that contained the venom proteins. Once such clades were identified, data sets containing representatives of nonvenom proteins, with particular emphasis on representing clades close to the venom proteins (except cystatin, natriuretic peptide, and nerve growth factor data sets, which included all nonvenom sequences), and all of
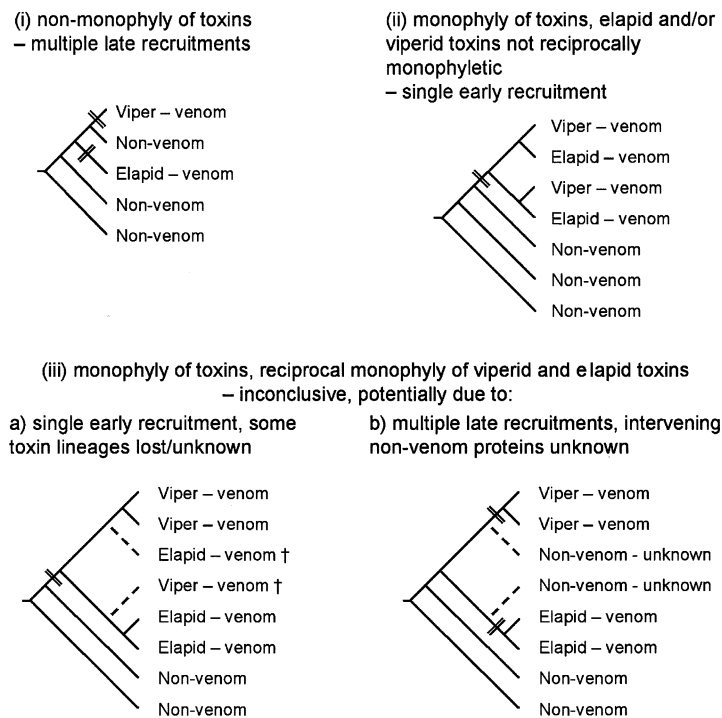
FIG. 2.—Schematic approach to interpreting patterns of recruitments of protein families into the venom proteome from gene trees. Double-bars represent recruitment events.

the venom proteins (except for PLA$_2$ and the α-chains and β-chains of the C-type lectins: because of the extremely large number of venom toxin sequences, representatives of the full breadth of gene phylogenetic diversity were selected) were analyzed using Bayesian inference implemented on MrBayes version 3.0b4 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck, in press). The method uses Markov chain Monte Carlo methods to generate posterior probabilities for each clade represented in the tree. The analysis was performed by running a minimum of 1×10^6 generations in four chains and saving every 100th tree. The log-likelihood score of each saved tree was plotted against the number of generations to establish the point at which the log-likelihood scores of the analysis reached their asymptote, and the posterior probabilities for clades established by constructing a majority rule consensus tree for all trees generated after the completion of the burn-in phase. Alignments can be obtained from the authors.

Results

The results of the phylogenetic analyses are summarized in table 1. Of the eight toxin families analyzed, five correspond to pattern (ii) of figure 2, in that the toxin proteins are monophyletic to the exclusion of nonvenom proteins, but viperid and elapid toxins are not reciprocally monophyletic. These are the BPTI-Kunitz (fig. 3), CRISP (fig. 4), M12B (fig. 5), and nerve growth factor (fig. 6) families and the GBL toxins. On the other hand, in PLA$_2$ (fig. 7) and natriuretic peptides (fig. 8), the elapid and

viperid toxins originate from independent recruitment events (pattern [i] in figure 2). The PLA$_2$ toxins in elapids were more similar to the "synovial-type" body PLA$_2$, whereas the viper venom PLA$_2$ were more similar to the "pancreatic-type" body PLA$_2$ sequences (group I and group II PLA$_2$, respectively [Heinrikson, Krueger, and Keim 1977]). The natriuretic toxins in elapid venoms were rooted deeply within the ANP/BNP lineage, whereas the viper natriuretic toxins were clearly evolved from within the CNP lineage (fig. 8).

The cystatins correspond to pattern (iii) of figure 2. Only single elapid and viperid sequences are available (fig. 9), so the reciprocal monophyly of the toxins of the two families cannot be tested. The observed toxin phylogeny is, thus, consistent with both single and multiple recruitments of these proteins into the venom proteome. However, the low distance values between the two sequences are indicative of a single recruitment event. Finally, the lectin toxins contained two venom clades, indicating that this protein family was recruited twice into the snake proteome (fig. 10). In the GBL (galactose-binding lectins) toxin clade, viper and elapid toxins were monophyletic relative to all other sequences but paraphylectic relative to each other, thus, consistent with a single recruitment event at the base of the Colubroidea tree (pattern [ii]). The α-chains and β-chains of the C-type lectins (found only in viper venoms) form a monophyletic gene clade relative to all other sequences but are also reciprocally monophylectic relative to each other, thus, appearing to have arisen from a single recruitment event postdating the Viperidae split off from the rest of the Colubroidea tree, yet with a very early gene duplication to form the α-chains and β-chains.

Table 1
Summary of the Results of the Phylogenetic Analyses

Toxin Family	Monophyly of Venom Proteins	Reciprocal Monophyly of Viperid and Elapid Toxins	Interpretation
CRISP	Yes	Elapid toxins polyphyletic, strongly supported	Single early recruitment
Cystatin	Yes	Unresolved —only one sequence each	Equivocal (although the low distance values are consistent with a single early recruitment event)
Kunitz	Yes	Elapid toxins paraphyletic, strongly supported	Single early recruitment
Lectins	Yes/No	Within galactose-binding lectins, elapid toxins paraphyletic, strongly supported; C-type lectins α-chains and β-chains in vipers only	Galactose-binding lectins; single early recruitment event; C-type lectins α-chains and β-chains resulting from one late, independent recruitment in vipers with rapid gene duplication to form α-chains and β-chains
M12B	Yes	Viperid toxins paraphyletic, strongly supported	Single early recruitment
Natriuretic peptides	No	Yes, although viper toxins are present in two unrelated types; the viper-only CNP-toxins and a second that may be homologous to the ANP/BNP-toxins found in elapid venoms	One independent recruitment (viper CNP toxins) and an early ANP/BNP toxin recruitment.
Nerve growth factor	Yes	Viperid toxins paraphyletic, strongly supported	Single early recruitment
Phospholipase A_2	No	Yes	Independent recruitments

Discussion

Our results provide new insights into the origin of venom in snakes by demonstrating that several toxin families were recruited into the venom apparatus of snakes before the diversification of extant colubroid snake lineages.

Origin and Recruitment of Toxin Families

Of the eight toxin families analyzed, five (Kunitz-type protease inhibitors, CRISP toxins, GBL toxins, M12B peptidases, and NGF toxins) correspond unequivocally to pattern (ii) of figure 2, which allows the rejection of the possibility that these proteins were recruited into the snake venom proteome after the lineage split between the vipers and the elapids (fig.11). In all, the lack of reciprocal monophyly of the viperid and elapid toxins is supported by high posterior probabilities in the Bayesian analyses. The low distance values between the sequences in the cystatin toxin gene family is also suggestive of a single early recruitment event. However, the reciprocal monophyly of viperid and elapid toxins, arising as a consequence of only one sequence being available from each lineage, does not allow rejection of the alternative of independent later recruitments. Additional sequences would be revealing.

Three of the toxin families analyzed, the lectin toxins, PLA_2, and the natriuretic peptides, are clearly the result of two independent recruitment events. In the case of the PLA_2 toxins, this confirms existing information on the evolution of the toxins independently in elapid (group I PLA_2) and viper venoms (group II PLA_2) (Heinrikson, Krueger, and Keim 1977; Dufton and Hider 1983) and also serves as a useful control for the possibility that apparent monophyly of all toxins within a protein family may be a result of homoplaseous convergence because of similar function

rather than common ancestry. The example of the PLA_2 toxins also illustrates the requirement for a toxin phylogeny to infer a single recruitment. Their widespread presence in snake venoms had been interpreted as homologous across all colubroids (e.g., Kochva 1987; Vidal 2002), but phylogenetic analyses show that this is actually the result of two independent recruitment events.

The lectin protein family was recruited once at the base of the Colubroidae tree (GBL toxins) and again in the Viperidae lineage subsesquent to its split from the rest of the advanced snakes (C-type lectins), with the vipers, thus, containing both C-type lectins as well as GBL, whereas all other lineages contain only GBL (fig. 10). In contrast, the actual points of recruitment of the group I PLA_2 and natriuretic toxin families remain unknown. Group I PLA_2 toxins have so far only been characterized and sequenced from elapid venoms. However, although there are as yet no PLA_2 toxin sequences from "colubrid" venoms, corresponding molecular masses were found in most "colubrid" venoms examined by mass spectrometry, including the basal Homalopsinae venoms (Fry et al. 2003c). This makes it likely that the group I PLA_2 toxins were recruited after the divergence of the vipers but before the divergence by other snake families, as is the case of the 3FTx family (Fry et al. 2003a, 2003c). However, until PLA_2 toxins from some of the "colubrid" lineages are adequately characterized and sequenced, the possibility of further independent recruitments cannot be excluded.

The snake natriuretic toxins were shown to be of two types (fig. 8). The sole elapid natriuretic toxin for which a full-length gene sequence was available (Ho et al. 1997) aligned within the ANP/BNP natriuretic peptide lineage, whreas the Viperidae natriuretic toxins aligned with the CNP natriuretic peptides (fig. 8). The sequences of the viper CNP natriuretic toxins are highly conserved, so much so
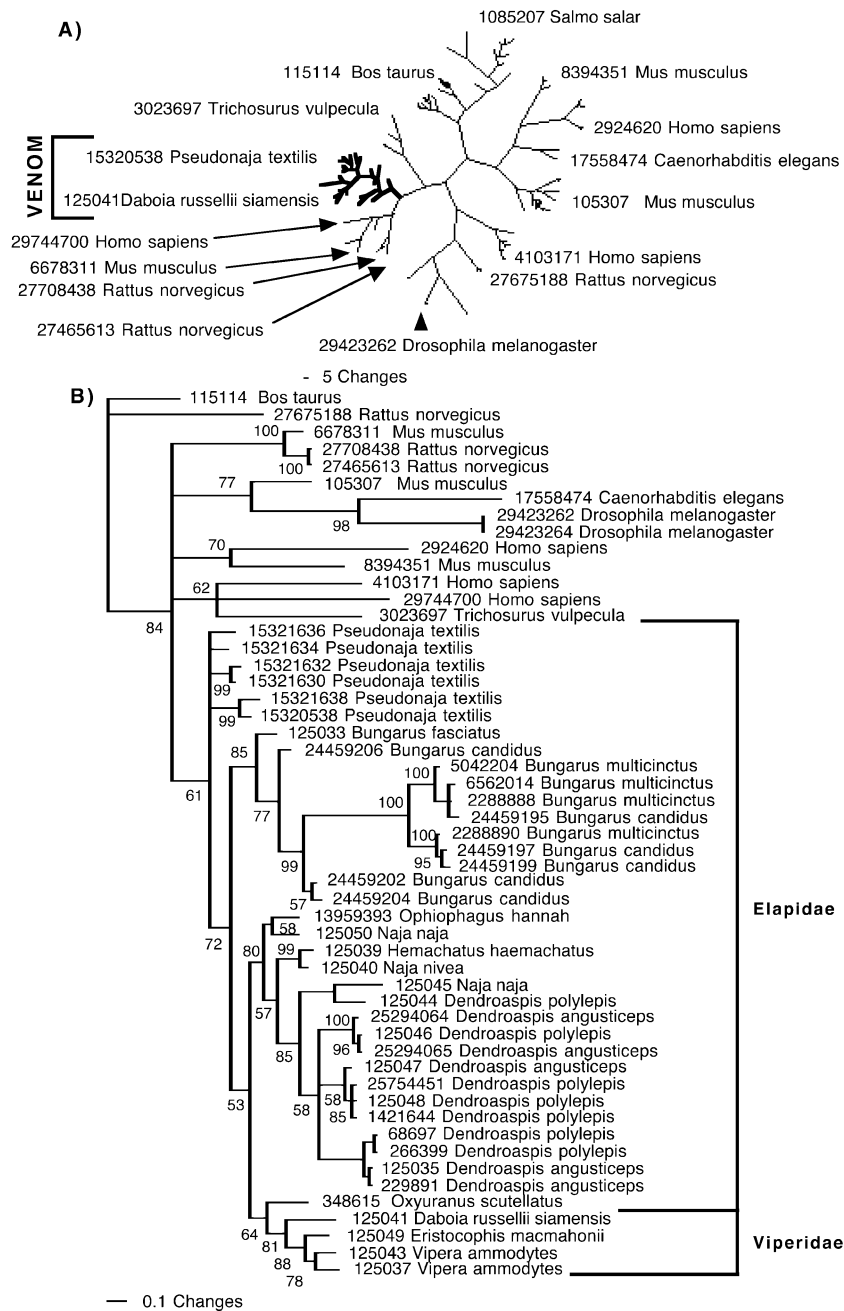
Fig. 3.—Molecular phylogenetic analysis of the BPTI/Kunitz-type proteinase inhibitor domain-containing proteins. (*A*) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (*B*) outgroup-rooted Bayesian analysis of all toxin sequences and representative nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

that the natriuretic toxin sequence from the venom gland of *Bothrops jararaca* (1580720 [Murayama et al. 1997]) is virtually identical to the natriuretic peptide gene transcript from the brain of the same species (Hayashi et al. 2003). Because we included only whole protein sequences in our phylogenetic analyses, we did not analyze partial sequences of divergent natriuretic toxins such as α-lebetin from the viper *Macrovipera lebetina* (Barbouche et al. 1996) and several toxins from the venoms of Australian elapids (32363239, 32363242, and 32363245 [Fry et al. 2002]), as only the final, highly processed short peptide sequences of

these toxins are known. However, all these toxins have the C-terminal extension characteristic of ANP/BNP natriuretic peptides, in contrast to the CNP natriuretic peptides, which terminate with the second cysteine. Thus, the ANP/BNP natriuretic toxins may be another ancient recruitment at the base of the Colubroidea tree, whereas the CNP natriuretic toxins are an independent recruitment that occurred after the vipers split off from the remainder of the advanced snakes, a scenario similar to that which also clearly occurred with the lectin toxins. However, a full-length transcript of α-lebetin (or a similar viper toxin) and additional elapid toxin
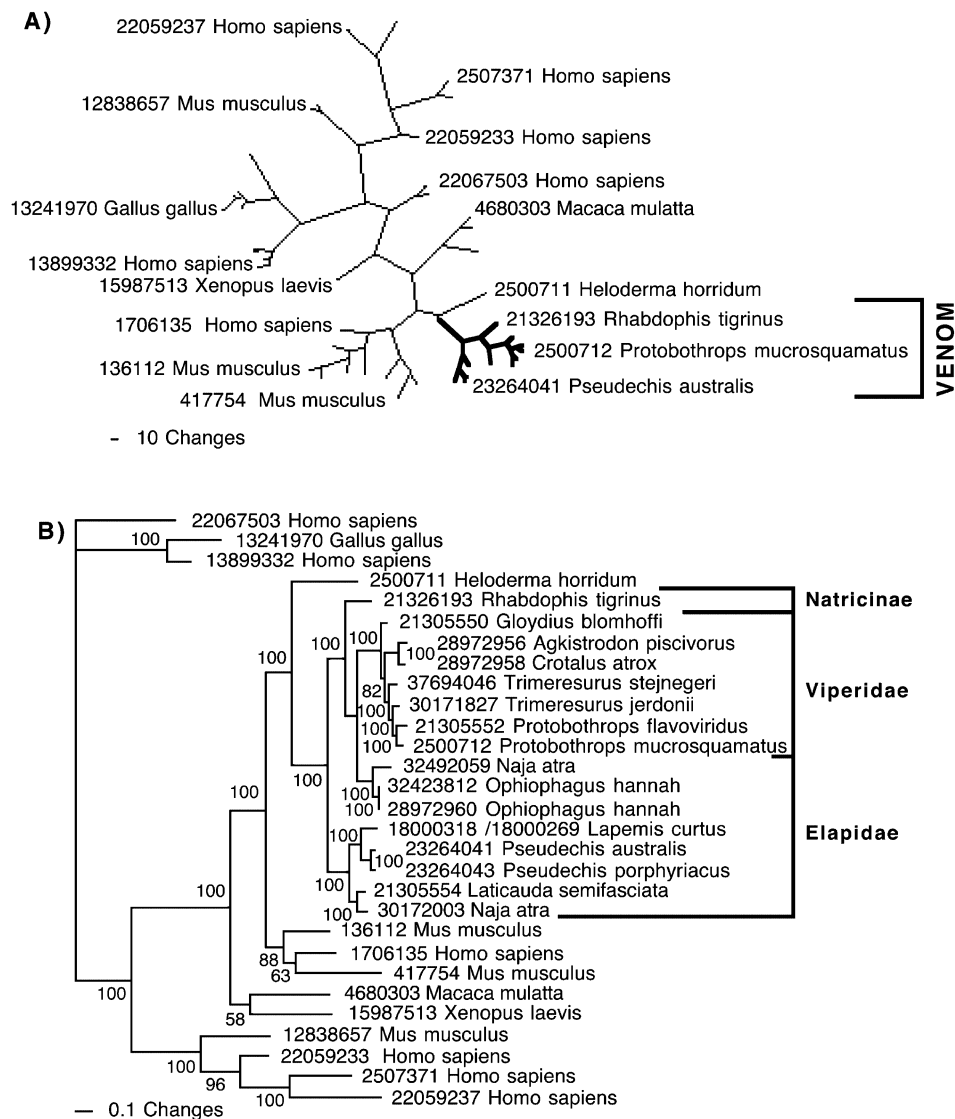
FIG. 4.—Molecular phylogenetic analysis of CRISP proteins. (*A*) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (*B*) outgroup-rooted Bayesian analysis of all toxin sequences and representative nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

sequences would be necessary to confirm the homology of the ANP/BNP toxins and the consequent basal recruitment.

The CRISP toxins from the snake and Helodermatid lizard venoms (fig. 4) represent independent recruitments of CRISP proteins for use as toxins. The venom-secreting structures in advanced snakes and *Heloderma* are different, nonhomologous structures (supralabial and infralabial glands, respectively), and the last common ancestor of *Heloderma* and advanced snakes would have been a basal varanoid (Forstner, Davis, and Arévalo 1995; Lee 1997) or even a basal scleroglossan (Rieppel et al. 2003) lizard devoid of a venomous function. However, it is interesting that, despite being the result of an independent recruitment event, helothermine represents the sister group to the snake CRISP proteins. Clearly, helothermine and snake venom CRISPs were recruited from closely related body proteins. This illustrates why phylogenetic patterns corresponding to pattern (iii) of figure 2, as found in the cystatin toxins

(fig. 9), cannot be taken as evidence of a single recruitment event, despite the low distance levels between the toxin sequences relative to the nontoxin sequences, which support the most parsimonious conclusion of a single recruitment.

Assembly of the Snake Venom Proteome

The pattern of recruitment of venom protein families revealed in this study shows that the last common ancestor of the extant colubroid snake radiation already had a complex venom containing at least five and possibly six or more of the toxin gene families shared by vipers and elapids today. Moreover, because some toxin lineages may have been lost in either vipers or elapids, this primitive colubroid may have had additional toxin families represented in its venom. Additional toxin families were recruited into the chemical arsenal soon after the basal
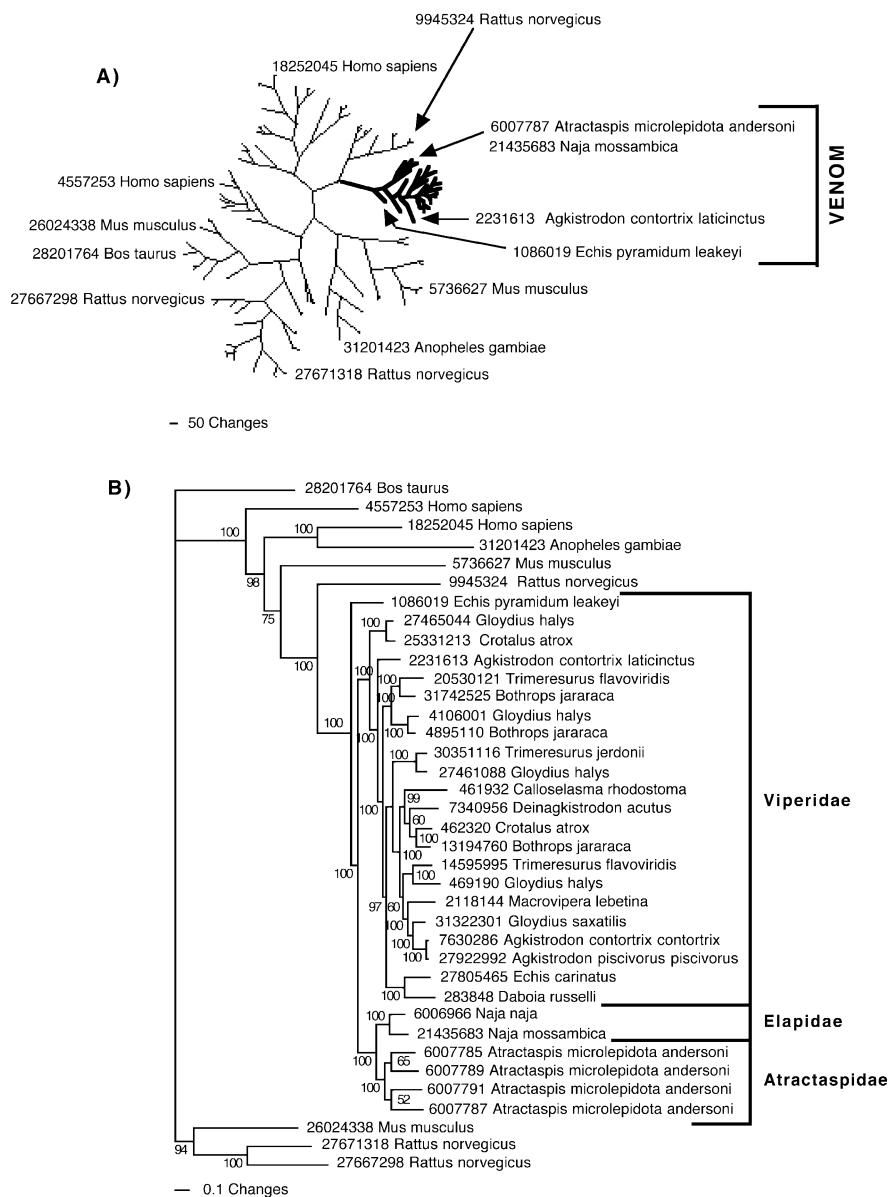
FIG. 5.—Molecular phylogenetic analysis of M12B peptidases. (A) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (B) outgroup-rooted Bayesian analysis of representative toxin and nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

lineage split between vipers and the remaining colubroids. The 3FTx family was recruited immediately after the vipers split from the remaining colubroid lineages (Fry et al. 2003a, 2003c) and the pancreatic-type (group I) PLA$_2$ toxins appear to have been recruited at approximately the same time, and both appear to be widespread across most colubroid lineages.

A number of other toxin families are presently known only from either the atractaspidids, the elapids, or the viperids and may have been recruited into the venom proteome later during the evolution of these lineages. The sarafotoxins appear to be unique to the Atractaspididae, having been isolated thus far from only *Atractaspis* species (e.g., Kloog et al. 1988). Toxin molecular scaffolds sequenced only from Elapidae venoms include acetylcho-

linesterase, cobra venom factor, factor Xa prothrombin–activating toxins, factor V toxins, prokinecitin-like peptides, wapins, and toxins containing the SPRY domain. Current Viperidae-only toxins include myotoxic peptides, S1 peptidases, vascular endothelial growth factor–like toxins, and waglerins. Although the full transcriptome of an elapid venom gland has not been determined, extensive cloning of a viper cDNA library revealed at least one conclusively new viper toxin but did not reveal any additional toxin types shared with elapid venoms (Junqueira-de-Azevedo and Ho 2002). L-amino oxidase activity has been reported widely from both Elapidae and Viperidae venoms, but, to date, only full-length Viperidae sequences have been reported (e.g., 5565692 from *Crotalus atrox*). A fragment from *Ophiophagus hannah* (6093637) entered as
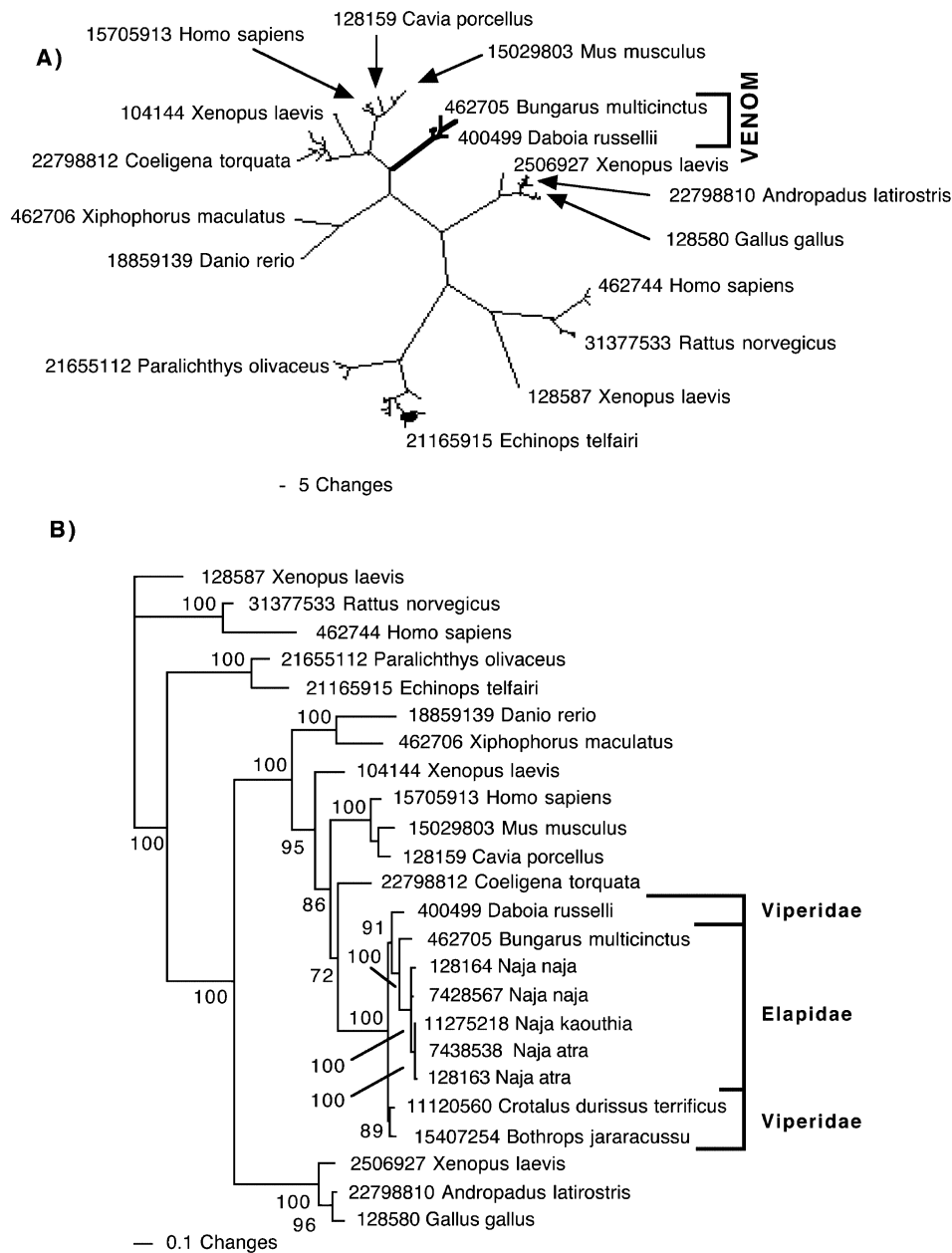
Fig. 6.—Molecular phylogenetic analysis of nerve growth factors. (A) Unrooted maximum-parsimony network and (B) outgroup-rooted Bayesian analysis of all toxin sequences and representative nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

a L-amino oxidase displayed homology to a Viperidae fragment (1085231 from *Calloselasma rhodostoma*), thus indicating that this toxin type is yet another basal recruitment.

The results of this study allow a number of inferences on the composition of the hitherto largely unstudied bulk of "colubrid" venoms. The fact that five to eight toxin gene families were recruited before the basal divergence of the colubroids suggests that they ought to be present in the venoms of most other colubroid clades as well. Because of the almost complete lack of well-characterized "colubrid" toxins, there is little in the way of firm, sequence-confirmed

evidence to test this hypothesis at present. However, although we only included complete protein sequences in our phylogenetic analysis, the available partial CRISP toxin sequences from several "colubrid" lineages confirm the ubiquity of this toxin type in Colubroidea venoms (Hill and Mackessy 2000; Fry et al. 2003c), and, similarly, the partial sequence of the potently prothrombin activating M12B toxin from the colubrine *Dispholidus typus* (Kamiguti et al. 2000) suggests the widespread presence of the M12B peptidases, including the disintegrin domain, within the Colubroidea, as would be predicted from a basal recruitment of these proteins.
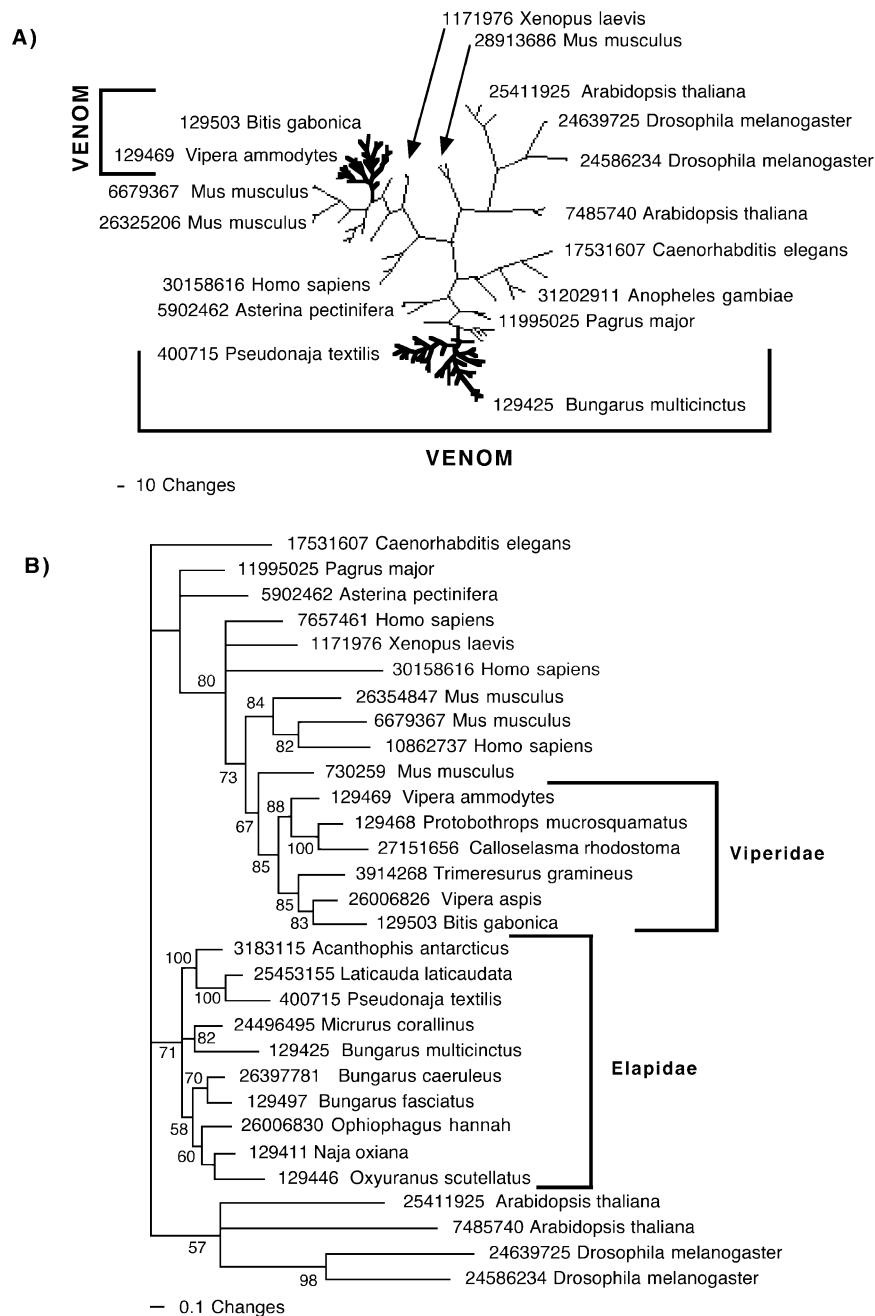
A)

B)

FIG. 7.—Molecular phylogenetic analysis of PLA$_2$. (*A*) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (*B*) outgroup-rooted Bayesian analysis of representative toxin and nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

In view of the close phylogenetic relationship between some "colubrid" lineages and the elapids, it is possible that other toxin classes currently believed unique to the Elapidae may eventually be found in some "colubrid" lineages, as has already happened with the 3FTx (Fry et al. 2003*a*, 2003*c*). This would indicate that they were recruited earlier in Colubroid history. On the other hand, based on our understanding of the phylogeny of the Colubroidea, it is unlikely that any of the toxins currently known only from the vipers and not from the elapids will be found in the "colubrids."

Although a significant number of snake venom toxin types are currently known, the presence of additional, novel toxin groups in many colubroid clades is likely. The full transcriptome of a viper cDNA library revealed at least one conclusively new viper toxin (Junqueira-de-Azevedo and Ho 2002). This reinforces the fact that, despite the status of the venom apparatus as an ancestral character of all Colubroidea and the basal recruitment of several toxin types, new toxin types continue to be recruited independently within the lineages. Consequently, although the venoms of the many almost entirely unexplored "colubrid"
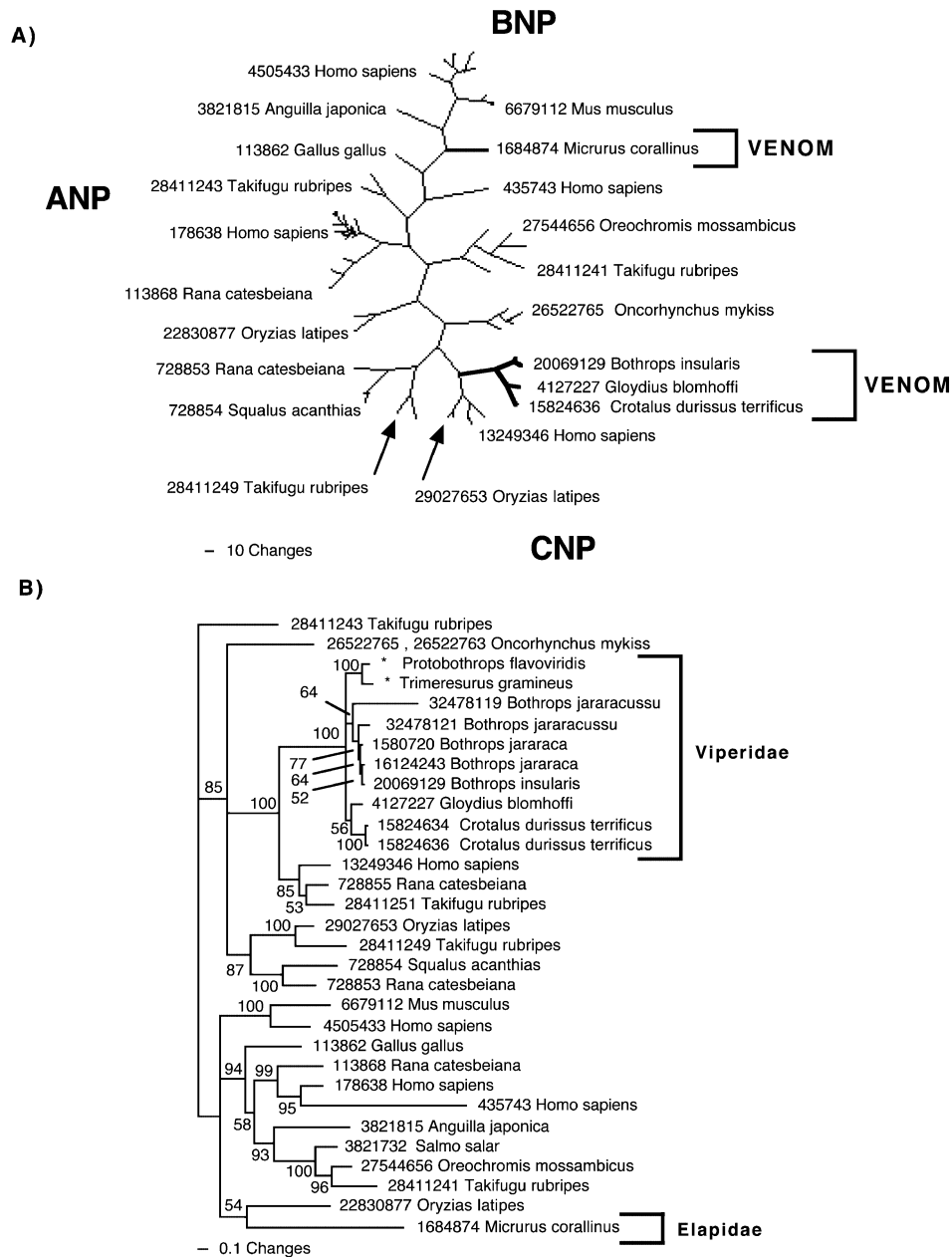
A)

BNP

4505433 Homo sapiens

3821815 Anguilla japonica

6679112 Mus musculus

113862 Gallus gallus

1684874 Micrurus corallinus   **VENOM**

28411243 Takifugu rubripes

ANP

435743 Homo sapiens

178638 Homo sapiens

27544656 Oreochromis mossambicus

28411241 Takifugu rubripes

113868 Rana catesbeiana

26522765 Oncorhynchus mykiss

22830877 Oryzias latipes

728853 Rana catesbeiana

20069129 Bothrops insularis

4127227 Gloydius blomhoffi

15824636 Crotalus durissus terrificus   **VENOM**

728854 Squalus acanthias

13249346 Homo sapiens

28411249 Takifugu rubripes

29027653 Oryzias latipes

— 10 Changes

CNP

B)

28411243 Takifugu rubripes

26522765 , 26522763 Oncorhynchus mykiss

100 ⌐ * Protobothrops flavoviridis

└ * Trimeresurus gramineus

64

32478119 Bothrops jararacussu

100 ⌐ 32478121 Bothrops jararacussu

1580720 Bothrops jararaca

77

16124243 Bothrops jararaca

64

20069129 Bothrops insularis

85   52

4127227 Gloydius blomhoffi

100

56 ⌐ 15824634 Crotalus durissus terrificus

100 └ 15824636 Crotalus durissus terrificus

Viperidae

13249346 Homo sapiens

85

728855 Rana catesbeiana

53

28411251 Takifugu rubripes

100 ⌐ 29027653 Oryzias latipes

28411249 Takifugu rubripes

87

728854 Squalus acanthias

100

728853 Rana catesbeiana

100

6679112 Mus musculus

4505433 Homo sapiens

113862 Gallus gallus

94   99

113868 Rana catesbeiana

178638 Homo sapiens

95

435743 Homo sapiens

58

3821815 Anguilla japonica

93

3821732 Salmo salar

100 ⌐ 27544656 Oreochromis mossambicus

96 └ 28411241 Takifugu rubripes

54

22830877 Oryzias latipes

1684874 Micrurus corallinus   **Elapidae**

— 0.1 Changes

Fig. 8.—Molecular phylogenetic analysis of natriuretic peptides. (*A*) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (*B*) outgroup rooted Bayesian analysis of all toxin sequences and representative nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein). *Protobothrops flavoviridis* and *Trimeresurus gramineus* sequences from Higuchi et al. 1999.

lineages can confidently be predicted to contain some of the more widespread toxin families, it also appears extremely likely that that many of the major lineages will be found to contain multiple, hitherto unknown, novel toxin families, some of which may be of considerable biomedical research interest.

Origin and Evolution of Venom in Snakes

Whereas morphological analyses (Underwood and Kochva 1993; Underwood 1997; Jackson 2003) and analyses of advanced snake phylogeny (Gravlund 2001;

Slowinski and Lawson 2002; Vidal 2002; Vidal and Hedges 2002; Kelly, Barker, and Villet 2003) have provided extensive evidence for the homology of the venom-secreting glands in all advanced snakes, much less was known about how the biochemical arsenal of snakes was assembled in the course of evolution. Our analyses show that many of the toxin gene families found in today's snake venoms were recruited into the venom proteome very early in the evolution of advanced snakes, before their major radiation (fig. 11), and that their presence in the venoms of multiple snake lineages is, thus, homologous. The homology of multiple venom families across the
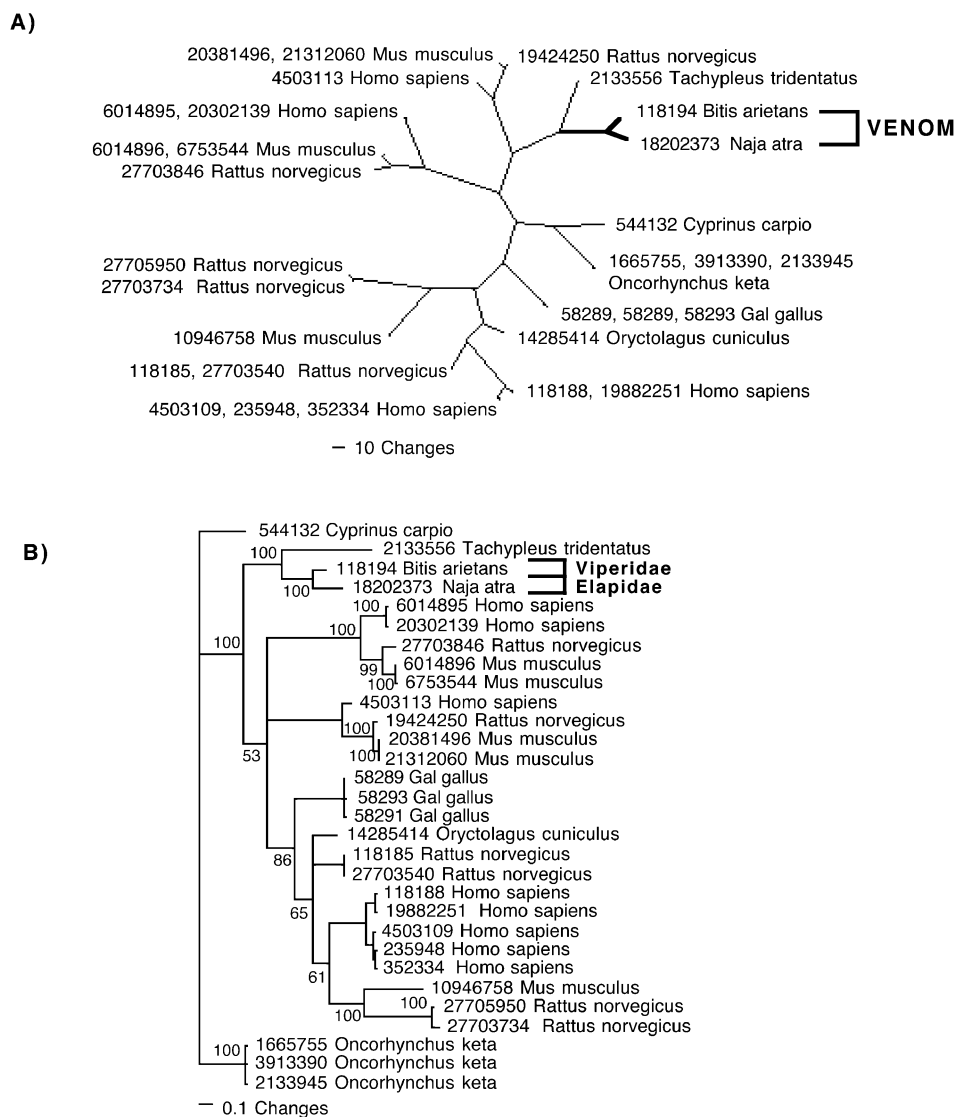
A)

B)

FIG. 9.—Molecular phylogenetic analysis of cystatins. (*A*) Unrooted maximum-parsimony network with all toxin and nontoxin sequences labeled and (*B*) outgroup rooted Bayesian analysis of all toxin and nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein).

venom glands of the most basally split lineages of colubroid snake represents strong additional evidence for the homology of the venom glands across the Colubroidea. This also reinforces the point that the distinction between "colubrid" Duvernoy's gland and the venom glands of the three front-fanged snake radiations is evolutionarily misleading (Fry et al. 2003*c*).

Our study has, thus, provided new insights into the evolutionary assembly of the complex and sophisticated biochemical arsenal of snakes and provided a partial insight into the composition of the earliest snake venoms. However, much remains to be learned about the origin of venom in snakes and, particularly, its functional context. The complexity of the earliest snake venoms demonstrated here suggests that the function of these early venoms may have been similar to that of modern "colubrid" venoms. Clearly, venoms evolved into complex and sophisticated

secretions soon after the initial evolution of serous supra-labial glands at the base of the colubroid radiation (Vidal 2002). However, many questions remain unanswered. For instance, were additional protein families recruited randomly into the snake venom proteome in the course of evolution, or were most families recruited early, followed by relative stasis later? Was the evolution of more specialized venom delivery systems associated with increased recruitment of new toxin gene families? Similarly, did the development of heat-seeking pits in the Viperidae also drive additional recruitment events? More information is needed on structure-function relationships in the almost totally neglected "colubrid" venoms. These may eventually allow inferences on the activities of the earliest snake venom toxins and venoms, and may provide new insights on their function and use. For example, the isolation and characterization from 'colubrids' of 3FTx revealed the
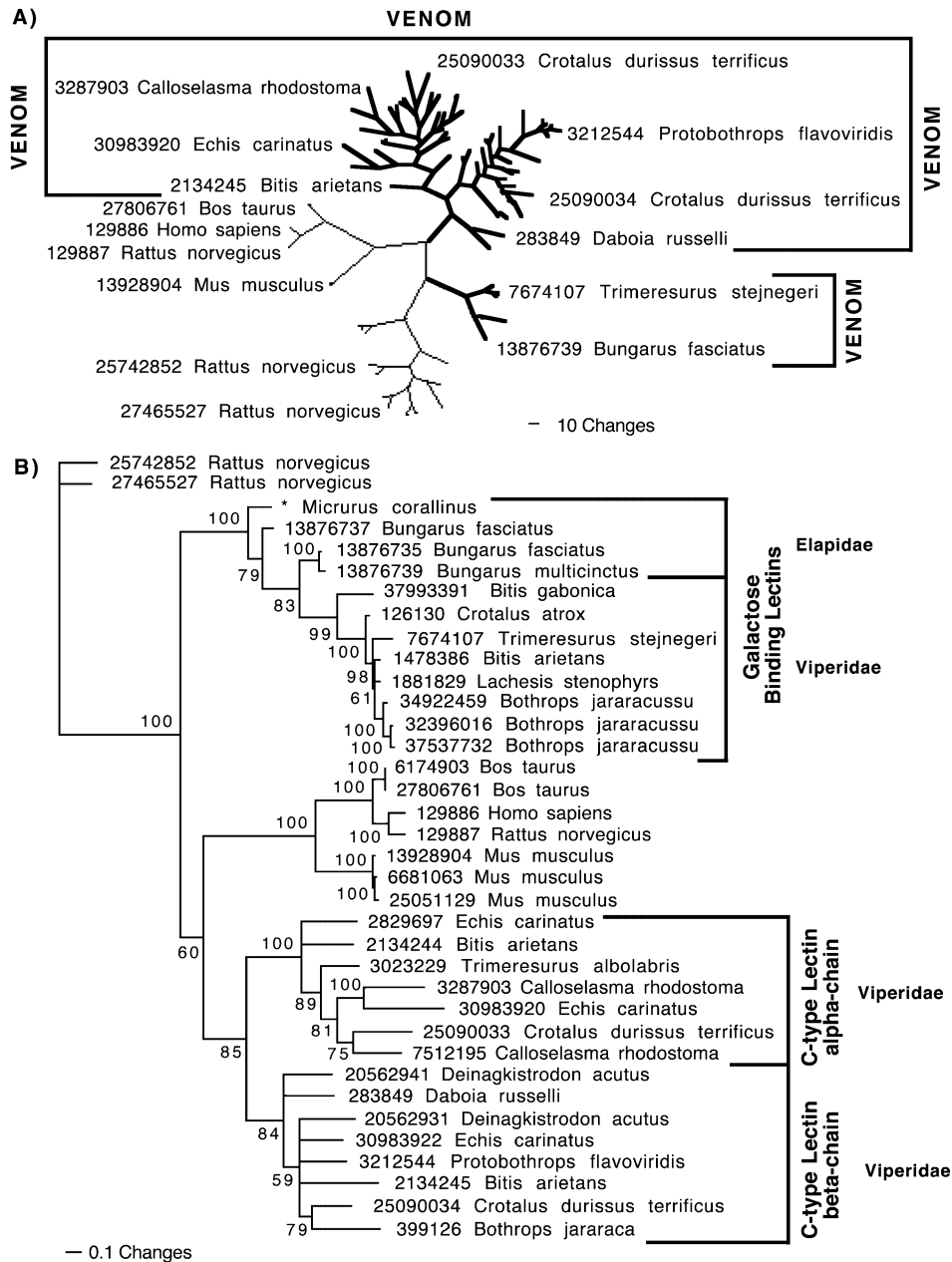
FIG. 10.—Molecular phylogenetic analysis of lectins. (A) Unrooted maximum-parsimony network with representative toxin and nontoxin sequences labeled and (B) outgroup rooted Bayesian analysis of all galactose-binding lectin toxin sequences and representative C-type lectin toxin sequence and nontoxin sequences. To minimize confusion, all proteins sequences are referred to by their NCBI accession numbers (http:// www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein). *Micrurus corallinus* sequence from Ho et al. (1995).

ancestral 10-cysteine framework, as well as demonstrating that α-neurotoxic activity was the basal activity of this toxin type (Fry et al. 2003a, 2003b, 2003c).

Resolution of the many outstanding questions will require evidence from a variety of sources. In particular, additional studies of "colubrid" venoms, particularly the documentation and characterization of the different toxin families by sequencing in transcriptome studies (Junqueira-de-Azevedo and Ho 2002), so far lacking even for the medically important elapids, are required to provide further insights into the diversity of toxin families present in the colubroid radiation. Additionally, we remain short of information on the biological role and function of venom in modern "colubrids," another topic that has been neglected by all but a few studies (Rodríguez-Robles and Thomas 1992; Hayes et al. 1992; Salomão and Laporta-Ferreira 1994). More such studies may provide further insights into the possible uses of relatively complex venoms coupled with "colubrid"-type venom delivery systems, which surely characterized the earliest colubroids.

In conclusion, we hope that the results of this study will stimulate more interest in the venoms of the hitherto
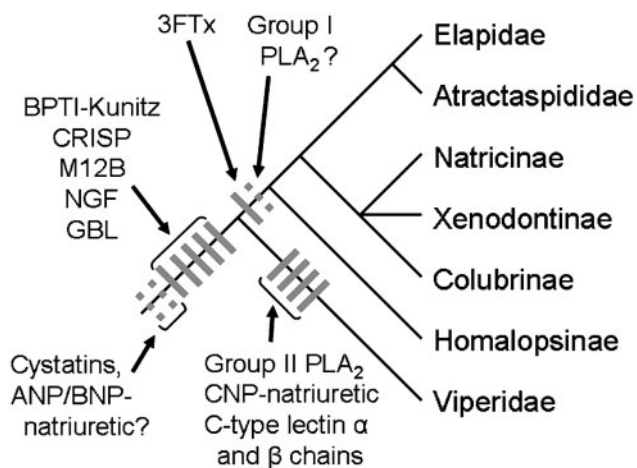
F$_{IG}$. 11.—Pattern of recruitment of toxin gene families into the snake venom proteome. The recruitment of cystatins is equivocal (although the low distance values between the single Elapidae and Viperidae toxins is consistent with a single origin). ANP/BNP natriuretic toxins have been fully sequenced only from Elapidae species, but partial sequences with apparent homology have been sequenced from Viperidae species. Toxin families hitherto known only from one lineage have not been mapped onto the tree.

neglected, medically unimportant majority of venomous snake lineages, with a view towards improving our understanding of one of the most sophisticated integrated weapon systems in the natural world.

Acknowledgments

Literature Cited

Barbouche, R., N. Marrakchi, P. Mansuelle, M. Krifi, E. Fenouillet, H. Rochat, and M. El Ayeb. 1996. Novel anti-platelet aggregation polypeptides from *Vipera lebetina* venom: isolation and characterization. FEBS Lett. **392**:6–10.

Chiszar, D., and H. M. Smith. 2002. Colubrid envenomations in the United States. J. Toxicol. Toxin Rev. **21**:85–104.

Dufton, M. J., and R. C. Hider. 1983. Classification of phospholipases A$_2$ according to sequence: evolutionary and pharmacological implications. Eur. J. Biochem. **137**:545–551.

Felsenstein, J. 2001 PHYLIP (phylogeny inference package). Version 3.6. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Forstner, M. R. J., S. K. Davis, and E. Arévalo. 1995. Support for the hypothesis of anguimorph ancestry for the suborder Serpentes from phylogenetic analysis of mitochondrial DNA sequences. Mol. Phylogenet. Evol. **4**:93–102.

Fry B. G., J. C. Wickramaratna, W. C. Hodgson, P. F. Alewood, R. M. Kini, H. Ho, and W. Wüster W. 2002. Electrospray liquid chromatography/mass spectrometry fingerprinting of *Acanthophis* (death adder) venoms: taxonomic and toxinological implications. Rapid Commun. Mass Spectrom. **16**:600–608.

Fry, B. G., N. G. Lumsden, W. Wüster, J. Wickramaratna, W. C. Hodgson, and R. M. Kini. 2003a. Isolation of a neurotoxin (alpha-colubritoxin) from a 'non-venomous' colubrid: evidence for early origin of venom in snakes. J. Mol. Evol. **57**:446–452.

Fry, B. G., W. Wüster, R. M. Kini, V. Brusic, A. Khan, D. Venkataraman, and A. Rooney. 2003b. Molecular evolution and phylogeny of elapid snake venom three finger toxins. J. Mol. Evol. **57**:110–129.

Fry, B. G., W. Wüster, S. F. R. Ramjan, T. Jackson, P. Martelli, and R. M. Kini. 2003c. Analysis of Colubroidea snake venoms by liquid chromatography with mass spectrometry: evolutionary and toxinological implications. Rapid Commun. Mass Spectrom. **17**:2047–2062.

Gravlund, P. 2001. Radiation within the advanced snakes (Caenophidia) with special emphasis on African opisthoglyph colubrids, based on mitochondrial sequence data. Biol. J. Linn. Soc. **72**:99–114.

Greene H. W. 1997. Snakes: the evolution of mystery in nature. University of California Press, Berkeley, Calif.

Hayashi, M. A., A. F. Murbach, D. Ianzer et al. (12 co-authors) 2003. The C-type natriuretic peptide precursor of snake brain contains highly specific inhibitors of the angiotensin-converting enzyme. J. Neurochem. **85**:969–977.

Hayes, W. K., P. Lavín Murcio, and K. V. Kardong. 1992. Delivery of Duvernoy's secretion into prey by the brown tree snake, *Boiga irregularis* (Serpentes: Colubridae). Toxicon **31**:881–887.

Heinrikson, R. L., E. T. Krueger, and P. S. Keim. 1977. Amino acid sequence of phospholipase A$_2$-alpha from the venom of *Crotalus adamanteus*: a new classification of phospholipases A$_2$ based upon structural determinants. J. Biol. Chem. **252**:4913–4921.

Higuchi S., N. Murayama, K. Saguchi, H. Ohi, Y. Fujita, A. C. Camargo, T. Ogawa, M. Deshimaru, and M. Ohno. 1999. Bradykinin-potentiating peptides and C-type natriuretic peptides from snake venom. Immunopharmacology **44**:129–35.

Hill, R. E., and S. P. Mackessy. 2000. Characterization of venom (Duvernoy's secretion) from twelve species of colubrid snakes and partial sequence of four venom proteins. Toxicon **38**:1663–1687.

Ho, P. L., M. B. Soares, T. Yamane, and I. Raw. 1995. Reverse biology applied to *Micrurus corallinus*, a South American coral snake. J. Toxicol. Toxin Rev. **14**:327–337.

Ho, P. L., M. B. Soares, T. Maack, I. Gimenez, G. Puorto, M. F. Furtado, and I. Raw. 1997. Cloning of an unusual natriuretic peptide from the South American coral snake *Micrurus corallinus*. Eur. J. Biochem. **250**:144–149.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Version 3.0b4. Bioinformatics **17**:754–755.

Jackson, K. 2003. The evolution of venom-delivery systems in snakes. Zool. J. Linn. Soc. **137**:337–354.

Junqueira-de-Azevedo, I. L. M., and P. L. Ho. 2002. A survey of gene expression and diversity in the venom glands of the pitviper snake *Bothrops insularis* through the generation of expressed sequence tags (ESTs). Gene **299**:279–291.

Kamiguti, A. S, R. D. G. Theakston, N. Sherman, and J. W. Fox. 2000. Mass spectrophotometric evidence for P-III/P-IV metallloproteinases in the venom of the boomslang (*Dispholidus typus*). Toxicon **38**:1613–1620.

Kelly, C. M. R., N. P. Barker, and M. H. Villet. 2003. Phylogenetics of advanced snakes (Caenophidia) based on four mitochondrial genes. Syst. Biol. **52**:439–459.

Kloog, Y, I. Ambar, M. Sokolovsky, E. Kochva, Z. Wollberg, and A. Bdolah. 1988. Sarafotoxin, a novel vasoconstrictor peptide: phosphoinositide hydrolysis in rat heart and brain. Science. **242**(4876):268–270.

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

———. 1965. The development of the venom gland in the opisthoglyph snake *Telescopus fallax* with remarks on *Thamnophis sirtalis* (Colubridae, Reptilia). Copeia **1965**:147–154.

———. 1978. Oral glands of the Reptilia. Pp. 43–162 *in* C. K. Gans and A. Gans, eds. Biology of the Reptilia. Vol. 8. Physiology B. Academic Press, New York.

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

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**:1244–1245.

Lee, M. S. Y. 1997. The phylogeny of varanoid lizards and the affinities of snakes. Phil. Trans. R. Soc. Lond. B Biol. Sci. **352**:53–91.

Murayama, N., M. Hayashi, H. Ohi et al. (11 co-authors). 1997. Cloning and sequence analysis of a *Bothrops jararaca* cDNA encoding a precursor of seven bradykinin-potentiating peptides and a C-type natriuretic peptide. Proc. Natl. Acad. Sci. USA **94**:1189–1193.

Rieppel, O., H. Zaher, E. Tchernov, and M. J. Polcyn. 2003. The anatomy and relationships of *Haasiophis terrasanctus*, a fossil snake with well-developed hind limbs from the Mid-Cretaceous of the Middle East. J. Paleontol. **77**:536–558.

Rodríguez-Robles, J. A., and R. Thomas. 1992. Venom function in the Puerto Rican racer, *Alsophis portoricensis* (Serpentes: Colubridae). Copeia **1992**:62–68.

Ronquist F., and J. P. Huelsenbeck. MrBayes3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Salomão, M. G., and I. L. Laporta-Ferreira. 1994. The role of secretions from the supralabial, infralabial, and Duvernoy's glands of the slug-eating snake *Sibynomorphus mikani*

(Colubridae: Dipsadinae) in the immobilization of molluscan prey. J. Herpetol. **28**:369–371.

Slowinski, J. B., and R. Lawson. 2002. Snake phylogeny: evidence from nuclear and mitochondrial genes. Mol. Phylogenet. Evol. **24**:194–202.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Beta version 4.0b10. Sinauer Associates, Sunderland, Mass.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**(24):4876–4882.

Underwood, G. 1967. A contribution to the classification of snakes. British Museum (Natural History), London.

———. 1997. An overview of venomous snake evolution. Pp. 1–13 *in* R. S. Thorpe, W. Wüster and A. Malhotra, eds. Venomous snakes: ecology, evolution and snakebite. Symposia of the Zoological Society of London, No. 70. Clarendon Press, Oxford, UK.

Underwood, G., and E. Kochva. 1993. On the affinities of the burrowing asps *Atractaspis* (Serpentes: Atractaspididae). Zool. J. Linn. Soc. **107**:3–64.

Vidal, N. 2002. Colubroid systematics: evidence for an early appearance of the venom apparatus followed by extensive evolutionary tinkering. J. Toxicol. Toxin Rev. **21**:21–41.

Vidal, N., and S. B. Hedges. 2002. Higher-level relationships of caenophidian snakes inferred from four nuclear and mitochondrial genes. C. R. Biol. **325**:987–995.

Yamazaki, Y., H. Koike, Y. Sugiyama, K. Motoyoshi, T. Wada, S. Hishinuma, M. Mita, and T. Morita. 2002. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. Eur. J. Biochem. **269**:2708–2715.