Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Effects of Genome Structure and Sequence on Variation and Evolution*

Adaptive radiation of venomous marine snail lineages and the accelerated evolution of venom peptide genes

Baldomero M. Olivera,[1] Maren Watkins,[1] Pradip Bandyopadhyay,[1] Julita S. Imperial,[1] Edgar P. Heimer de la Cotera,[2] Manuel B. Aguilar,[2] Estuardo López Vera,[3] Gisela P. Concepcion,[4,5] and Arturo Lluisma[4,5]

[1]Biology Department, University of Utah, Salt Lake City, Utah. [2]Laboratorio de Neurofarmacología Marina, Departamento de Neurobiología Celular y Molecular, Instituto de Neurobiología, Universidad Nacional Autónoma de México, Campus Juriquilla, Mexico. [3]Instituto de Ciencias del Mar y Limnologia, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Coyoacán, Mexico. [4]Marine Science Institute, University of the Philippines, Diliman, Philippines. [5]Philippine Genome Center, University of the Philippines, Diliman, Philippines

Address for correspondence: Baldomero M. Olivera, Biology Department, University of Utah, 257 S. 1400 E., Salt Lake City, UT 84112. olivera@biology.utah.edu

An impressive biodiversity (>10,000 species) of marine snails (suborder Toxoglossa or superfamily Conoidea) have complex venoms, each containing approximately 100 biologically active, disulfide-rich peptides. In the genus *Conus*, the most intensively investigated toxoglossan lineage (∼500 species), a small set of venom gene superfamilies undergo rapid sequence hyperdiversification within their mature toxin regions. Each major lineage of Toxoglossa has its own distinct set of venom gene superfamilies. Two recently identified venom gene superfamilies are expressed in the large Turridae clade, but not in *Conus*. Thus, as major venomous molluscan clades expand, a small set of lineage-specific venom gene superfamilies undergo accelerated evolution. The juxtaposition of extremely conserved signal sequences with hypervariable mature peptide regions is unprecedented and raises the possibility that in these gene superfamilies, the signal sequences are conserved as a result of an essential role they play in enabling rapid sequence evolution of the region of the gene that encodes the active toxin.

Keywords: venom peptides; accelerated evolution; Conidae; Turridae

Introduction

Venom has evolved independently many times in different phylogenetic lineages. The biodiversity of venomous animals that live in diverse habitats today may well exceed 10^5 species. Terrestrial groups such as snakes, spiders, scorpions, bees, and wasps are the most well known;[1,2] however, the marine environment has a comparable biodiversity of venomous animals, with those that frequently injure humans (such as jellyfish) being the most familiar.[3] Venomous marine snails, including the cone snails (genus *Conus*; family Conidae), constitute a significant fraction of the total biodiversity of all venomous animals (>10,000 species).[4] Two other groups of marine snails, the terebrids or auger snails (family Terebridae) and the turrids (family Turridae, s.l.) are venomous, and together with the cone snails, traditionally constitute the suborder Toxoglossa (or alternatively, the superfamily Conoidea).

Marine snail venoms are complex

A noteworthy feature of marine snail venoms is their complexity; it was estimated, based on the initial biochemical purification of cone snail venoms, that each cone snail species has a repertoire of between 50–200 different venom components, mostly small disulfide-rich peptides.[5–8] Only a subset of these are expressed at one time in the venom of a single individual. However, on the basis of recent mass spectrometry analyses, it has been suggested that a much larger number of venom components are present, and that ∼200 is a gross underestimate.[9] What has yet to be definitively established is the number of gene products encoded by distinct genes that are used in envenomation. Biass *et al.*[10] suggested that

doi: 10.1111/j.1749-6632.2012.06603.x

some of the >1,000 components they detected by mass spectrometry from *Conus consors* venom may be processing intermediates, or variants in the extent of posttranslational modification, and propeptides encoded by the same gene, and even degradation fragments that might be artifacts of venom collection and storage. Until the number of different genes encoding distinct peptides has been established, this issue remains unsettled. With the increasing number of transcriptome and genome analyses of cone snail venoms,[8,11–13] a determination of the number of venom peptides encoded by different genes used for envenomation should be forthcoming in the near future.

Cone snail venom peptides are examples of startlingly rapid evolution

A property of cone snail venom peptides that is now well established is their greatly accelerated evolution; essentially, each cone snail species has its own distinctive complement of venom peptides.

In general, only two or three gene superfamilies account for the majority of the peptides found expressed in the venom of each *Conus* species; however, which gene families predominate depends on the phylogenetic clade within the genus *Conus* to which that species belongs (see Fig. 1). In the past decade, the gene superfamilies that encode cone snail venom peptides have become increasingly well defined.[7] Most major gene superfamilies are found across the entire genus *Conus*, although a few seem to have a more restricted distribution.

The distribution of peptides in different gene superfamilies for the six *Conus* species that have yielded at least 30 different peptides from transcriptome analyses is shown in Figure 1A. Three of these are fish-hunting cone snails that belong to different clades: *C. bullatus* (*Textilia* clade), *C. consors* (*Pionoconus* clade), and *C. geographus* (*Gastridium* clade). One snail hunting species, *C. textile* (*Cylinder* clade) is shown, as well as two worm-hunting species, *C. litteratus* (*Strategeconus* clade) and *C. pulicarius* (*Puncticulis* clade).[16] Two species in Figure 1 have the greatest number of different peptides, *C. pulicarius* (82 sequences) and *C. geographus* (62 sequences). These peptides are from 16 different gene superfamilies; while some superfamilies are well represented in both (i.e., the O-superfamily[17,18]), peptides from seven of the superfamilies were found in only one of the two species.

Dramatic juxtapostion of conserved and hypervariable regions in *Conus* venom peptide gene superfamilies

A venom peptide is first translated as a larger precursor,[19,20] which can readily be divided into three regions: the signal sequence or preregion, an intervening or proregion, and, at the C-terminal end, the mature peptide region always present in single copy.[a] Throughout the genus, a superfamily retains the same conserved signal sequence at the N-terminus, juxtaposed against a hypervariable mature peptide (core) region at the C-terminus. With some exceptions, the number and arrangement of Cys residues is a conserved feature of all peptides in the same gene superfamily.[7]

Hypervariability of the biologically active toxin includes synonymous coding positions

In all major cone snail venom peptide gene superfamilies, there is a juxtaposition of an extremely conserved signal sequence region with the hypervariable mature toxin region; the C-terminal core sequences undergo accelerated evolution at an unprecedented rate. This striking feature was noted the very first time genes encoding *Conus* peptides were cloned.[19] It is important to note that the high rate of amino acid change that occurs in the core region is not merely because of extremely strong selection on individual toxin amino acids because the rate of change in synonymous codons is also greatly elevated.[22]

After undergoing posttranslational modification, the mature toxin peptide is released from the precursor by proteolytic cleavage. The posttranslationally modified mature peptide is the biologically active gene product injected into a targeted animal and binds specifically to a molecular target—an ion channel or receptor, and therefore, it would be this segment of the translated precursor, which is subject to intense diversifying selection. This is the

[a]Recently, it has been suggested that, for all posttranslationally modified but ribosomally synthesized peptides, no matter what their origin, the proregion be referred to as the "leader," and that the mature peptide region, which can be posttranslationally modified, as the "core peptide."[21]

A

Framework	Cys pattern	Superfamily	species: C.lit	C.tex	C.bul	C.pul	C.cons	C.geo
I	CC-C-C	A	3		10		14	12
III	CC-C-C-CC	M	8[a]	1	3	4	8	2
V	CC-CC	T	13			5	5	6
VI/VII	C-C-CC-C-C	O	8[b]	24	14	41[d]	15	18
VIII	C-C-C-C-C-C-C-C-C-C	S					3	5
IX	C-C-C-C-C-C	P	3	5		3	1	
XI	C-C-CC-CC-C-C	I				10[e]		1
XIV	C-C-C-C	J, L	2		1	5		5
XV	C-C-CC-C-C-C-C	V	1[c]			1		
XVI	C-C-CC		1[c]					
XIX	C-C-C-CCC-C-C-C-C					1		
Conantokin	Non-disulfide-rich					3	1	2
Contryphan	Non-disulfide-rich		1		1	6		1
Contulakin	Non-disulfide-rich		2			1	4	1
Conkunitzin[f]	(Large)				1	2	7	1
Conophysin[f]	(Large)							1
Con-ikot-ikot[f]	(Large)							7

B

Figure 1. (A) Frequency of venom peptides. Comparison of relative frequency of conotoxin gene superfamilies in various Conus venom duct transcriptomes (Cys frameworks and superfamilies are indicated). (*C. lit*, *C. litteratus*),[14] (*C. tex*, *C. textile*),[15] (*C. bul*, *C. bullatus*),[8] (*C. pul*, *C. pulicarius*),[12] (*C. cons*, *C. consors*),[11] (*C. geo*, *C. geographus*).[13] [a]Includes sequences with three Cys residues; [b]includes sequences with two Cys residues; [c]the preregion (i.e., signal peptide) sequences indicate that this peptide belongs to a different superfamily; [d]includes O1 and some O3 superfamily sequences; [e]includes I1 and I2 sequences; [f]includes large peptides. (B) Shells of the cone snails analyzed above: (left to right) *C. textile*, *C. litteratus*, *C. geographus*, *C. bullatus*, *C. consors*, and *C. pulicarius*.

biological rationale offered for why hypermutation is focally localized to the mature toxin region.

The propeptide region is not hypervariable

The propeptide (leader) region seems to be conserved but undergoes a moderate degree of mutation; these leader sequences are much more conserved than the mature peptide but far less conserved than the adjoining N-terminal signal sequences. The propeptide regions contain binding sites for posttranslational modification enzymes, and thus are believed to facilitate the proper posttranslational processing of the core peptide region. Specifically, it has been demonstrated for one posttranslational modification enzyme, which catalyzes the carboxylation of glutamate to γ-carboxyglutamate, that the recognition sequence within the propeptide region facilitates recruitment of the modification enzyme.[23] Evidence that disulfide isomerases bind to the propeptide has also been presented.[24]

The signal sequence is, unexpectedly, highly conserved

In contrast to every other known gene superfamily of secreted polypeptides, in *Conus* peptide gene superfamilies there is a striking sequence conservation of the signal sequence at the N-terminus of

the precursor. This is unexpected because signal sequences do not require a specific amino acid sequence to function as secretion signals, but rather only need to have a general hydrophobic character. The unprecedented degree of sequence conservation strongly suggests that conopeptide gene superfamily signal sequences play additional important roles.

Specific examples of the juxtaposition of conserved and hypervariable regions are given in Figure 2. (Two groups of *Conus* venom peptides are shown: A-superfamily peptides, which generally have two disulfide linkages—the majority of those characterized are competitive nicotinic acetylcholine receptor antagonists; and S-superfamily peptides, which in contrast, have five disulfide bonds, and the molecular targets of these peptides are mostly unknown.)

The signal sequence and the portion of the mature toxin sequence encompassing the five central cysteine residues are magnified in the figure, and the corresponding nucleic acid sequences are shown in Figure 2B. This provides a particularly striking juxtaposition of absolute conservation in signal sequences, while in the interval of the toxin region shown, not a single amino acid is conserved except for the cysteine residues. Thus, the problem that remains, which has not been incisively addressed, is a mechanism of accelerated evolution resulting in almost absolute sequence conservation at the N-terminus, absent even a silent mutation, but an unprecedented rate of hypermutation at the C-terminal core region.

Rapid diversification of sequence enables rapid access to new prey

The data set on *Conus* venom peptides comprises what may be the most extensive set of rapidly diverging gene products for a single genus of animals. In some cases, thousands of different peptides in the same superfamily, from many different species, have been elucidated. The phylogenetic relationship between various *Conus* species has been determined, so that the pattern of divergence in a particular gene superfamily can be correlated with phylogenetic relationships between species (see, for example, Puillandre *et al.*).[25] For a substantial fraction of the peptides, there has been sufficient physiological characterization to allow the consequences of hypermutation to be assessed not only structurally but also functionally. An example of this type of functional assessment is provided by the five peptides whose precursors are shown in Figure 2, and produced by the *Conus* species shown in Figure 3A.

Two gene superfamilies are represented, the A- and the S-superfamilies. The two S-superfamily peptide sequences shown in Figure 2 are the only examples in this superfamily where molecular targets have been identified: σ-conopeptide GVIIIA,[26] blocks the 5HT3 receptor (a ligand-gated ion channel that has serotonin as the native agonist), while αS-conopeptide RVIIIA[27] is a nicotinic receptor antagonist.

Functionally, the peptides shown fall into two classes: σ-conopeptide GVIIIA is an inhibitor of the serotonin 5HT3 receptor, whereas the other four, αS-conopeptide RVIIIA, α-conotoxins MII,[28] EI,[29] and CI[30] are nicotinic receptor antagonists (despite the fact that the first belongs to a different gene superfamily from the other three). Furthermore, α-conotoxin MII diverges functionally from the three other nicotinic antagonists in that it is targeted to a neuronal nicotinic receptor subtype (the α3β2 nicotinic receptor is the likely physiologically relevant target). The three other peptides all inhibit the muscle nicotinic receptor and paralyze fish.

Rapid convergent evolution across clades

Interestingly, there is an enormous sequence divergence between the latter three peptides although all come from fish-hunting cone snail venoms. There is compelling evidence[31] that the shift from an ancestral worm-hunting species to fish hunting took place independently in the new world to generate the *Chelyconus* clade (which includes *C. ermineus*, the source of α-conotoxin EI), and in the IndoPacific region (to generate the *Pionoconus* clade, which includes *C. catus*, the source of α-conotoxin CI). Although the selective pressure to evolve nicotinic receptor antagonists specifically targeted to the fish neuromuscular junction were presumably similar for the two peptides, because these peptides evolved from quite different worm-hunting ancestors, their evolutionary trajectories were very different, resulting in considerable divergence in their amino acid sequences, but convergence of their structures. Although it inhibits the same molecular target, the third peptide, αS-conotoxin RVIIIA is strikingly different structurally because an entirely different peptide superfamily was recruited for the same physiological purpose.

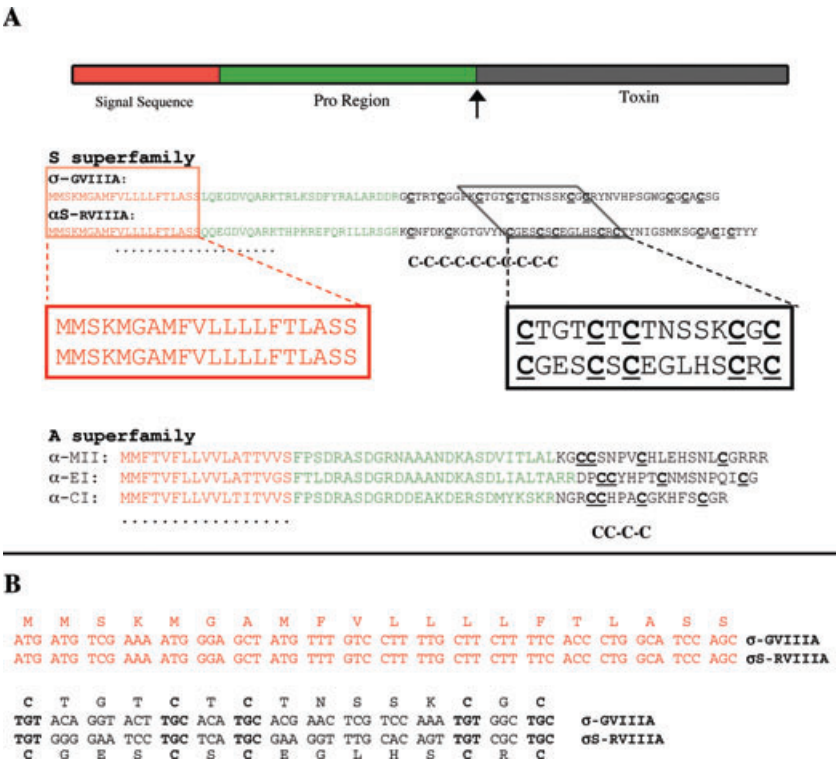

Figure 2. (A) Five precursor sequences from two Conus peptide gene superfamilies. The canonical organization of conopeptide precursors is shown diagrammatically, with the arrow indicating the proteolytic cleavage site that generates the mature toxin ("core") peptide. Two members of the S-superfamily are shown, with mature peptides containing five disulfide linkages (10 Cys residues). Three members of the A-superfamily are illustrated, with mature peptides that have two disulfide crosslinks (four Cys residues). The signal sequence and a segment of the mature toxins of the S-superfamily peptides have been magnified to illustrate the total conservation in the signal sequence region, and the absence of any conservation of non-Cys amino acids in the mature toxin region. (B) The DNA sequences encoding the magnified aa sequences are shown; note the complete conservation of the nucleic acid sequences for the σ-GVIIIA and the αS-RVIII signal sequence regions; note the invariant cysteine codons. The molecular targets of these peptides are discussed in the text. The species from which the venom peptides are derived are illustrated in Figure 3A. Accession numbers: GVIIIA: FJ959110; RVIII: FJ959114; a-MII: P56636; a-EI: P50982; a-CI: FJ868069.1.

Accelerated evolution of venom peptides in other toxoglossan lineages

The vast majority of published studies on toxoglossan venoms focuses on cone snails, despite the fact that these snails are only a minor fraction of the total number of toxoglossan species. The venoms of most large toxoglossan clades have not been investigated at all; the only other group, apart from *Conus*, where a sufficient number of species have been examined so venom peptide evolution can be assessed is the family Turridae. Anatomical data and molecular phylogeny reveal that the classical family Turridae (s.l.) is a polyphyletic taxonomic assemblage, and there have been several proposals to break it up into multiple family groups. Many of the genera previously assigned to Turridae have now been transferred to other families, such as Drillidae and Raphitomidae. The total number of species in the family Turridae (s.s.), as presently restricted to a smaller number of genera by Bouchet *et al.*,[4] are probably comparable to the number of cone snail species. We discuss two superfamilies of turripeptides (as venom peptides from species in this family are designated).[32]

Turridae gene superfamilies are distinct from those expressed in cone snail venom

One issue is whether the major venom gene superfamilies of cone snails are also major components of the venoms of other venomous molluscs. At present, extensive transcriptome data have been published for only one species in the family

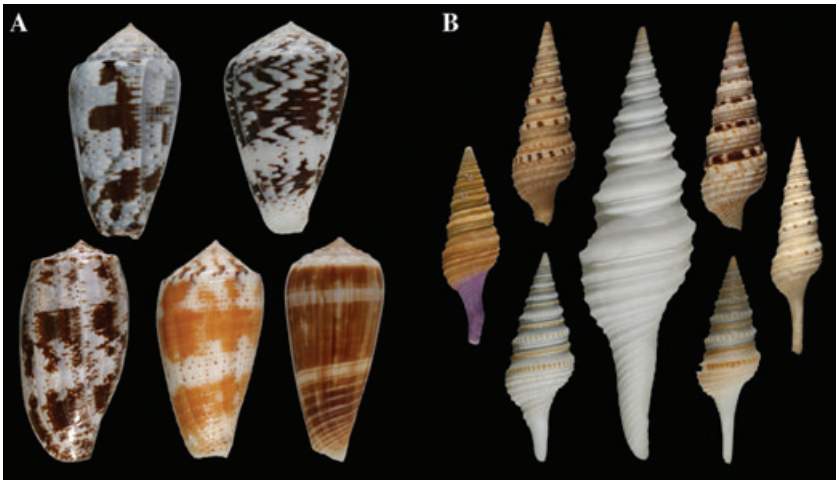

Figure 3. Shells of species analyzed in Figures 2 and 4. Shells are shown of cone snail and turrid species from which gene superfamily sequences discussed in this paper were obtained. The shells are not shown to scale. (A) Conus species from which conopeptide precursor sequences in Figure 2 were derived. Top row: *C. catus* (α-conotoxin CI); *C. ermineus* (α-conotoxin EI). Bottom row: *C. geographus* (σ-conotoxin GVIIIA); *C. magus* (α-conotoxin MII); and *C. radiatus* (αS-conotoxin RVIIIA). All of these species are believed to be fish hunters. *C. catus* and *C. magus* belong to the *Pionoconus* clade; *C. ermineus* to the *Chelyconus* clade, *C. geographus* to the *Gastridium* clade, and *C. radiatus* to the *Phasmoconus* clade. (B) Species in the family Turridae from which the sequences in Figure 4 are derived. Top row: *Xenuroturris cingulifera*; *X. olangoensis*. Large central shell, *Polystira albida*. Middle row: *Gemmula lisajoni*; *Lophiotoma albina*. Bottom row: *G. sogodensis*; *G. speciosa*. All of these species were collected in the Philippines, except for *P. albida*, which was collected from trawling in deep water in the Gulf of Mexico.

Turridae, *Lophiotoma olangoensis*.[33] *L. olangoensis* venom is comparable in complexity to cone snail venoms; some of the turripeptides in the venom are related to each other and belong to the same gene superfamily. However, a major conclusion was that there was little overlap between the gene superfamilies of cone snails, and the putative gene superfamilies that encode peptides expressed in *L. olangoensis* venom, suggesting that the expression of cone snail gene superfamilies does not generally extend to other major lineages of Toxoglossa.

Turripeptide superfamiliy expansion

An overview framework has emerged for how the diverse peptides in the venoms of the ∼500 different *Conus* species have been generated.[6] As new species of cone[6,7] snails evolve, the type of focal hypermutation shown in Figure 2 occurs, primarily in a small subset of the gene superfamilies shown in Figure 1; the consequence is that the complement of peptides found in the venom of each particular *Conus* species diverges significantly from peptides in all other *Conus*. Does this hold true for other toxoglossan lineages as well?

Because the initial analysis of the *L. olangoensis* transcriptome, other species in the family Turridae have been analyzed,[34] so the similarities and differences in peptides from multiple species can be evaluated. Evidence for a gene superfamily subject to accelerated evolution was previously presented for a turripeptide gene superfamily expressed in the genus *Gemmula*.[34] More recent results demonstrate that a family of turripeptides that contain three disulfide bonds is expressed in four different genera of Turridae: the relevant precursor sequences are shown in Figure 4. It is clear that the polypeptides shown in Figure 4A belong to the same gene superfamily, given their highly conserved signal sequences. Although the mature toxin region is hypervariable, the arrangement of the six cysteine residues is conserved.

Gene superfamilies distributed across the entire Turridae undergo accelerated evolution

A preliminary molecular phylogeny of Turridae[35] has revealed that a basal clade in the family is the genus *Polystira*. This is a new world group, with several exceptionally large species found in both the Panamic region and the Caribbean. In the tree

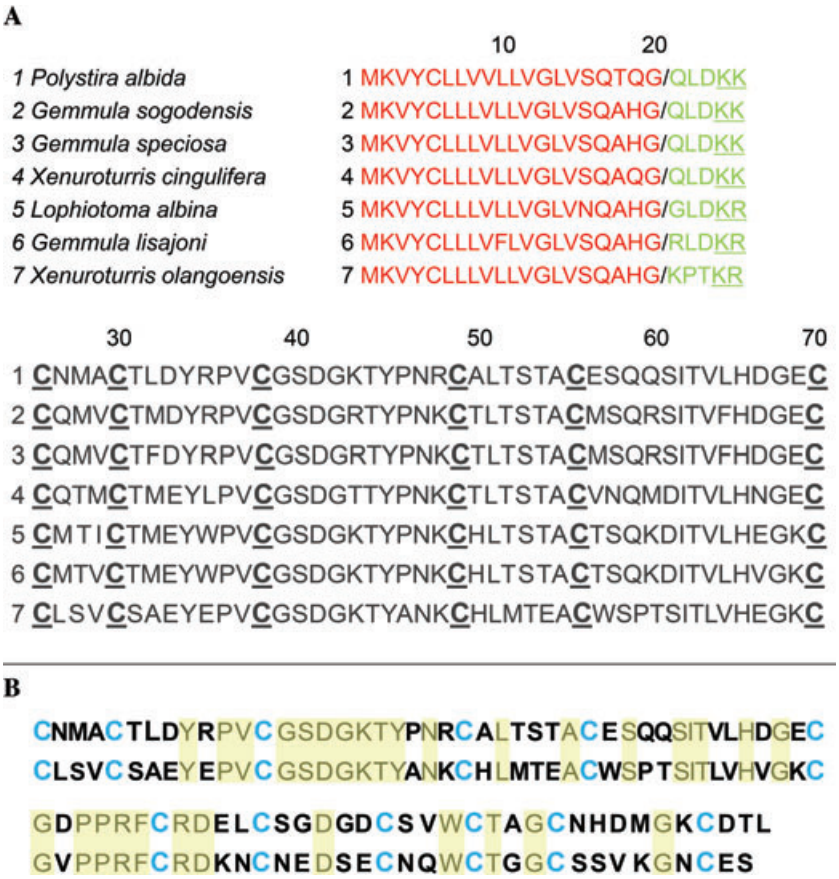

Figure 4. Turripeptide gene superfamilies. (A) Turripeptide sequences from seven species are aligned. All the sequences have a total length of 70 aa, with a conserved signal sequence of 20 aa. Predicted mature peptides have 3 disulfide linkages, in a P-like conopeptide framework (C–C–C–C—C—C). Note that compared to most conopeptide precursors, this turripeptide gene superfamily propeptide region is extremely short (and possibly even absent). (B) Two predicted mature peptide sequences from the superfamily shown in A, compared to two turripeptides from the Pg superfamily previously described by Heralde *et al.*[34] from two *Gemmula* species, *G. sogodensis*, and *G. kieneri*. Note that although both sets of peptides have six Cys residues, there is no alignment between the two groups, and these define two different P-like turripeptide superfamilies. In contrast to the sequences shown in A, the precursors of the Pg superfamily have a well-defined propeptide region of 23 aa.

presented by Heralde *et al.* this clade is the most distant from Indo-Pacific genera such as *Lophiotoma*, *Xenuroturris* and the Indo-Pacific branch of *Gemmula*. Because the genus *Polystira* is phylogenetically distant from the other groups, the presence of the precursor in *Polystira albida* venom shown in Figure 4A is strongly indicative that this gene superfamily is broadly expressed across the Turridae. Thus, as in *Conus*, there seem to be gene superfamilies distributed across the entire family Turridae that undergo accelerated evolution.

In Figure 4B, the predicted mature toxin from two peptides in the Pg turripeptide superfamily[34] are compared to two predicted mature peptides from the turripeptide superfamily in Figure 4A. Although both groups have a P-like conopeptide framework, they share no sequence homology with the P-conopeptide family, nor with each other. These are distinct from any peptides from *Conus* venoms.

Biological perspectives: radiation of biodiverse lineages

Lineage-specific acceleration of gene superfamily evolution

The biological implications of the molecular genetics discussed above are briefly discussed here. The data suggest that diversification of venom peptides

in each major lineage of venomous toxoglossans occurs through the accelerated evolution of a few gene superfamilies that are characteristic of that lineage. There is an unusual conservation of signal sequences in such rapidly evolving gene superfamilies, compared to what is found in other families of secreted polypeptides. In all venom peptide precursors, the conserved signal sequences are strikingly juxtaposed against a rapidly evolving mature toxin region at the C-terminus, with only the cysteine codons conserved. A curious feature of the pattern of hypermutation, first pointed out by Conticello and Fainzilber,[15,36] is that within the hypervariable core region, not only are the cysteine residues conserved, but at each specific position, the codon encoding a Cys residue is invariant, with no mutational cycling between the two alternative Cys codons (see Fig. 2B for a specific example). It was suggested that mutations in the toxin region were introduced through the action of an error-prone DNA polymerase.

High conservation of signal sequence DNA suggests an undiscovered role

The data obtained so far suggest that although there is little overlap in the specific gene superfamilies expressed between the different major toxoglossan lineages, the extraordinary conservation of signal sequences is a general feature of toxoglossan venom peptide gene superfamilies. This sequence conservation strongly suggests additional functions besides being a standard secretion signal; two general types of hypotheses can be formulated.

Cell biological hypotheses: refined targeting

The first class of possibilities, which we call the cell biological hypotheses is that there may be greater specialization of secretory pathways in venom ducts compared to other tissues. Thus, if conopeptide or turripeptide precursors belonging to a particular gene superfamily were packaged with a specific subset of enzymes that promoted the correct folding and posttranslational modification of peptides in that superfamily, then the conserved sequences may not just direct a precursor polypeptide to pass through the endoplasmic reticulum, but may specify docking at a specific locus to package the precursor with specific chaperones, disulfide isomerase isoforms, modification enzymes, etc. This more refined targeting may require that signal sequences be recognized by sets of specific proteins. The essence of the hypothesis is that the signal sequence does not interact just with the generic secretory system, but that additional recognition has evolved for each superfamily. This would provide positive selection for conserving specific amino acid sequences in each superfamily, although it would not explain the complete conservation of the DNA sequence as observed in Figure 2B. However, mechanisms similar to the conservation of sequence ribosomal RNA gene clusters could lead to DNA sequence conservation.

Genetic metacode hypotheses: functional constraints acting on the DNA

The other class, that we refer to as genetic metacode hypotheses is that the signal sequence region is an important determinant for modulating the rate of sequence change in the mature toxin segment. Thus, there may be some feature of the nucleic acid sequences encoding the signal sequence that increases the potential for hypermutation in the mature core sequence; one specific model would be to make the mature peptide region a preferred locus for double strand breaks or a hot spot for recombination. This type of additional function could account for the signal sequence conservation observed in some gene superfamilies even at the nucleic acid level (such as in the striking example of the absence of synonymous substitutions in the S-superfamily given in Fig. 2B). This could be rationalized by this type of hypothesis, first formulated by Caporale.[37] In one specific variation of these hypotheses, the requirement to be recognized by a biochemical mechanism that maintains the generation of hypermutability in the core peptide region is, paradoxically, the underlying explanation for signal sequence conservation. The two general classes of explanations are not mutually exclusive.

Lineage selection: favored by the rapid evolution of exogenes

As is discussed elsewhere,[38] venom genes are a readout of the biotic interactions of a particular toxoglossan species with its prey, predators, and competitors. Thus, as changes in the environment occur, novel ecological opportunities arise and new species can evolve. The ability of a lineage to opportunistically colonize new ecological niches as these arise depends on how quickly an ancestral species can become maximally fit for a new niche that becomes available. We suggest that a major component of the genetic change that accompanies speciation events in biodiverse lineages is optimization of biotic

interactions for a new ecological niche by mutationally "tuning" those gene families that determine the fitness of a newly evolving species to deal with novel predators, prey, and competitors. Thus, in a specific lineage of animals, new species can be generated rapidly if the genes most relevant to success in a new ecological niche have the potential to be rapidly tuned to the new environment. In the venomous toxoglossa, this rapid tuning potential has apparently been achieved through the fascinating combination of extremely rapid sequence change in the mature toxin region accompanied by an almost complete suppression of base changes in the signal sequence region. The challenge is understanding what leads to this pattern—the availability of several complete toxoglossan genome sequences could clearly provide some useful insights.

As has been suggested elsewhere, toxoglossan venom peptide genes belong to a larger general class that we refer to as exogenes[38]—these are the genes that encode gene products that do not function endogenously, but that act on other organisms. We have suggested that the accelerated evolution of exogenes generally accompanies the rapid speciation that occurs as an explosive adaptive radiation generates a biodiverse lineage, such as has occurred for the major lineages of venomous marine snails.

Acknowledgments

The work described in this report was supported by NIH Grant GM48677 (to B.M.O.) and by ICBG Grant 1U01TW008163, Margo Haygood (PI), Oregon Health and Science University. This work was further supported by Grant IX211904 (to M.B.A.) and IN213808–3 (to E.L.V.) from the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México (PAPIIT-UNAM), and Grants 75809 and 153915 (to E.L.V.) from the Consejo Nacional de Ciencia y Tecnología (CONACYT). We would like to thank Dr. Adolfo Gracia Gasca and the officers and crew of the RV Justo Sierra for their help in the trawling operations to obtain turrid snails during the oceanographic campaign COBERPES III. We thank Eric Schmidt for insightful discussions, and the stimulating and insightful suggestions of Lynn Caporale are also gratefully acknowledged.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Halstead, B.W. 1988. *Poisonous and Venomous Marine Animals of the World (Second Revised Edition). Princeton.* The Darwin Press, Inc. New Jersey.
2. Tu, A.T., ed. 1984. *Handbook of Natural Toxins Volume 2. Insect Poisons, Allergens, and Other Invertebrate Venoms*. Marcel Dekker, Inc. New York, NY.
3. Tu, A.T., ed. 1988. *Handbook of Natural Toxins Volume 3. Marine Toxins and Venoms*. Marcel Dekker, Inc. New York, NY.
4. Bouchet, P. *et al*. 2011. A new operational classification of the Conoidea (Gastropoda). *J. Molluscan Stud*. **77:** 273–308.
5. Olivera, B.M. *et al*. 1985. Peptide neurotoxins from fish-hunting cone snails. *Science*. **230:** 1338–1343.
6. Olivera, B.M. 1997. *Conus* venom peptides, receptor and ion channel targets and drug design: 50 million years of neuropharmacology (E.E. Just Lecture, 1996). *Mol. Biol. Cell*. **8:** 2101–2109.
7. Terlau, H. & B.M. Olivera. 2004. *Conus* venoms: a rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **84:** 41–68.
8. Hu, H. *et al*. 2011. Characterization of the Conus bullatus genome and its venom-duct transcriptome. *BMC Genom.* **12:** 60.
9. Davis, J., A. Alun Jones & R.J. Lewis. 2009. Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides.* **30:** 1222–1227.
10. Biass, D. *et al*. 2009. Comparative proteomic study of the venom of the piscivorous cone snail Conus consors. *J. Proteomics.* **72:** 210–218.
11. Terrat, Y. *et al*. 2012. High-resolution picture of a venom gland transcriptome: case study with the marine snail Conus consors. *Toxicon.* **59:** 34–46.
12. Lluisma, A.O. *et al*. 2012. Novel venom peptides from the cone snail Conus pulicarius discovered through next-generation sequencing of its venom duct transcriptome. *Marine Genom.* **5:** 43–51.
13. Hu, H. *et al*. 2012. Elucidation of the molecualr envenomation strategy of the cone snail Conus geographus through transcriptome sequencing of its venom duct. Submitted.
14. Pi, C. *et al*. 2006. Diversity and evolution of conotoxins based on gene expression profiling of *Conus litteratus*. *Genomics*. **88:** 809–819.
15. Conticello, S.G. *et al*. 2001. Mechanisms for evolving hypervariability: the case of conopeptides. *Mol. Biol. Evol.* **18:** 120–131.
16. Tucker, J.K. & M.J. Tenorio. 2009. Systematic Classification of Recent and Fossil Conoidean Gastropods: Conchbooks. Germany.
17. McIntosh, J.M. *et al*. 1995. A new family of conotoxins that blocks voltage-gated sodium channels. *J. Biol. Chem.* **270:** 16796–16802.
18. Olivera, B.M. 2002. *Conus* venom peptides: reflections from the biology of clades and species. *Annu. Rev. Ecol., Evol. System.* **33:** 25–42.
19. Woodward, S.R. *et al*. 1990. Constant and hypervariable regions in conotoxin propeptides. *EMBO J.* **1:** 1015–1020.

20. Colledge, C.J. *et al*. 1992. Precursor structure of ω-conotoxin GVIA determined from a cDNA clone. *Toxicon.* **30:** 1111–1116.
21. Van der Donk *et al*. Ribosomally synthesized and post-translationally modified peptide natural products, in Natural Product Reports. 2012. Accepted.
22. Olivera, B.M. *et al*. 1999. Speciation of cone snails and interspecific hyperdivergence of their venom peptides. Potential evolutionary significance of introns. *Ann. N.Y. Acad. Sci.* **870:** 223–237.
23. Bandyopadhyay, P.K. *et al*. 1998. Conantokin-G precursor and its role in γ-carboxylation by a vitamin K-dependent carboxylase from a *Conus* snail. *J. Biol. Chem.* **273:** 5447–5450.
24. Buczek, O., B.M. Olivera & G. Bulaj. 2004. Propeptide does not act as an intramolecular chaperone but facilitates protein disulfide isomerase-assisted folding of a conotoxin precursor. *Biochemistry.* **43:** 1093–1101.
25. Puillandre, N. *et al*. 2012. Molecular phylogeny, classification and evolution of conopeptides. *J. Mol. Evol.* Submitted.
26. England, L.J. *et al*. 1998. Inactivation of a serotonin-gated ion channel by a polypeptide toxin from marine snails. *Science.* **281:** 575–578.
27. Teichert, R.W., E.C. Jimenez & B.M. Olivera. 2005. αS-Conotoxin RVIIIA: a structurally unique conotoxin that broadly targets nicotinic acetylcholine receptors. *Biochemistry* **44:** 7897–7902.
28. Cartier, G.E. *et al*. 1996. α-Conotoxin MII (a-Ctx-MII) interaction with neuronal nicotinic acetylcholine receptors. *Soc. Neurosci. Abst.* **22:** 268.
29. Martinez, J.S. *et al*. 1995. α-Conotoxin EI, a new nicotinic acetylcholine receptor-targeted peptide. *Biochemistry.* **34:** 14519–14526.
30. Puillandre, N., M. Watkins & B.M. Olivera. 2010. Evolution of *Conus* peptide genes: duplication and positive selection in the A-superfamily. *J. Mol. Evol*. **70:** 190–202.
31. Imperial, J. *et al*. 2007. Using chemistry to reconstruct evolution: on the origins of fish-hunting in venomous cone snails. *Proc. Am. Philos. Soc*. **151:** 185–200.
32. Olivera, B.M., J. Imperial & G.P. Concepcion. 2012. Venom peptides from Conus and other Toxoglossan Marine Snails. In *Handbook of Biologically Active Peptides*. A.J. Kastin, Ed. Elsevier, Inc. Accepted.
33. Watkins, M., D.R. Hillyard & B.M. Olivera, 2006. Genes expressed in a turrid venom duct: divergence and similarity to conotoxins. *J. Mol. Evol.* **62:** 247–256.
34. Heralde, F.M., 3rd *et al*. 2008. A rapidly diverging superfamily of peptide toxins in venomous Gemmula species. *Toxicon.* **51:** 890–897.
35. Heralde, F.M., 3rd *et al*. 2010. The Indo-Pacific Gemmula species in the subfamily Turrinae: aspects of field distribution, molecular phylogeny, radular anatomy and feeding ecology. *Philip. Sci. Let*. **3:** 21–34.
36. Conticello, S.G. *et al*. 2000. Position-specific codon conservation in hypervariable gene families. *Trends Genet.* **16:** 57–59.
37. Caporale, L.H. 1984. Is there a higher level genetic code that directs evolution? *Mol. Cell. Biochem.* **64:** 5–13.
38. Olivera, B.M. 2006. Conus peptides: biodiversity-based discovery and exogenomics. *J. Biol. Chem.* **281:** 31173–31177.