J Mol Evol (2003) 57:110–129
DOI: 10.1007/s00239-003-2461-2

JOURNAL OF MOLECULAR EVOLUTION

Molecular Evolution and Phylogeny of Elapid Snake Venom Three-Finger Toxins

B.G. Fry,[1,2] W. Wüster,[3] R.M. Kini,[2] V. Brusic,[4] A. Khan,[4] D. Venkataraman,[4] A.P. Rooney[5]

[1] Australian Venom Research Unit, Department of Pharmacology, University of Melbourne, Parkville, Vic 3010, Australia
[2] Department of Biological Sciences, National University of Singapore, Singapore 119260
[3] School of Biological Sciences, University of Wales, Bangor LL57 2UW, Wales, UK
[4] BioDiscovery Group, Laboratories for Information Technology, 21 Heng Mui Keng Terrace, Singapore 119613
[5] Mississippi State University, Department of Biology, P.O. Box GY, Mississippi State, MS 39762, USA

Received: 27 August 2002 / Accepted: 18 February 2003

Abstract. Animal venom components are of considerable interest to researchers across a wide variety of disciplines, including molecular biology, biochemistry, medicine, and evolutionary genetics. The three-finger family of snake venom peptides is a particularly interesting and biochemically complex group of venom peptides, because they are encoded by a large multigene family and display a diverse array of functional activities. In addition, understanding how this complex and highly varied multigene family evolved is an interesting question to researchers investigating the biochemical diversity of these peptides and their impact on human health. Therefore, the purpose of our study was to investigate the long-term evolutionary patterns exhibited by these snake venom toxins to understand the mechanisms by which they diversified into a large, biochemically diverse, multigene family. Our results show a much greater diversity of family members than was previously known, including a number of subfamilies that did not fall within any previously identified groups with characterized activities. In addition, we found that the long-term evolutionary processes that gave rise to the diversity of three-finger toxins are consistent with the birth-and-death model of multigene family evolution. It is anticipated that this ''three-finger toxin toolkit'' will prove to be useful in providing a clearer picture of the diversity of investigational ligands or potential therapeutics available within this important family.

Key words: Venom — Three-finger toxin — Multigene family — Elapidae

Correspondence to: Dr. Bryan Grieg Fry;
email: bgf@unimelb.edu.au

Introduction

Many venom components are invaluable in molecular, biochemical, and biomedical research due to their specificity and potency. The variation in the biochemical composition of snake venom occurs between closely related species or even within a species itself (Jiménez-Porras 1964; Glenn et al. 1983; Yang et al. 1991; Assakura et al. 1992; Daltry et al. 1996; Fry et al. 2002). The great diversity of snake venom toxins is due to their mode of evolution, which is subject to frequent duplication of toxin-encoding genes that is sometimes followed by functional and structural diversification (Moura-da-Silva et al. 1995; Slowinski et al. 1997; Afifiyan et al. 1999; Chang et al. 1999; Kordis and Gubensek 2000) and accelerated rates of sequence evolution (e.g., Kini and Chan 1999; Nakashima et al. 1995). This diversification is possibly a result of selection for the ability to kill and digest different prey (e.g., Daltry et al. 1996) or as part of a predator–prey arms race (e.g., Poran et al. 1987; Heatwole and Poran 1995). Thus, a common theme in venom evolution is a multiplicity of toxins

with different actions that are encoded by multigene families.

Understanding the evolution of snake toxin multigene families has practical as well as theoretical applications. For example, an understanding of how a toxin multigene family evolves, coupled with a knowledge of the species' systematics and natural history, can help predict the occurrence of toxins in taxonomic groups whose venom has not been biochemically characterized. In addition, such an approach can predict the likely activity of toxins that are rooted among other toxins with better-characterized activities. Such studies might also highlight evolutionary isolated toxins that might have novel modes of action and would, therefore, be of special interest as investigational ligands.

The three-finger toxins of elapids (sea snakes and cobras) form a broad superfamily of nonenzymatic polypeptides. We became interested in the three-finger toxin family of snake venom peptides because (1) they encompass a large variety of toxins with different functional activities and are therefore interesting from a molecular evolutionary perspective, and (2) they are of interest to a wide range of biochemical and biomedical researchers. The members of this multigene family contain 60–74 amino acid residues and are rich in disulfide bonds, with four such bonds being conserved in all family members (Endo and Tamiya 1987). All proteins in this family, therefore, have a similar pattern of protein folding that consists of three loops extending from a central core containing the four conserved disulfide bridges (e.g., Ménez 1998; Tsetlin 1999) resulting in an uncanny resemblance to three fingers, hence the name "three-finger" toxin. Despite their overall similarity in structure, these polypeptides differ from each other in their biological activities. The endogenous three-finger peptides of vertebrates that play a significant role in cell–cell adhesion may be the ancestors of the three-finger toxins (Fleming et al. 1993; Gumley et al. 1995). Related peptides are used in the complement system (CD59) and lymphocytes (Ly6) and are also secreted in the brain (Lynx1). Due to the intensive use of the snake venom three-finger toxins as investigational ligands in biomedical and biochemical research, a large number have been characterized and sequenced, making this class of toxins particularly valuable for molecular evolutionary studies. Understanding the evolutionary mechanisms generating the variety of three-finger toxins is important from the perspective of biomedical researchers who wish to characterize the diverse functional activities of these toxins. Therefore, the aim of this study is to understand the long-term evolutionary processes that resulted in the structural diversification of three-finger toxins and to provide a phylogenetic framework for the investigation of these proteins, which may also guide the search for novel toxins with activities of particular interest.

Materials and Methods

We analyzed 276 three-finger toxin amino acid sequences from snakes in the family Elapidae. All sequences used in this study were obtained from SWISS-PROT/TreEMBL (http://www.expasy.org/sprot) except for several sequences that were obtained from the literature: Type A muscarinic toxins ml toxin 2 (Carsi and Potter 2000) and MT5 (Jolkkonen 1996), Type B muscarinic toxin (Carsi et al. 1999), and bulongin (Kini et al. unpublished results). To simplify sequence nomenclature and minimize confusion, we refer to toxins by their accession numbers in the text (Table 1). We used the program CLUSTAL-X (Thompson et al. 1997) to align the sequences, followed by visual inspection of the resultant alignment for errors. The final alignment consisted of 123 amino acid sites. The 75% consensus sequences were determined using Consensus (http://www.bork.embl-heidelberg.de:8081/Alignment/consensus.html). A copy of the full sequence alignment can be obtained by emailing the first author.

Phylogenetic trees were reconstructed using the maximum parsimony (MP) and neighbor-joining (NJ) (Saitou and Nei 1987) methods. Due to the large number of taxa in our study, we conducted our phylogenetic analyses in two steps. First, both MP and NJ trees were constructed to test for congruent clustering patterns on the basis of both topology and the reliability in interior branches as assessed by bootstrap values. In this manner, we identified clades of interest that could be further analyzed in more detail. Once such clades were identified, they were analyzed separately using both MP and NJ methods. MP heuristic searches were conducted by implementing 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. NJ searches were conducted using amino acid p distances, as the simple p distance generally gives better results in phylogenetic inference than more complicated distance measures for minimum evolution methods such as NJ (Takahashi and Nei 2000). Statistical reliability was assessed using 100 and 1000 bootstrap replications for MP and NJ searches, respectively.

The results of our phylogenetic analyses were used to classify the three-finger toxins into groups on the basis of their phylogenetic relationships and their demonstrated mode of action, as far as known. In all cases, the LY-6 sequences Q14210 and P35459 were utilized as outgroup taxa. All analyses were performed using the computer program PAUP* (Swofford 2002).

To calculate the number of events of gene loss and gene duplication, we used the gene tree parsimony approach, implemented in the program GeneTree version 1.3.0 (Page 2001). The aim of the method is to reconcile the gene tree with the organismal tree in a manner requiring the fewest assumptions of gene duplication and gene loss. The gene tree used was the NJ tree obtained as above, because the GeneTree software requires a fully resolved, dichotomous tree. The use of the gene tree parsimony method requires an organismal tree for the species at hand. The phylogeny of the elapid snakes has been investigated by a number of researchers (Slowinski et al. 1997; Keogh 1998; Keogh et al. 1998; Slowinski and Keogh 2000), but there is as yet no comprehensive, robustly supported phylogenetic hypothesis for the entire family, and no analysis has ever included all the taxa from which we have obtained toxin sequences for this paper. Still, to provide background relevant to the three-finger toxin multigene family analyses, we thought it would be helpful to show a putative species tree for the taxa included in this study (Fig. 1). We drew this tree on the basis of the ML tree of Slowinski and Keogh (2000). Taxa not represented in our toxin database were pruned from the tree, and taxa represented in our

Table 1. Swiss-prot accession numbers for components in each group

Acn-esterase inhibitors	P01403, P25681
Antiplatelet toxins	P28375, P81946, P01413
L-Type Ca^{2+} blockers	P01414, Q9PS74, P22947, P25684, P25683
Type IA cytotoxins	P01467, P01468, P01469, P01470, P25517, P01459, P01460, P01457, P01455, P01456, P01465, P01466, P01462, P01464, P01463, P01461, P01458, P01442, P01443, Q98961, P07525, P01445, P01446, P01447, P24780, Q9PSN2, P01453, P01454, P01448, P01452, Q98960, Q9DGH9, Q98959, Q9PST3, O93472, Q9PST4, P01440, O73856, O93473, Q02454, P01444, Q9PS23, O93471, Q91125, Q98958, P01449, Q91135, Q91136, Q98957, Q9PS24, P01451, Q98956, Q98962, P24779, P49124, P80245, P49123, Q98965, P01441
Type IB cytotoxins	P01471, P24776, P24777
Type I α-ntx	P01420, P01421, P01422, P01423, P25675, P01424, P01429, P01425, P01433, P80958, P01427, P01428, P01426, P34076, P01430, Q9PTT0, Q9PSN6, P14613, Q9DE57, Q9W6X0, O57326, O57327, P82849, P01431, P01432, P34075, P19958, P19960, P32879, P19959, P25494, P25493, P01438, P25492, P01437, P25497, P01434, P10457, Q9PRJ0, Q9PRJ7, Q9PRJ5, Q9PWJ4, P10455, Q9PRJ6, P10456, Q9YGC7, Q9YGX0, P10458, P10460, Q9YGC2, Q9YGC4, P25495, P10459, Q9PRJ3, Q9YGW8, Q9YGW9, P25496, P01435, P01419, P01417, P01418, P01416, P80548, P01412, P10808
Type II α-ntx	P01385, P34073, P01384, P01378, Q9PRI6, Q9YGD1, Q9PRI7, Q9W726, P80965, P80156, P01383, P01382, P01394, P01395, P01393, P01397, P25667, P01396, P01386, P01387, P07526, P25674, P01390, P01389, P01388, P25672, P25673, P25668, P25669, P25671, P01391, O42257, P13495, Q9W7J5, P01381, P01380, P14612, P25670, P82662, P01379, O93496, P15815
Type III α-ntx	Q9W7J9, Q9W7K2, Q9W7J8, Q9W7K0, Q9W7K1, Q9W7J6, Q9W7J7
κ-neurotoxins	O12961, O12962, P15816, P01398, P15817, Q9W729
Type A muscarinic toxins	P80970, m 2, P80495, P25518, MT5, P81031, Q9PSN1, P81030, P80494, P18328
Type B muscarinic toxins	Carsi et al. (1999)
Type C muscarinic toxins	P82462, P82463
Synergistic toxins	P01408, P01409, P01410, P01411, P01407, P17696
Orphan group I	Q9W717
Orphan group II	O93422, Q9YGI2, Q9YGI1, P29180, P01401, P01400, P29179, P01399, P25680, P29181, P29182, Q9YGI4, Q9YGI7, Q9W7I3, Q9W7I4, O42256, O42255, P25679, P82935
Orphan group III	Q9YGI8, P83346
Orphan group IV	P15818, P81783, Q9YGI9
Orphan group V	Q9YGJ0, O12963, Q9YGH9
Orphan group VI	P25676, P01415
Orphan group VII	P43445
Orphan group VIII	Q9W727, Q9DEQ3, P82464
Orphan group IX	Q9YGI0, bucoxin, P79688
Orphan group X	P18329
Orphan group XI	P01406, P01405, P01404
Orphan group XII	Q9PRI1, Q9PUB7
Orphan group XIII	P19004
Orphan group XIV	P24778
Orphan group XV	Q91996, P14554, Q91126, Q91137, Q9W716, P14541, P01472, P49122, P01473, P01474
Orphan group XVI	P19003
Orphan group XVII	Q9YGH0, Q9PW19, bulongin
Orphan group XVIII	P14534
Orphan group XIX	P81782, P25682
Orphan group XX	P25677, P25678, P01402

database but not the original tree were grafted into the tree based on literature data (Slowinski 1994, 1995; Slowinski et al. 1997; Keogh 1998; Keogh et al. 1998) and our own data (*Dendroaspis*, *Naja*).

Results

Three-Finger Phylogeny

Both MP (Fig. 2) and NJ (Fig. 3) trees for all the entire set of sequences were highly congruent with respect to group-level composition. In all gene trees, the conventionally recognized, major functional groups of toxins, characterized by activity type and specific functional motifs, formed monophyletic groups. Bootstrap support for most gene clades was low, most likely as a result of the short length of the toxin sequences and the high number of alignment gaps in the amino acid sequences. However, in addition to the clades of toxins of known function, our analysis identified 20 distinct clades of toxins lacking specific functional motifs but having unique 75% consensus sequences (Fig. 4), for which the biological activity remains unknown (Table 2). These groups were termed *orphan groups* and numbered. This designation is intended to be temporary and should be replaced as soon as functional data are determined.

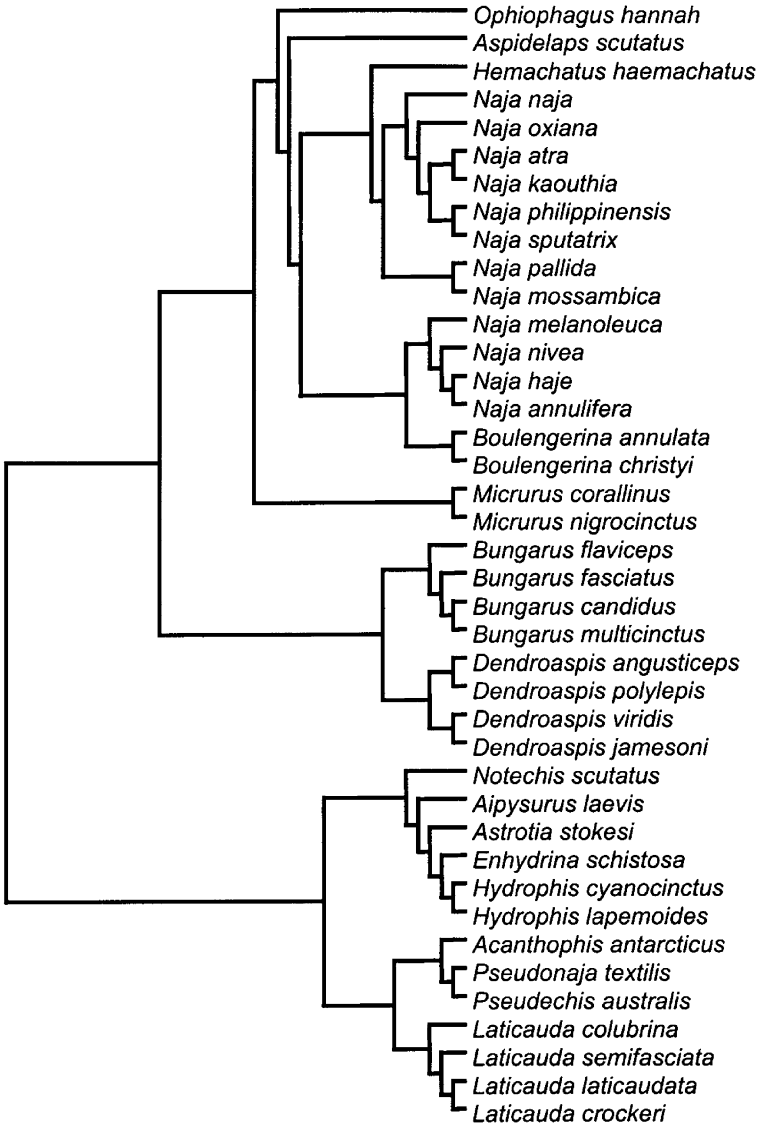

Fig. 1. Putative species tree of members of the family Elapidae included in this study. The tree was derived from Slowinski and Keogh (2000). Note, in particular the basal split between an Asian–African–American group and an Australasian/marine group. These are often recognized as the subfamilies Elapinae and Hydrophiinae, respectively.

Major clades were subjected to further analyses for a more thorough investigation of relationships of the individual toxins (Figs. 5–9).

The use of the gene tree parsimony approach shows that reconciling the gene tree with the organismal tree requires 201 assumptions of gene duplication and 516 assumptions of gene loss or, alternatively, 586 incidences of deep coalescence.

Discussion

Patterns of Multigene Family Evolution

The analyses presented here show that the three-finger toxins evolve through a process of gene duplication, and shifts in protein function are normally associated with gene duplication events. Our gene tree parsimony analysis conservatively estimated that mapping the gene tree revealed by our analyses onto the putative species tree (Fig. 1) would require 201 assumptions of gene duplication and 516 assumptions of gene loss. While the latter figure is most likely due to inadequate sampling in many elapid taxa (see caveats at the end of this section), the former is almost certainly an underestimate, for the same reason. Nevertheless, our data thus show very clearly the importance of gene duplications in the evolution of this toxin gene family.

The patchy sampling of elapid toxin sequences also impeded attempts to use the toxin gene phyelogeny to gain additional understanding of the organismal phylogeny of the elapids. Attempts to use GeneTree to infer the organismal tree requiring the fewest assumptions of gene duplication were aborted after the program-specified maximum of 15,000 trees equally most parsimonious was identified. A strict consensus revealed only one resolved note, which placed *Micrurus corallinus* as the sister taxon to all other elapids (including *M. nigrocinctus*), a result entirely inconsistent with published data (Slowinski

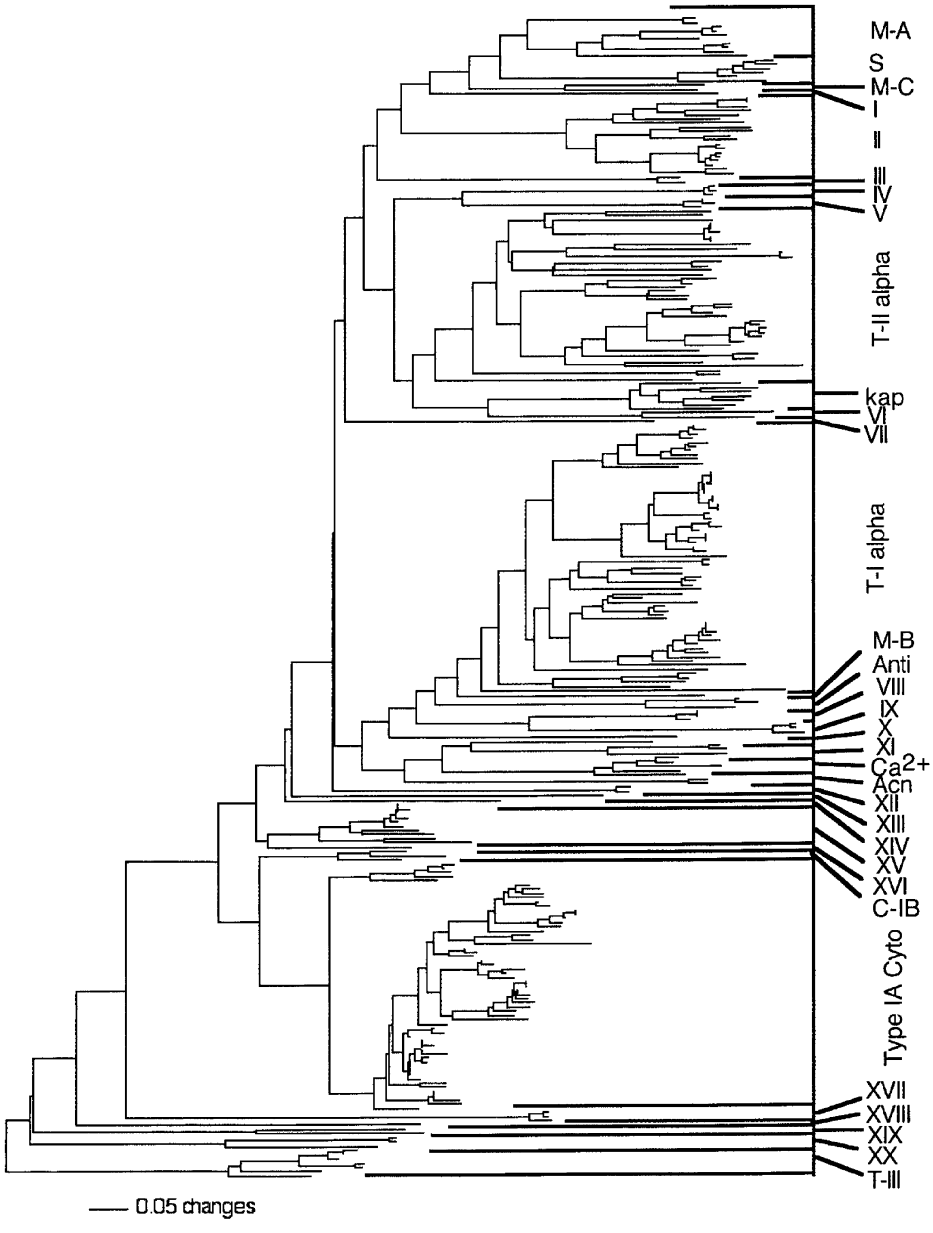

Fig. 2. NJ tree for the three-finger toxin superfamily. Groups are Type A muscarinic toxins (M-A), Type B muscarinic toxins (M-B), Type C muscarinic toxins (M-C), synergistic toxins (S), Type I α-neurotoxins (Type I α). Type II α-neurotoxins (Type II α), κ-neurotoxins (kap), Antiplatelet toxins (anti), L-type calcium channel blocking toxins (L), acetylcholinesterase inhibiting toxins (Acn), Type IA cytotoxins (Type IA cyto), Type IB cytotoxins (C-B), Type III α-neurotoxins (T-III), and orphan groups I–XX. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

1995; Slowinski et al. 2001) but easily explainable due to the fact that the only toxins sampled for this species were two highly divergent peptides (orphan group XII).

One of the most conspicuous findings of this study is that a substantial proportion of sequenced three-finger toxins in the Elapidae belongs to clades with as yet largely unknown functional properties. These orphan groups were defined through comparison of 75% consensus sequences (Fig. 4), physical properties, and presence/absence of known functional motifs. No fewer than 20 such orphan groups, containing 67 individual toxins, were identified in this study. Since past sequencing efforts are likely to have been biased in favor of toxins with known biological activities, it seems likely that these orphan-group toxins are underrepresented in the database analyzed here. The orphan groups are of interest in that they may potentially contain toxins of novel and potentially interesting modes of action, which may be of interest from a pharmacological point of view or as investigational tools. The accession numbers for each group are listed in Table 1.

In light of the increased diversity of the α-neurotoxins, the well-characterized "short-chain" and "long-chain" groups were renamed Type I α-neurotoxins and Type II α-neurotoxins, respectively. The new group of α-neurotoxins from the genus *Pseudonaja* (Gong et al. 1999) we then designated Type III α-neurotoxins. For example, initial evidence suggests that at least one of the toxins in orphan group IV (P81783) may be a reversible neurotoxin

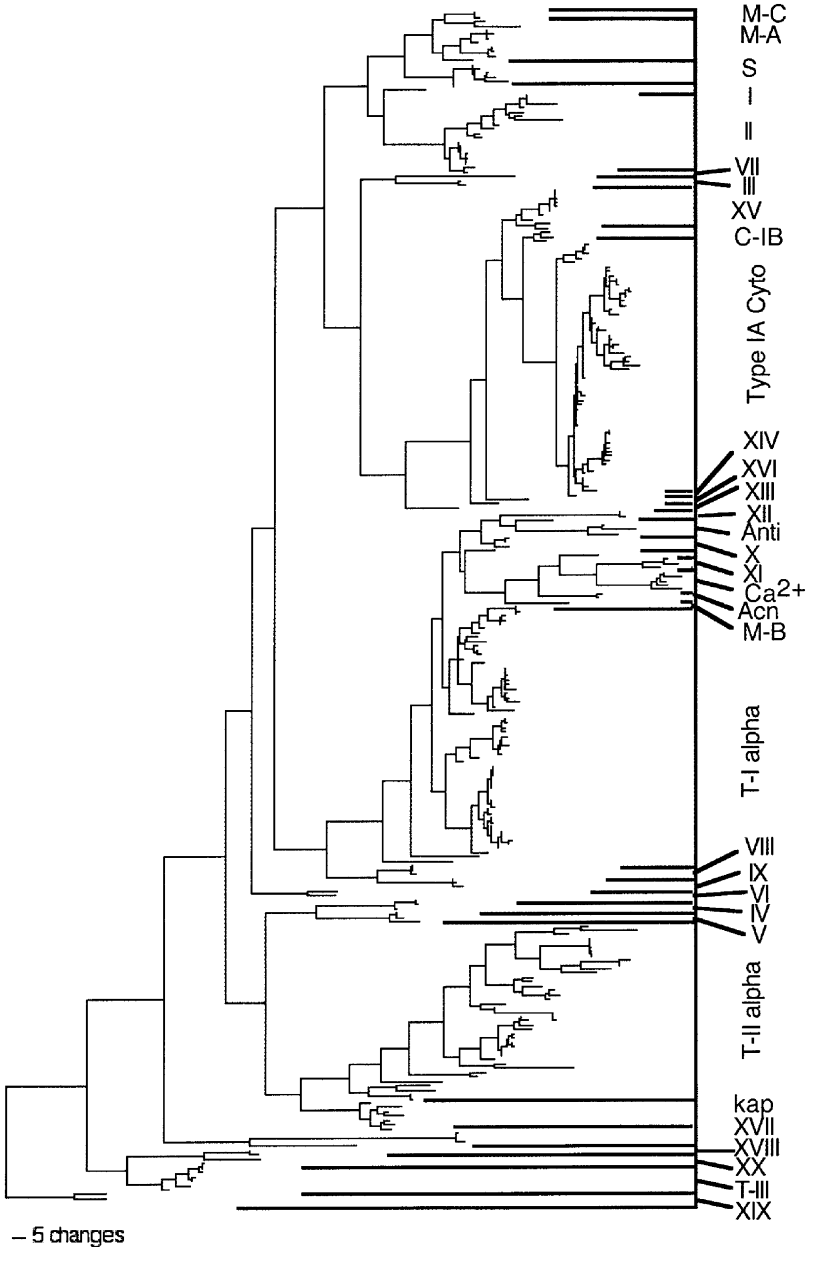

Fig. 3. MP tree for the three-finger toxin superfamily. See Fig. 1 for group labelling and composition. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

(Nirthanan et al. 2002), and thus this entire group may ultimately be designated the Type IV neurotoxins.

The different toxin clades identified in this study vary considerably in the taxonomic breadth of their distribution (Table 3). Some toxin groups have representatives in many of the genera examined here, such as the Type I and II α-neurotoxins. These groups obviously emerged quite early during elapid evolutionary history. For example, the division between terrestrial Australian elapids and sea snakes and terrestrial African and Asian elapids is quite ancient (Slowinski et al. 1997; Slowinski and Keogh 2000) and may represent the most basal division within the Elapidae, yet the members of both clades possess Type I and Type II α-neurotoxins.

However, it is also evident that some toxin groups are restricted to specific organismal clades, such as the Type IB cytotoxins of *Hemachatus* or the fasciculins of *Dendroaspis* species. Furthermore, taxon-specific toxin clusters sometimes form within functionally uniform toxin clades, such as the Type I α-neurotoxins of *Laticauda* species. There are three potential scenarios that could produce this pattern: (1) the toxins emerged prior to the divergence of the taxa in which they are currently found and were lost in the other lineages but remained in the current lineage in which they are found; (2) the toxins emerged subsequent to the divergence of the taxa in which they are found and are unique to those taxa; or (3) the toxins are present in other genera but have not been sequenced yet. Under the first scenario, the toxins will

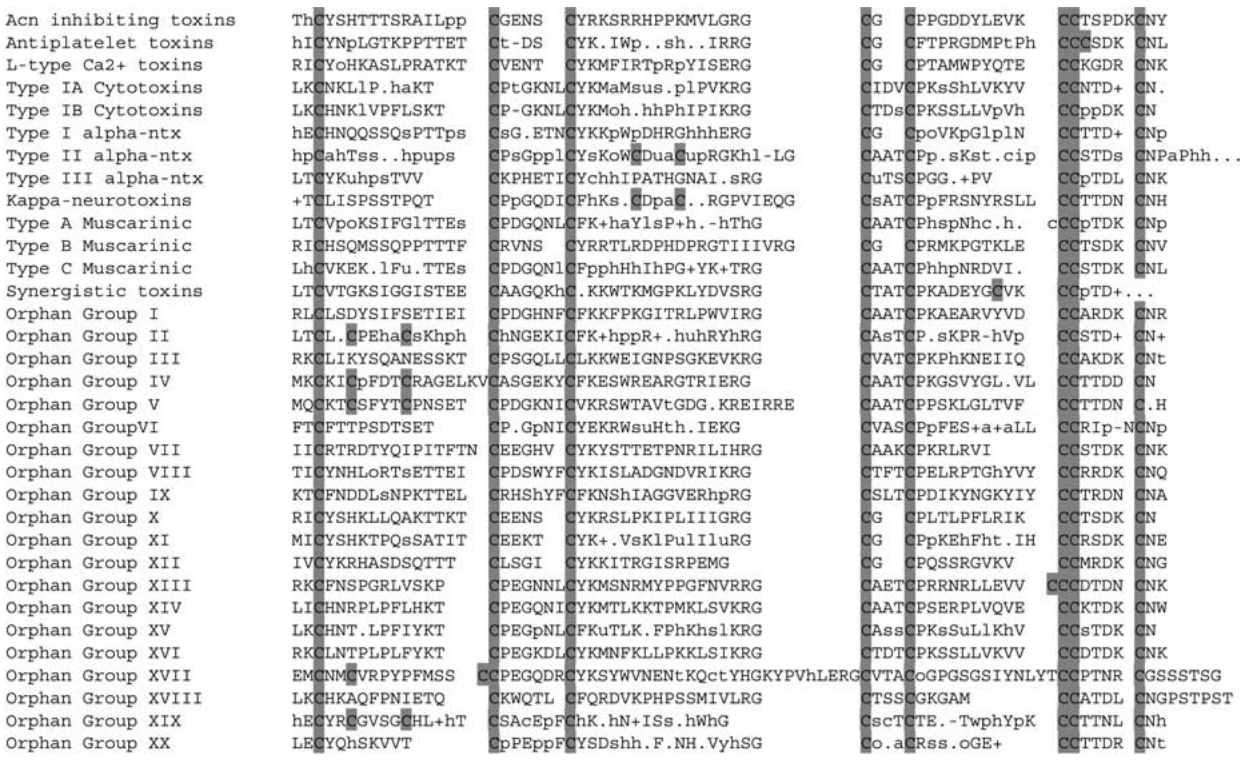

Fig. 4. Cysteine-aligned 75% consensus sequences of the toxins present within each group. Alcohol = o = S, T; aliphatic = l = I, L, V; any = . = A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; aromatic = a = F, H, W, Y; charged = c = D, E, H, K, R; hydrophobic = h = A, C, F, G, H, I, K, L, M, R, T, V, W, Y; negative = – = D, E; polar = p = C, D, E, H, K, N, Q, R, S, T; positive = + = H, K, R; small = s = A, C, D, G, N, P, S, T, V; tiny = u = A, G, S; turnlike = t = A, C, D, E, G, H, K, N, Q, R, S, T. Cysteines are highlighted in gray.

appear to be more divergent from one another and would most likely occupy an isolated position in the gene tree, because they would have evolved prior to the divergence of the species under consideration. Moreover, one would expect to find traces of these families in the shape of pseudogenes in the genomes of taxa in which the toxins are not presently expressed. Under the second scenario, we would expect toxins not to have diverged extensively since they would have evolved recently, and they should be rooted among other functionally and structurally similar toxins from related taxa, and homologous pseudogenes would be absent in other groups. However, even under the second scenario the toxins could be extremely divergent if the split was ancient (e.g., the genus *Dendroaspis* splitting off from the other elapids), and in such cases, it is only the presence or absence of relictual pseudogenes that can provide evidence that will discriminate between scenario 1 and scenario 2. Importantly, given the lack of study of the venoms of many genera of elapid snakes, the third scenario cannot be excluded for many groups of toxins or, more importantly, for many elapid taxa for which few sequences are available.

On the basis of the clustering patterns shown in the phylogenetic trees, the first scenario best explains the emergence of taxon-specific toxin groups (e.g., the Type I cytotoxins of *Naja*). The second scenario is favored in cases where taxon-specific clusters have emerged within groups. For example, the synergistic toxins of *Dendroaspis* appear to have evolved after *Dendroaspis* split from other terrestrial African and Asian elapids. In fact, their closest relatives are the Type A muscarinic toxins of *Dendroaspis*. Similarly, the large subclade of *Laticauda*-specific Type I α-neurotoxins almost certainly evolved in this manner. The third scenario is favored where a group of toxins has been found in a few phylogenetically extremely divergent groups, as in the case of the muscarinic toxins, in which toxins affecting this receptor have been found in the very divergent *Dendroaspis* and *Naja* venoms. Thus it is likely that muscarinic toxins are present in other elapid venoms, being a case of inadequate sampling rather than genuine absence.

All these patterns are possible under the birth-and-death mode of multigene family evolution (Nei et al. 1997; Rooney et al. 2002). According to this process, gene families are created through the process of gene duplication. Over time, some genes get deleted from the genome, through processes such as unequal crossing-over, while some become nonfunctional and degenerate into pseudogenes. As a result, paralogous groups of genes are generated across taxonomic lines if the gene duplication events giving rise to these

Table 2. Functional activity of each group

Acn-esterase	Potent acetylcholinesterase inhibitors
Antiplatelet toxins	Competatively bind to platelet GPIIb/IIIa receptor
L-Type Ca^{2+} blockers	Block L-type calcium channels
Type IA cytotoxins	Activities include cadiotoxic and cytotoxic effects
Type IB cytotoxins	Cardiotoxic, hemolytic and hypotensive
κ-neurotoxins	Antagonistes of α3β2 neuronal nicotinic acetylcholine receptor subtype
Type A muscarinic toxins	Bind to muscarinic acetylcholine receptors
Type B muscarinic toxins	Bind to muscarinic acetylcholine receptors
Type C muscarinic toxins	Bind to muscarinic acetylcholine receptors
Type I α-neurotoxins	Antagonists of α1 neuromuscular nicotinic acetylcholine receptor subtypes
Type II α-neurotoxins	Antagonists of α1 and α7 neuromuscular nicotinic acetylcholine receptor subtypes
Type III α-neurotoxins	Neuromuscular nicotinic acetylcholine receptor antagonists
Synergistic toxins	Alone are nonactive but form complexes with α-neurotoxins to enhance neurotoxicity
Orphan group I	Unknown
Orphan group II	Unknown; some low-level α-neurotoxic activity evident
Orphan group III	Unknown
Orphan group IV	Unknown
Orphan group V	Unknown
Orphan group VI	Unknown
Orphan group VII	Unknown
Orphan group VIII	Unknown
Orphan group IX	Unknown
Orphan group X	Unknown
Orphan group XI	Unknown
Orphan group XII	Unknown
Orphan group XIII	Unknown
Orphan group XIV	Unknown
Orphan group XV	Unknown; some low-level cytotoxic activity
Orphan group XVI	Unknown
Orphan group XVII	Unknown
Orphan group XVIII	Weak and reversible neuromuscular nicotinic acetylcholine receptor antagonists
Orphan group XIX	Unknown
Orphan group XX	Unknown

groups took place before their divergence. This is what is observed in cases 1 and 2. Searches for pseudogenes associated with toxins that are not expressed in a particular species or taxonomic group would be revealing in this context and would constitute a useful test of the birth-and-death model of gene evolution.

According to case 3, recent gene duplication produces a cluster/group of toxins, which can explain why they appear to be closely related. Of course, a broader taxonomic sampling may help to refine this picture somewhat. Nevertheless, these birth-and-death patterns explain why a number of groups of toxins are restricted to genera representing relatively long-isolated lineages (see Slowinski and Keogh [2000] and reference therein). For example, κ-neurotoxins, orphan groups III, IV, VI, VII, VIII, XIII, XVI, and XXI, are restricted to the kraits (*Bungarus*); moreover, where kraits have been shown to have proteins belonging to groups also found in other genera (e.g., Type I neurotoxins), the toxins derived from *Bungarus* form their own, separate monophyletic group. The same applies to the mambas (*Dendroaspis*): synergistic, antiplatelet toxins, L-type calcium channel blockers, the sole Type B muscarinic toxin, and orphan groups XI and XII are unique to this genus, and mamba toxins within other groups tend to form discrete clades (e.g., Type I α-neurotoxins). Indeed, the mamba Type I α-neurotoxins are not resolved well within this group and may represent a subgroup. Similarly, within the Type I α-neurotoxins, an entire toxin clade is made up of *Laticauda* toxins, and another of sea snake neurotoxins along with Australian terrestrial snake toxins, which is consistent with at least some phylogenetic hypotheses about this group (e.g., Keogh 1998). The Type III α-neurotoxins appear to be unique to the genus *Pseudonaja*, as extensive LC/MS analysis has not revealed the presence of components within this mass range (6100–6300 daltons) in the venom of other members of the Hydrophiinae (Fry et al. 2002; B.G. Fry, unpublished results).

Thus, our study of three-finger toxin evolution shows that a birth-and-death model best describes the evolution of this large multigene family. The three-finger toxins were recruited into the venom proteome of the elapid snakes early, before the divergence of even the most basal clade of extant elapid snakes, and diverged early to form a broad superfamily. However, this superfamily continues to diversify, as shown

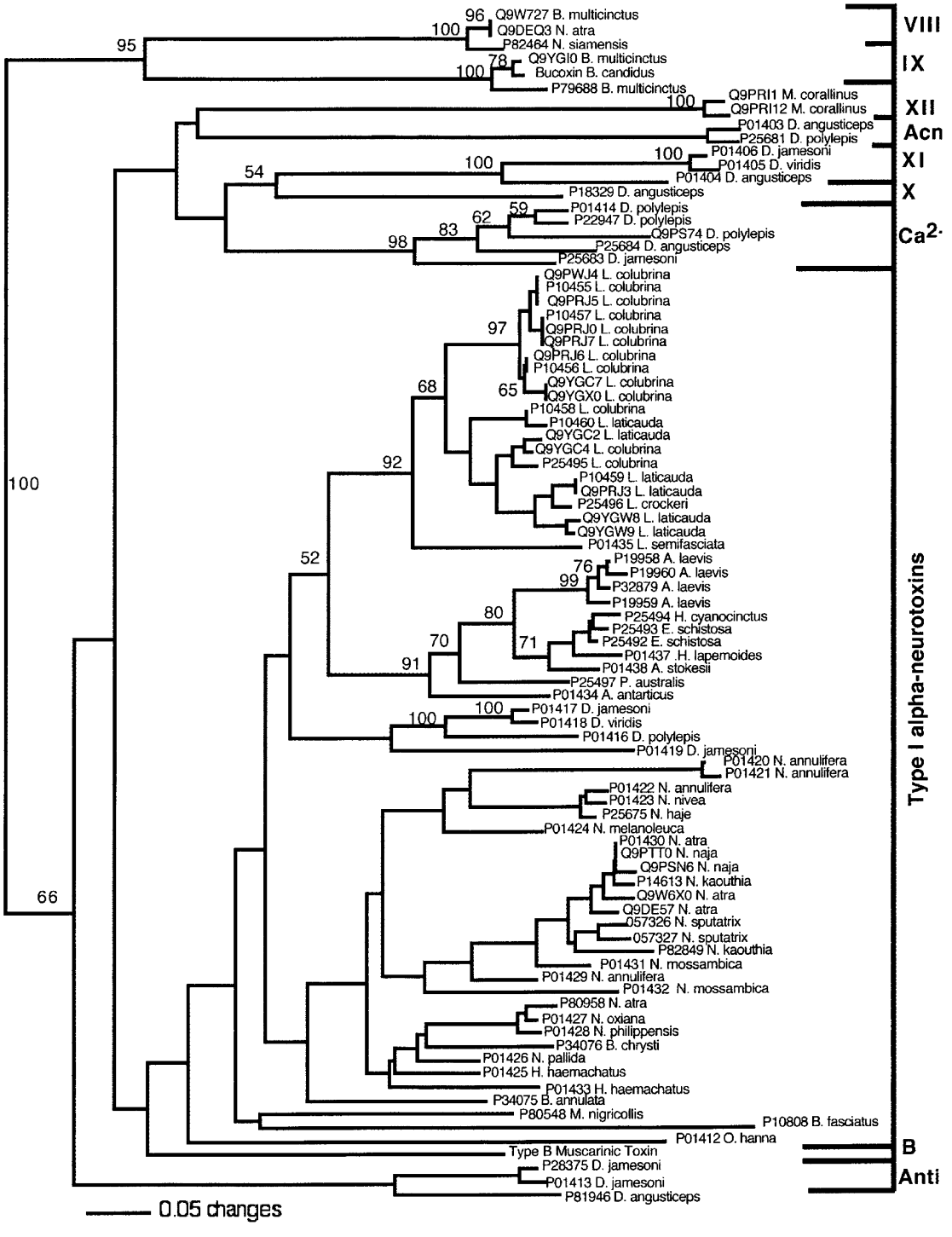

Fig. 5. NJ tree for Type I α-neurotoxins, acetylcholinesterase inhibiting toxins (Acn), antiplatelet toxins (Anti), L-type calcium channel toxins (L), Type B muscarinic toxins (B), and orphan groups VIII, IX, X, XI, and XII. Bootstrap values are the result of 1000 replicates. Only bootstrap values 50% or greater are shown. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

by our finding of taxon-specific gene clusters. These evolutionary patterns are similar to what has been observed in multigene families involved in the adaptive immune response (e.g., immunoglobulins and major histocompatibility complex genes [Nei et al. 1997]). It is believed that gene duplication and subsequent divergence contribute to an organism's ability to react to a wide range of foreign antigens. In an analogous manner, snake toxins must react with diverse compounds in their prey. Thus, a birth-and-death mode of evolution may generate a suite of toxins to allow snake predators to adapt to a variety

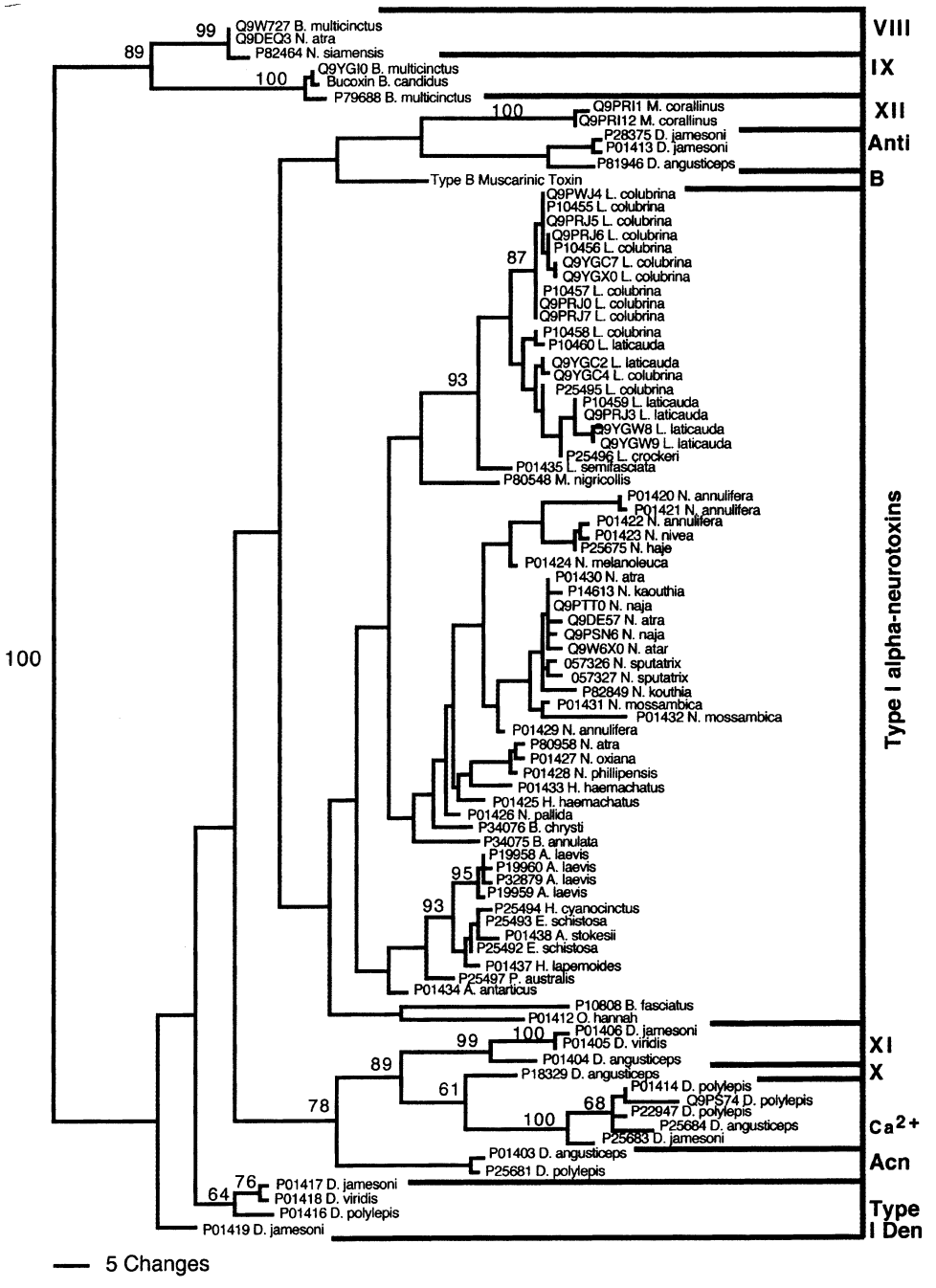

Fig. 6. MP tree for Type I α-neurotoxins (*Dendroaspis* toxins are labeled Type I Den), acetylcholinesterase inhibiting toxins (Acn), antiplatelet toxins (Anti), L-type calcium channel toxins (L), Type B muscarinic toxins (B), and orphan groups VIII, IX, X, XI, and XII. Only bootstrap values 50% or greater are shown. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

of different prey species. We note that another snake toxin multigene family, the phospholipase A_2's, appear to show evidence of birth-and-death evolution (Slowinski et al. 1997), although a more thorough analysis is needed to confirm this.

Inferred Structure–Function Relationships

While significant differences are evident in the 75% consensus sequences (Figure 4), a level of conservation of overall physical characteristics is evident among the three-finger toxins (Table 4). The groups range in size from 57 amino acids (orphan group XII) to 82 (orphan group XVII), but with the vast majority of the groups being between 60 and 65 residues. The p*I*'s are usually slightly basic but range from acidic (orphan group VI) to strongly basic (cytotoxins and orphan groups III, X, XV and XVI).

The number of cysteines, and thus disulfide bonds, is also quite conserved, with the majority containing

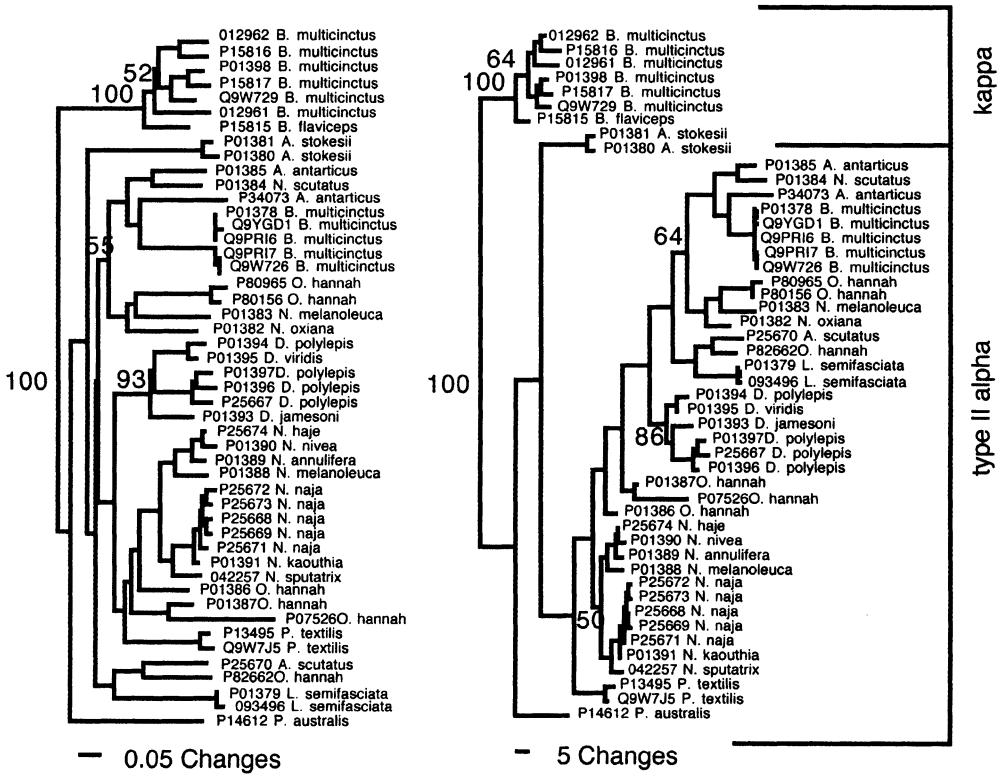

Fig. 7. Results of (**A**) NJ and (**B**) MP analysis of Type II α-neurotoxins and κ-neurotoxins. Only bootstrap values 50% or greater are shown. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

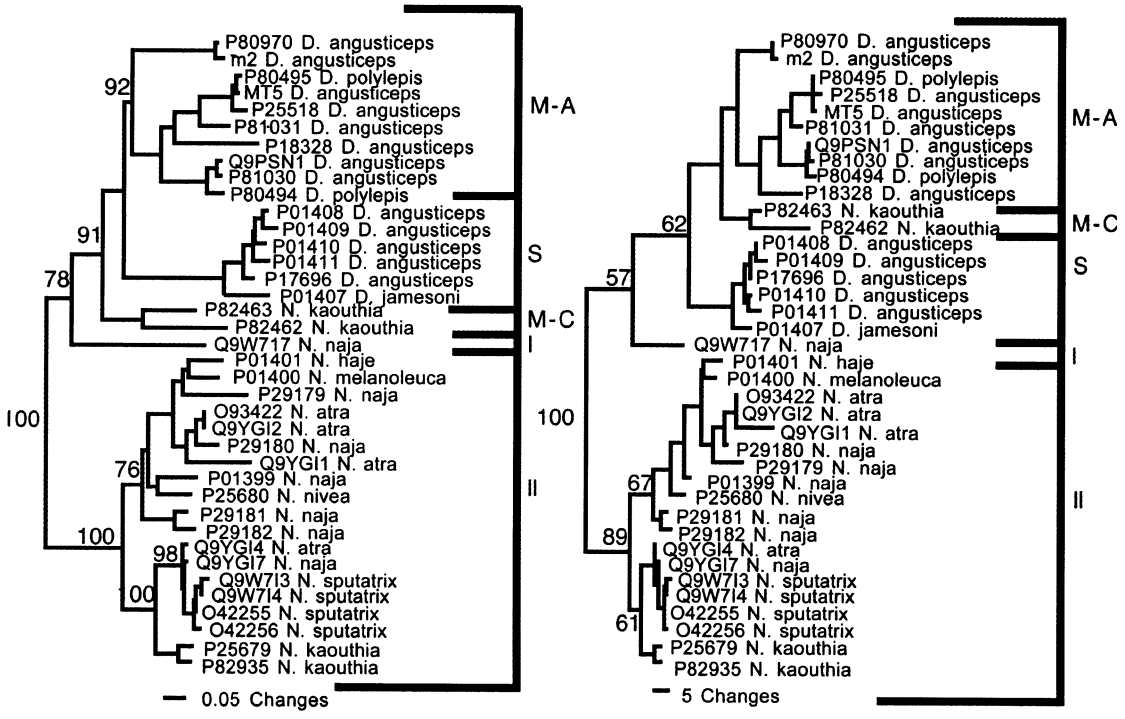

Fig. 8. Results of (**A**) NJ and (**B**) MP analysis of Type A muscarinic toxins (M-A), synergistic toxins (S), Type C muscarinic toxins (M-C), and orphan groups I and II. Only bootstrap values 50% or greater are shown. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

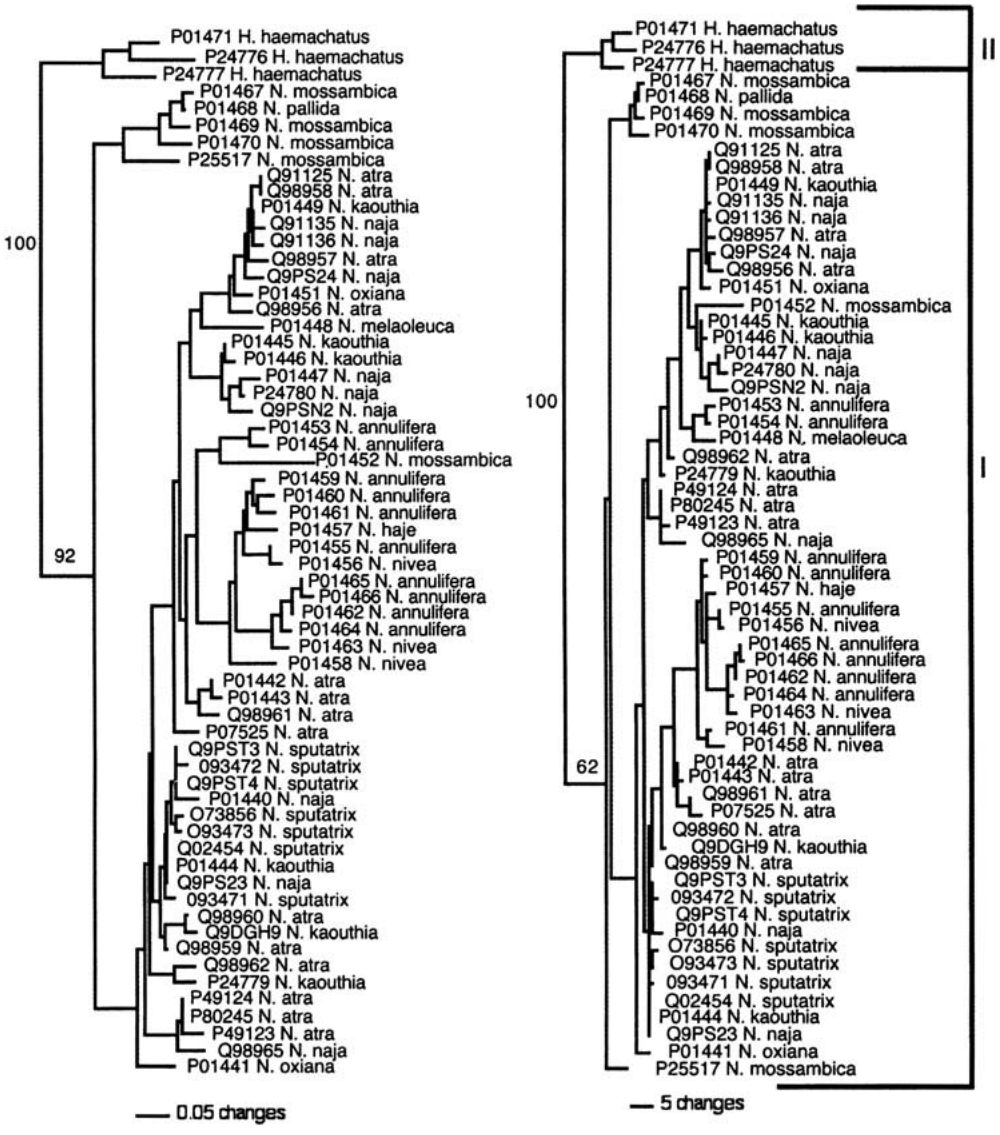

Fig. 9. Results of (**A**) NJ and (**B**) MP analysis of Type IA and Type IB cytotoxins. Only bootstrap values 50% or greater are shown. Outgroup sequences (Q14210 and P35459) were removed from the final tree image, although they were included in the analysis to root the phylogeny.

eight cysteines and four disulfide bonds. Only 6 of the 34 groups contain 10 cysteines and five disulfide bonds. However, 10 cysteines is the ancestral condition as evidenced by being highly conserved in diverse three-finger peptides such as LY-6 (Q14210, P35459, Q99JA5, Q9CXN2, Q64253, Q16553, Q90986, O94772, Q9WUC3, P05533, P35460, P35461, P09568, Q9WU67, Q63317 and Q63318), CD59 (O55186, Q920G6, P58019, Q920G7, P27274, P51447, P46657, Q00996, P47777, Q28216, Q28785, P13987, O62680, and O77541), Lynx-1 (Q9WVC2), and the xenoxins (P38951, P38952, and Q09022). Consequently, starting from the N terminus of the sequences, we designate these ancestral cysteines C^1–C^{10}.

In the three-finger toxins, only eight of these basal cysteines (C^1, C^4–C^{10}) are highly conserved, with the majority of toxins having the cysteine pattern –C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}– and the spacing of these eight ancestral cysteines being highly conserved. C^2 and C^3 are found only in orphan groups II, IV, V, and XIX. Orphan group XVII contains C^2 but lacks C^3. The spacing between the eight ancestral cysteines is also highly conserved (Table 5). It is worth noting that the three-finger peptide found on the skin of the hagfish (Q9UAD1) has only eight cysteines, with the ancestral C^2 and C^3 cysteines missing just as in many snake venom three-finger toxins.

The more recently evolved cysteines in the snake venom three-finger toxins are divergent in location (Fig. 4) and thus evolved independently. In the synergistic group, for example, all have a new cysteine located between C^7 and C^8 but in only one toxin (P17696) does the ancestral C^{10} remain. Of the Type I

Table 3. Genera from which each group has been isolated

Acetylcholinesterase inhibiting toxins	*Dendroaspis*
Antiplatelet toxins	*Dendroaspis*
L-Type calcium channel blocking toxins	*Dendroaspis*
Type IA cytotoxins	*Naja*
Type IB cytotoxins	*Naja*
κ-neurotoxins	*Bungarus*
Type A muscarinic toxins	*Dendroaspis*
Type B muscarinic toxins	*Dendroaspis*
Type C muscarinic toxins	*Naja*
Type I α-neurotoxins	Ubiquitous
Type II α-neurotoxins	Ubiquitous
Type III α-neurotoxins	*Pseudonaja*
Synergistic toxins	*Dendroaspis*
Orphan group I	*Naja*
Orphan group II	*Naja*
Orphan group III	*Bungarus*
Orphan group IV	*Bungarus*
Orphan group V	*Bungarus*
Orphan group VI	*Hemachatus*, *Naja*
Orphan group VII	*Bungarus*
Orphan group VIII	*Naja* and *Bungarus* (although the identity of the entries is suspect)
Orphan group IX	*Bungarus*
Orphan group X	*Dendroaspis*
Orphan group XI	*Dendroaspis*
Orphan group XII	*Micrurus*
Orphan group XIII	*Aspidelaps*
Orphan group XIV	*Hemachatus*
Orphan group XV	*Naja*
Orphan group XVI	*Aspidelaps*
Orphan group XVII	*Bungarus*
Orphan group XVIII	*Bungarus*
Orphan group XIX	*Bungarus* and *Dendroaspis*
Orphan group XX	*Hemachatus* and African *Naja*

α-neurotoxins (previously known as short-chain α-neurotoxins), only the Australo-Papuan/marine elapids species have the recently evolved cysteine located adjacent to C^1 (i.e., $–C^1C^x–$). However, this characteristic motif is lacking in the sea kraits (*Laticauda*). This would support the monophyly of Australo-Papuan elapids and sea snakes to the exclusion of *Laticauda*, a notion consistent with some (Keogh, 1998) but not other (Keogh et al. 1998; Slowinski and Keogh, 2000) reconstructions of elapid phylogeny.

The system of nomenclature devised by us for recently evolved cysteines is based upon which of the two basal cysteines they fall between (i.e., $C^{5/6}$ reflects the cysteine located between the basal C^5 and C^6), and if multiple new cysteines are present within two basal cysteines, then starting from the cysteine closest to the N terminus, they are designated –A, –B, etc. (i.e., $C^{5/6A}$, $C^{5/6B}$, etc.) (Table 6). If basal cysteines are lost, then the nomenclature should be based upon the remaining basal cysteines. Thus the new cysteine in the Australo-Papuan/marine elapids is designated as $C^{1/4\text{-A}}$ while that in orphan group XVII is $C^{2/4\text{-A}}$. Orphan group XVII also has a second recently evolved cysteine, which is designated $C^{9/10\text{-A}}$. Care must be taken in interpreting between groups that have evolved cysteines within the same two basal cysteines. For example, while the $C^{5/6\text{-A}}$ and $C^{5/6\text{-B}}$ cysteines in the Type II α-neurotoxins and the κ-neurotoxins are homologous, this may not necessarily be the case with new groups that will no doubt be discovered as more sequences become available. Careful examination of residues flanking the cysteines will be invaluable in aiding the determination of the relative relationships.

As shown by the well-studied classes, the level of toxicity as well as the specific activity of the different groups is quite variable. Swiss-Prot entries range from extremely potent such as the α-neurotoxins (intravenous LD_{50} of 0.07–0.2 mg/kg) to virtually nontoxic such as orphan group VII (intravenous LD_{50} of 250 mg/kg). There are functional differences in variability between modes of testing (intramuscular, intravenous, intraperitoneal, and subcutaneous). The α-neurotoxins have essentially the same LD_{50} whatever the mode tested, whereas orphan group XI is 75 times more toxic when injected intraperitoneally compared to subcutaneously. Even within a group, profound differences can occur. In the acetylcholinesterase inhibiting toxins, P01403 has a LD_{50} of >20 mg/kg by intravenous injection, while P25681 has a LD_{50} of 2.1 mg/kg by intravenous injection (Viljoen and Botes 1973; Joubert and Taljaard 1978).

However, lethality is a poor indicator of bioactivity. A molecule can be potently active without being strongly toxic. Thus, venom components that are weakly toxic may be potently bioactive in a manner not yet assayed for. Thus orphan group II contains within it toxins previously referred to as "weak neurotoxins" but we have dropped all reference to this name for the reasons outlined above. Nevertheless, it is intriguing to note that toxin groups with the ancestral cysteines C^2 and C^3 still present (orphan groups II, IV, V, and XIX) are the least toxic.

Other than scattered lethality testing, nothing is known about the activities of the majority of the orphan groups. Orphan group II has some low-level α-neurotoxic activity evident, some low-level cytotoxic activity is evident in orphan group XV, and orphan group XVIII has been reported to be weak and reversible neuromuscular nicotinic acetylcholine receptor antagonists. However, as the bioactivities of these three orphan groups are far from resolved, placement into defined functional groups at this time would be premature.

Groups with the potential to contain toxins with divergent activities are the Type I α-neurotoxins, Type II α-neurotoxins, orphan group II, Type A muscarinic toxins, and the cytotoxins.

Table 4. Average physical characteristics of each group

Group	AA[a]	MW	Cys	p*I*
Acetylcholinesterase inhibiting toxins	61	6803	8	8.23
Antiplatelet toxins	60	6815	8	7.45
L-Type Ca^{2+} blockers	60	6837	8	8.55
Type IA cytotoxins	60	6780	8	9.50
Type IB cytotoxins	61	6817	8	9.39
κ-neurotoxins	66	7290	10	7.60
Type A muscarinic toxins	65	7372	8	7.47
Type B muscarinic toxins	64	7215	8	8.87
Type C muscarinic toxins	65	7324	8	7.91
Type I α-neurotoxins	60	6794	8–9	8.56
Type II α-neurotoxins	70	7800	10	7.5
Type III α-neurotoxins	58	6281	8	7.78
Synergistic toxins	63	6745	8–9	8.30
Orphan group I	65	7418	8	8.23
Orphan group II	65	7501	10	8.54
Orphan group III	65	7207	8	9.81
Orphan group IV	66	7334	10	7.82
Orphan group V	68	7518	10	8.02
Orphan group VI	61	6967	8	6.10
Orphan group VII	62	7151	8	8.62
Orphan group VIII	65	7620	8	7.91
Orphan group IX	65	7478	8	8.01
Orphan group X	59	6653	8	9.15
Orphan group XI	60	6782	8	8.53
Orphan group XII	57	6212	8	8.62
Orphan group XIII	64	7308	9	8.69
Orphan group XIV	63	7504	8	8.85
Orphan group XV	62	6944	8	9.41
Orphan group XVI	63	7164	8	9.47
Orphan group XVII	82	9078	10	7.53
Orphan group XVIII	63	6834	8	8.23
Orphan group XIX	63	7333	10	7.65
Orphan group XX	59	6715	8	7.07

[a]Amino acids.

Phylogenetically the Type I α-neurotoxins contain further divisions that may be reflective of taxonomical rather than functional divisions (Figs. 4 and 5) and they have been used for phylogenetic studies in the past (Slowinski et al. 1997). While residues identified as being essential for postsynaptic activity are conserved (Antil et al. 1999; Antil-Delbeke et al. 2000), many other residues are not. The high degree of variability in other residues may be indicative of variable activities upon peripheral nerve transmission in different animals. The overall homology between the toxins is only 30–40% and differences also exist in the number and location of prolines. Particularly notable is the presence of $C^{1/4A}$ adjacent to C^1 in the toxins isolated from the poorly studied Australo-Papuan/sea snake species. This may allow for the dimerization of the toxins or alternative cysteine connectivity. It is unclear at this time whether this difference is significant enough for these toxins to become a subgroup of the Type I α-neurotoxins. However, it is notable that despite containing the invariant functional residues characteristic of the Type I α-neurotoxins, the toxins from *Dendroaspis* (P01416, P01417, P01418, and P01419) do not resolve well within this clade. Indeed, by MP analysis they actually fall outside of the group. Further functional testing may reveal differences significant enough to justify a separate subgrouping or even full grouping. However, at this time the Type I α-neurotoxins remain a single clade but one urgently in need of in-depth functional analysis.

In contrast to the taxonomically well-ordered Type I α-neurotoxins (which also have a high level of conservation of residues identified as being essential for activity), the Type II α-neurotoxins are particularly heterogeneous (Fig. 7). This group contains a number of poorly defined phylogenetic divisions and a high level of sequence diversity. Indeed, many toxins lack residues that have been previously shown to play important roles in the recognition of the acetylcholine receptor (Antil-Delbeke et al. 2000). As with the Type I α-neurotoxins, the vast majority of these components have not been functionally tested (even for LD_{50}). Therefore, at this time for both groups of α-neurotoxins the relationship between phylogeny and functionality is unclear.

Table 5. Cysteine spacing of the conserved eight ancestral cysteines

Group	C^1–C^4	C^4–C^5	C^5–C^6	C^6–C^7	C^7–C^8	C^8–C^9	C^9–C^{10}
Acn-esterase inhibitors	13	4	16	1	11	0	5
Antiplatelet toxins	13	4	16	1	11	0	4
L-Type Ca^{2+} blockers	13	4	16	1	10	0	4
Type IA cytotoxins	10	6	16	3	10	0	4
Type IB cytotoxins	11	6	14	3	10	0	4
κ-neurotoxins	10	6	20	3	11	0	4
Type A muscarinic	13	6	17	3	11	0	4
Type B muscarinic	13	4	20	1	10	0	4
Type C muscarinic	13	6	17	3	10	0	4
Type I α-ntx	13	6	16	1	10	0	4
Type II α-ntx	12	6	20	3	11	0	4
Type III α-ntx	9	6	17	3	7	0	4
Synergistic toxins	13	6	17	3	10	0	4
Orphan group I	13	6	17	3	10	0	4
Orphan group II	13	6	17	3	10	0	4
Orphan group III	13	6	17	3	10	0	4
Orphan group IV	15	6	16	3	11	0	4
Orphan group V	13	6	20	3	10	0	4
Orphan group VI	10	6	15	3	11	0	4
Orphan group VII	14	4	17	4	7	0	4
Orphan group VIII	13	6	16	3	11	0	4
Orphan group IX	13	6	16	3	11	0	4
Orphan group X	13	5	16	1	11	0	4
Orphan group XI	13	4	15	1	11	0	4
Orphan group XII	12	4	14	1	10	0	4
Orphan group XIII	11	6	17	3	11	0	4
Orphan group XIV	11	6	17	3	10	0	4
Orphan group XV	11	6	17	3	10	0	4
Orphan group XVI	11	6	18	3	10	0	4
Orphan group XVII	13	6	26	3	13	0	4
Orphan group XVIII	11	5	17	3	5	0	4
Orphan group XIX	13	6	14	3	11	0	4
Orphan group XX	8	4	17	2	8	0	4

While a high level of overall sequence similarity is evident in orphan group II, there exists a significant division within the group (Fig. 8). This division is supported by high bootstrap values and may be indicative of functional differences. However, the low distance levels between the clades makes formalizing the subgroups premature at this time without functional data to support the divisions.

Muscarinic toxins have previously been been divided into Type A and Type B, with all toxins being from *Dendroaspis* (mamba) venom. (Karlsson et al. 2000). The relationship of the cobra muscarinic toxins has not been examined previously. Phylogenetically, the Type A muscarinic toxins are distinct from the *Naja* (cobra) muscarinic toxins (Fig. 8). Consequently the cobra toxins represent a distinct group in their own right and thus are placed into the Type C muscarinic group. The occurrence of muscarinic toxins in *Naja* and *Dendroaspis*, two phylogenetically distant groups, suggests that these toxin types should also occur in other Old World elapid taxa from which they have not yet been characterized. It is likely that activities not only will be diverse in receptor subtype preference but also may prove to be antagonistic as well as agonistic. It is reasonable to hypothesise that the well defined divisions within the Type A muscarinic toxin group may herald significant differences in functional activities. As the majority of these toxins have been assayed simply for binding, the activity (i.e., agonistic vs antagonistic) has not been determined for many.

It is extremely interesting that the Type B muscarinic toxin (Carsi et al. 1999) shares little homology with the Type A muscarinic toxins. Indeed, this toxin has a much stronger affinity for the clade containing the Type I α-neurotoxin and toxins with similar structures. This indicates either great radiation of the muscarinic toxins within the mambas or a case of convergent evolution for the same receptor, i.e., functional homology vs functional homoplasy. This makes the muscarinic toxins an excellent functional evolution case study. From the point of view of searching for novel investigational ligands, the muscarinic toxins are also a most satisfactory group as a whole.

The Type IA cytotoxins are represented by a large group of toxins from mostly Asian species of *Naja* (cobra) (Fig. 9). It has been proposed previously that the cytotoxins are grouped into two types, the P types and the S types, based on the relative presence and

Table 6. Cysteine pattern of each group

Acetylcholinesterase inhibiting toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Antiplatelet toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
L-Type calcium channel blocking toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type IA cytotoxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type IB cytotoxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
κ-neurotoxins	–C^1–C^4–C^5–$C^{5/6A}$–$C^{5/6B}$–C^6–C^7–C^8C^9–C^1
Type A muscarinic toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type B muscarinic toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type C muscarinic toxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type I α-neurotoxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}– or –$C^1C^{1/4A}$–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Type II α-neurotoxins	–C^1–C^4–C^5–$C^{5/6A}$–$C^{5/6B}$–C^6–C^7–C^8C^9–C^{10}–
Type III α-neurotoxins	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Synergistic toxins	–C^1–C^4–C^5–C^6–C^7–$C^{7/8A}$–C^8C^9– or –C^1–C^4–C^5–C^6–C^7–$C^{7/8A}$–C^8C^9–C^{10}–
Orphan group I	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group II	–C^1–C^2–C^3–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group III	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group IV	–C^1–C^2–C^3–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group V	–C^1–C^2–C^3–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group VI	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group VII	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group VIII	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group IX	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group X	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XI	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XII	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XIII	–C^1–C^4–C^5–C^6–C^7–$C^8C^9C^{9/10A}$—C^{10}–
Orphan group XIV	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XV	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XVI	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XVII	–C^1–C^2–$C^{2/4A}C^4$–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XVIII	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XIX	–C^1–C^2–C^3–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–
Orphan group XX	–C^1–C^4–C^5–C^6–C^7–C^8C^9–C^{10}–

location of prolines and serines (Dufton and Hiden 1991). However, these divisions were not supported by the results of our phylogenetic analysis. While conserved divisions exist within the Type IA cytotoxins, these toxins all contain the characteristic cytotoxins functional motifs MxM and IDV (Stevens-Truss and Hinman 1996; Kumar et al. 1999) and therefore the phylogenetic subdivisions may not, in this case, be reflective of functional divisions. However, it is worth noting the level of divergence of the Type IA cytotoxins from the African spitting cobras *Naja mossambica* and *N. pallida* (P01467, P01468, P01469, P01470, P25517). These toxins form a clade separate from the main cytotoxin group (including being separated from other African cobra species). Comparative assaying would be required to determine if there are differences in potency or specificity between these toxins and the main group of Type IA cytotoxins and therefore whether these toxins from the African spitting cobra toxins indeed represent a functional subgroup. In contrast, the toxins P01471, P24776, and P24777 from the rinkhals (*Hemachatus haemachatus*) form a well-separated clade. These toxins not only are clearly phylogenetically distinct but also have changes in the MxM and IDV functional motifs. The first methionine is present in all the toxins but the second is lacking in P24776. The IDV cytotoxic motif has been replaced with TDA (P01471 and P24776) or TDT (P24777). As such, both phylogeny and changes in functional motifs justify these toxins being placed in the Type IB cytotoxin subgroup. The pharmacology of the entire cytotoxin clade is poorly resolved and much work remains to be done to elucidate the activities. In particular, the ability to interfere specifically with cell adhesion processes remains to be determined for these toxins but we consider it likely that this will ultimately be shown and that specific receptors may be targeted.

Ramifications of Inferred Structure–Function Relationships

By combining a molecular evolutionary approach with information on biochemical properties, we were able to make inferences on three-finger toxin structured–function relationships. An example of this is provided by the acetylcholinesterase inhibiting toxins. The two mamba toxins P01403 (*Dendroaspis angusticeps*) and P25681 (*Dendroaspis polylepis*) differ only by M/I and TN/KD, respectively. The first two substitutions (M/I and T/K) are hydrophobic for hydrophobic and polar for polar so could be considered

conserved substitutions. The third substitution (N/D) is polar for polar but is uncharged being changed to charged. As discussed earlier, the murine LD_{50} values differ by over 10-fold despite the high degree of homology between the toxins. A comparison of lethality against different likely mamba prey species may be revealing in this context, especially in view of documented differences between the diet of *Dendroaspis polylepis* (mostly mammals) and that of *D. angusticeps* (mostly birds) (Branch et al. 1995).

Another example of the usefulness of phylogenetic analysis in identifying groups of interest for structure–function analysis is the *Dendroaspis* toxin P01419. While this toxin is strongly aligned with but slightly distinct from the α-neurotoxins from *Dendroaspis* (P01416, P01417, and P01418), the functionally important residue in reactive loop II (K) (Fillet et al. 1993) has been changed to S. This toxin also lacks the invariant E in the later part of the molecule, having K in its place. It is possible that two residue changes are responsible for this toxin being almost 100-fold less toxic than comparable Type I α-neurotoxins (Joubert and Taljaard 1979). It remains to be determined whether these residue changes may have affected the three-dimensional structure of the toxin in such a way that its binding to the nicotinic acetylcholine receptor may be dramatically lessened. Further pharmacological testing may answer the question whether this toxin is essentially devoid of affinity for the nAChR yet may bind elsewhere and, as such, may prove to be another useful investigational ligand.

Previously, we stated that a molecular evolutionary analysis of toxin multigene families might have important ramifications for other fields of toxin research, including biomedicine and the search for useful investigational ligands. We found that three-finger toxins were identified as part of discrete groups even though they had been entered in the databases as members of other groups. For example, P25518 from *Dendroaspis polylepis* was entered into Swiss-Prot as a synergistic-type toxin. However, it contains all of the invariant residues of the muscarinic toxins and aligns deeply within this clade. Indeed as it differs by only two residues from P80495 (E/D and N/E) and three residues from MT-5 (I/K, E/D, and N/E), it is thus logical to conclude that this is a muscarinic toxin. Intriguingly, despite differing by only one residue, P80495 and MT-5 have considerable differences in receptor subtype binding affinity (Karlsson et al. 2000). With P25518 differing from MT-5 in a slightly different manner than P80495 does, functional testing may shed more light as to which residues are essential for affinity to the different receptor subtypes. This may continue the history displayed by this toxin group of being extremely useful investigational ligands.

A similar situation occurred within the L-type calcium channel blocking toxins. Toxin P25683 from *Dendroaspis jamesoni* (Jameson's mamba) was entered into Swiss-Prot as a short neurotoxin homologue. However, it showed a high degree of homology to the proven L-type channel blocking toxins from other mamba species and aligned strongly within this group. Intriguingly this toxin differs appreciably in the region of C^1–C^4 from other members of the group. In place of the conserved CYO HKASLPRATKTC this toxin has CYTHKSLQAKTTKSC. As the structure–function relationships of these toxins have not been fully elucidated, it is unclear at this time if these differences have any effect upon potency or specificity. This is certainly a toxin worthy of in-depth study. Another notable result of considerable research interest is that orphan groups X and XI consistently cluster with the L-type calcium channel inhibiting toxins and, as such, may represent a larger clade of ion channel toxins.

An example of problematic naming was the entire group that was renamed in this study as orphan group XV. This group is made up of venom components from *Naja* species (Cobras). Despite lacking the invariant residues of cytotoxins and not having any demonstrated cytotoxic activity, some of these toxins had previously been referred to as CLBPs (cytotoxin-like basic peptides) (Inoue et al. 1987) as well as less-cytotoxic basic polypeptides (LCBP) (Takechi et al. 1985). As these toxins are functionally and phylogenetically distinct from the cytotoxins, the names only serve to promote confusion. This entire group is thus moved into orphan group XV until such time as the activity can be elucidated and a proper system of nomenclature devised. An example of premature naming was Q9YGJ0 from orphan group V. This and two other toxins also from *Bungarus multicinctus* (O12963 and Q9YGH9) form a phylogenetically distinct group of unknown activity. Only one of the toxins, Q9YGJ0, has been tested and pharmacological studies were limited to intravenous LD_{50} testing (0.15 mg/kg) and observations of intracerebroventricular injections with "laboured breathing" the only result reported (Aird et al. 1999). Nothing was determined about the mechanism of action. However, authors concluded that the toxin acted antagonistically upon the nicotinic acetylcholine receptor and placed this toxin into a new group, the "γ-neurotoxin class." We consider this designation as premature since neurotoxicity has not been confirmed, let alone determination of postsynaptic neurotoxicity at nicotinic acetylcholine receptor subtypes or binding sites distinct from those targeted by α- or κ-neurotoxins. Further to this, the presence of the RGD motif in Q9YGJ0 means that antiplatelet activity additionally cannot be ruled out. Consequently, we place this toxin, and the closely related toxins O12963 and Q9YGH9, into orphan group V until such time as the mechanism of action is eluci-

dated and the toxins named accordingly; e.g., Type (n) α-neurotoxins if α-neurotoxicity is determined, moved back into γ-neurotoxins if evidence for a distinct mode of neurotoxic action is produced or named accordingly if a novel mode of action is revealed.

Caveats and Conclusions

Several caveats need to be considered in an investigation such as this. First, any analysis of this kind can only be as accurate as the identification of the toxins involved. The field of toxinology has had a notoriously disastrous track record of taxonomic confusion and inaccuracy, with the result that the identification of a substantial proportion of venoms and venom components is likely to be questionable or erroneous (Wüster and McCarthy 1996). These errors are likely to be confined primarily to the lowest taxonomic levels (i.e., among closely related and frequently confused species, such as the Asiatic *Naja*), but nonetheless, other errors are possible. An example of this can be found in orphan group VIII, made up of one *Bungarus* (Krait) and two *Naja* toxins (Cobra). The *Bungarus* toxin (Q9W727) is identical to the reported sequence of one of the *Naja* toxins (Q9DEQ3). This is extraordinary considering the taxonomically extremely divergent snakes from which they were isolated. However, in light of the fact that the same laboratory reported all three sequences, the possibility of laboratory error or an error in database input cannot be ruled out.

Another important caveat, with potential impact particularly on the interpretation of evolutionary patterns, is that the toxin sequence data used here have not been compiled in a systematic manner, with a strong biological bias in the sequencing of toxins to date. The venoms of some species have been studied intensively, and a large number of different toxins sequenced. On the other hand, the venoms of other species have remained largely unstudied, at least as far as toxin sequences are concerned. In the present study, no fewer than 131 of 263 sequenced toxins (49.8%) come from the single genus *Naja*, and 222 of 263 toxins (84.4%) come from the four genera *Naja*, *Bungarus*, *Dendroaspis*, and *Laticauda*. On the other hand, Australian terrestrial elapids are grossly underrepresented, with only 19 toxin sequences (7.2%). The confirmed presence of three-finger toxins in diverse genera such as *Acanthophis*, *Aipysurus*, *Laticauda*, *Oxyuranus*, *Notechis*, *Pseudechis*, and *Pseudonaja* suggests that they should be widespread in this clade, and the lack of sequences is more likely to be due to a lack of research than a lack of the toxins. Rigorous and systematic LC/MS analysis of Australian elapid venoms shows that components with molecular weights of 6–8 kDa are ubiquitous (B.G. Fry, unpublished results). In addition, fragments of toxins were not included in the analysis. This necessitated the deletion of toxins from divergent groups such as *Maticora* and *Micropechis*. In any case, studies relying on the interpretation of gene trees must remain subject to the fundamental logical consideration that the absence of evidence of certain toxins in certain groups cannot necessarily be taken as evidence of their absence: The toxins concerned may well be present, but not yet fully sequenced, and thus missing from the tree.

In summary, this study has provided evidence for the birth-and-death model of sequence evolution in the three-finger toxins, as well as providing a phylogenetic framework for future work on this important family of snake venom toxins. It is anticipated that this "three-finger toxin toolkit" will prove to be useful in providing a clearer picture of the diversity of investigational ligands available within this important class of toxin.

Acknowledgments. We dedicate this paper to the memory of our friend Dr. Joseph B. Slowinski, who died from snakebite in September of 2001 in Myanmar in the pursuit of new species of cobra. We are grateful for the financial assistance of the Australia and Pacific Science Foundation and the Melbourne Aquarium.

References

Afifiyan F, Armugam A, Tan CH, Gopalakrishnakone P, Jeyaseelan K (1999) Postsynaptic alpha-neurotoxin gene of the spitting cobra. *Naja nafa sputatrix*: Structure, organization, and phylogenetic analysis. Genome Res 9(3):259–266

Aird SD, Womble GC, Yates JR, 3rd Griffin PR (1999) Primary structure of γ-bungarotoxin, a new postsynaptic neurotoxin from venom of *Bungarus multicinctus*. Toxicon 37(4):609–625

Antil S, Servent D, Ménez A (1999) Variability among the sites by which curaremimetic toxins bind to *Torpedo* acetylcholine receptor, as revealed by identification of the funcation residues of α-cobratoxin. J. Biol Chem 274(49):34851–34858

Antil-Delbeke S, Gaillard C, Tamiya T, Corringer P-J, Changeux J-P, Servent D, Ménez A (2000) Molecular determinants by which a long chain toxin from snake venom interacts with the neuronal a7-nicotinic acetylcholine receptor. J Biol Chem 275(38):29594–29601

Assakura MT, Furtado MFD, Mandelbaum FR (1992) Biochemical and biological differentiation of the venoms of the Lancehead vipers (*Bothrops atrox*, *Bothrops asper*, *Bothrops marajoensis* and *Bothrops moojeni*). Comp Biochem Physiol 102B:727–732

Branch WR, Haagner GV, Shine R (1995) Is there an ontogenetic shift in mamba diet? Taxonomic confusion and dietary records for black and green mambas (*Dendroaspis*: *Elapidae*). Herpetol Nat Hist 3:171–178

Carsi JM, Potter LT (2000) m1-toxin isoforms from the green mamba (*Dendroaspis angusticeps*) that selectively block m1 muscarinic receptors. Toxicon 38:187–l98

Carsi JM, Valentine HH, Potter LT (1999) m2-toxin: A selective ligand for m2 muscarinic receptors. Mol Pharmacol 56:933–937

Chang L, Lin S, Huang H, Hsiao M (1999) Genetic organization of alpha-bungarotoxins from *Bungarus multicinctus* (Taiwan banded krait): Evidence showing that the production of

α-bungarotoxin isotoxins is not derived from edited mRNAs. Nucleic Acids Res 27(20):3970–3975

Daltry JC, Wüster W, Thorpe RS (1996) Diet and snake venom evolution. Nature 379:537–540

Dufton MJ, Hider RC (1991) The structure and pharmacology of elapid cytotoxins. In: Harvey AL (ed) Snake toxins. Pergamon Press, New York

Endo T, Tamiya N (1987) Current view on the structure-function relationship of postsynaptic neurotoxins from snake venoms. Pharmacol Ther 34(3):403–451

Felsenstein J (2001) PHYLIP (phylogeny inference package) version 36. Department of Genetics, University of Washington, Seattle (http://evolutiongeneticswashingtonedu/phyliphtml)

Fleming, TJ, Ohigin C, Malek TR (1993). Characterization of two novel Ly-6 genes. Protein sequence and potential structural similarity to a-bungarotoxin and other neurotoxins. J Immunol 150:5379–5390

Fry BG, Wickramaratna JC, Hodgson WC, Alewood PF, Kini RM, Ho H, Wüster W (2002) Electrospray liquid chromatography/mass spectrometry fingerprinting of *Acanthophis* (death adder) venoms: Taxonomic and toxinological implications. Rapid Commun Mass Spectrom 16(6):600–680

Glenn JL, Straight RC, Wolfe MC, Hardy DL (1983) Geographical variation in *Crotalus scutulatus scutulatus* (Mojave rattlesnake) venom properties. Toxicon 21:119–130

Gong N, Armugam A, Jeyaseelan K. (1999) Postsynaptic short-chain neurotoxins from *Pseudonaja textilis*. cDNA cloning, expression and protein characterization. Eur J Biochem 265(3):982–989

Gumley TP, McKenzie IFC, Sandrin MS (1995). Tissue expression, structure and function of the murine Ly-6 family of mole-cules. Immunol Cell Biol 73:277–296

Heatwole H, Poran NS (1995) Resistances of sympatric and allopatric eels to sea snake venoms. Copeia 1:136–147

Jiménez-Porras JM (1964) Intraspecific variations in composition of venom of the jumping viper, *Bothrops nummifera*. Toxicon 2:187–195

Jolkkonen M (1996) Muscarinic toxins from *Dendroaspis* (mamba) venoms. Peptides selective for subtypes of muscarinic acetylcholine receptors. Acta Universitatis Upsaliensis. Comprehensive summaries of Uppsala dissertations from the Faculty of Science and Techonology 183

Joubert FJ, Taljaard N (1978) The complete primary structure of toxin C from *Dendroaspis polylepis polylepis* (black mamba) venom. S Afr J Chem 31:107–110

Joubert FJ, Taljaard N (1979) Complete primary structure of toxin CM-1C from *Hemachatus haemachatus* (Ringhals) venom. S Afr J Chem 32:73–77

Karlsson E, Jokkonen M, Mulugeta E, Onali P, Adem A (2000) Snake toxins with high selectivity for subtypes of muscarinic acetylcholine receptors. Biochimie 82:793–806

Keogh JS (1998) Molecular phylogeny of elapid snakes and a consideration of their biogeographic history Biol J Linn Soc 63:177–203

Keogh JS, Shine R, Donnellan S (1998) Phylogenetic relationships of terrestrial Australo-Papuan elapid snakes (Subfamily Hydrophiinae) based on cytochrome b and 16SrRNA sequences. Mol Phylogenet Evol 10:67–81

Kini RM, Chan YM (1999) Accelerated evolution and molecular surface of venom phospholipase A(2) enzymes. J Mol Evol 48:125–132

Kordis D, Gubensek F (2000) Adaptive evolution of animal toxin multigene families. Gene 261:43–52

Kumar TKS, Pandian STK, Jayaraman G, Peng HJ, Yu C (1999) Understanding the structure, function and folding of cobra toxins. Proc Natl Sci Counc ROC(A) 23(1):1–19

Ménez A (1998) Functional architectures of animal toxins: A clue to drug design? Toxicon 36(11):1557–1572

Moura-da-Silva AM, Paine MJ, Diniz MR, Theakston RD, Crampton JM (1995) The molecular cloning of a phospholipase A_2 from *Bothrops jararacussu* snake venom: Evolution of venom group II phospholipase A2's may imply gene duplications. J Mol Evol 41(2):174–179

Nakashima K, Nobuhisa I, Deshimaru M, et al. (1995) Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A_2 isozyme genes. Proc Natl Acad Sci USA 92(12):5605–5609

Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA 94(15):7799–7806

Nirthanan S, Charpantier E, Gopalakrishnakone P, Gwee MCE, Khoo H-E, Cheah L-S, Bertrand D, Kini RM (2002) Candoxin, a novel toxin from *Bungarus candidus*, is a reversible antagonist of muscle (alphabetagammadelta) but a poorly reversible antagonist of neuronal alpha 7 nicotinic acetylcholine receptors. J Biol Chem 277:17811–17820

Page RD (2001) Gene Tree, version 130. University of Glasgow, Glasgow (available at http://taxonomyzoologyglaacuk/rod/rodhtml)

Pillet L, Tremeau O, Ducancel F, Drevet P, Zinn-Justin S, Pinkasfeld S, Boulain JC, Ménez A (1993) Genetic engineering of snake toxins. Role of invariant residues in the structural and functional properties of a curaremimetic toxin, as probed by site-directed mutagenesis. J Biol Chem 268(2):909–916

Poran NS, Coss RG, Benjamini E (1987) Resistance of California ground squirrels (*Spermophilus beecheyi*) to the venom of the northern Pacific rattlesnake (*Crotalus viridis oreganus*): A study of adaptive variation. Toxicon 25(7):767–777

Rooney AP, Piontkivska H, Nei M (2002) Molecular evolution of the nontandemly repeated genes of the histone 3 multigene family. Mol Biol Evol 19(1):68–75

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

Slowinski JB (1994) A phylogenetic analysis of *Bungarus* (Elapidae) based on morphological characters. J Herpetol 28:440–446

Slowinski JB (1995) A phylogenetic analysis of the New World coral snakes (*Elapidae: Leptomicrurus*, *Micruroides*, and *Micrurus*) based on allozymic and morphological characters. J Herpetol 29:325–338

Slowinski JB, Keogh JS (2000) Phylogenetic relationships of elapid snakes based on cytochrome b mtDNA sequences. Mol Phylogenet Evol 51:157–164

Slowinski JB, Knight A, Rooney AR (1997) Inferring species trees from gene trees: A phylogenetic analysis of the Elapidae (Serpentes) based on the amino acid sequences of venom proteins. Mol Phylogenet Evol 8:349–362

Slowinski JB, Boundy J, Lawson R (2001) The phylogenetic relationships of Asian coral snakes (Elapidae: *Calliophis* and *Maticora*) based on morphological and molecular characters. Herpetologica 57:233–245

Stevens-Truss R, Hinman CL (1996) Chemical modification of methionines in a cobra venom cytotoxin differentiates between lytic and binding domains. Toxicol App Pharmacol 139:234–242

Swofford DS (2002) PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10 for Macintosh. Sinauer Associates, Sunderland, MA

Takahashi K, Nei M (2000) Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. Mol Biol Evol 17:1251–1258

Takechi M, Tanaya Y, Ahayashi K (1985) Amino acid sequence of a cardiotoxin-like basic peptide (CLBP) with low cytotoxicity isolated from the venom of Formsan Cobra (*Naja naja atra*). Biochem Int 11:795–802

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25(24):4876–4882

Tsetlin V (1999) Snake venom alpha-neurotoxins and other 'three-finger' proteins. Eur J Biochem 264(2):281–286

Viljoen CC, Botes DP (1973) Snake venom toxins. The purification and amino acid sequence of toxin F-VII from *Dendroaspis angusticeps* venom. J Biol Chem 248:4915–4919

Wüster W, McCarthy CJ (1996) Venomous snake systematics: Implications for snakebite treatment and toxinology. In: Bon C, Goyffon M (eds) Envenomings and their treatments. Fondation Mérieux, Lyon, pp. 13–23

Yang CC, Chang LS, Wu FS (1991) Venom constituents of *Notechis scutatus scutatus* (Australian tiger snake) from differing geographic regions. Toxicon 29:1337–1344