THE ROYAL SOCIETY

Received 12 September 2003
Accepted 3 February 2004
Published online 20 April 2004

Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*

Thomas F. Duda Jr[1*] and Stephen R. Palumbi[2]

[1]*Naos Marine Laboratory, Smithsonian Tropical Research Institute, Box 2072, Balboa, Ancón, Republic of Panama*
[2]*Hopkins Marine Station, Stanford University, Ocean View Boulevard, Pacific Grove, CA 93950-3094, USA*

Differential expression of gene-family members is typically associated with the specific development of certain tissues and organs, but its importance in the ecological adaptation of organisms has rarely been investigated. Several specialized feeding modes have evolved within the predatory marine gastropod genus *Conus*, including molluscivory and piscivory. Based on phylogenetic investigations of *Conus* species, it has been concluded that piscivory arose at least twice in this genus. Moreover, molecular analyses of conotoxin mRNA transcripts reveal that piscivores from independent evolutionary lineages express the same subset of four-loop conotoxins, contrary to phylogenetic expectations. These results demonstrate that differential expression of gene-family members can play a key role in adaptive evolution, particularly during shifts to new ecological niches.

Keywords: gene family; adaptive evolution; differential expression; *Conus*; conotoxins

1. INTRODUCTION

The evolution of gene families has come to be regarded as an important means by which organisms evolve adaptively (Ohno 1970; Ohta 1991, 1994; Hughes 1994, 1999; Li 1997). A widely supported view is that gene products of duplicated loci can acquire or specialize for new functions because the genes that encode them are redundant (Ohno 1970, 1973; Li 1997; Hughes 1999). Among the 500+ species of the gastropod *Conus*, many have specialized diets, preying on particular taxa such as polychaetes, molluscs, hemichordates or fishes (Kohn 1959, 1968, 1981; Marsh 1971; Kohn & Nybakken 1975; Nybakken 1979; Reichelt & Kohn 1985). *Conus* uses a venom that contains a cocktail of neurotoxic peptides, termed conotoxins, to stun prey by blocking muscle and neuronal ion channels and receptors (Endean & Rudkin 1965; Olivera *et al.* 1985, 1990, 1991), and rapid immobilization of prey is essential to successful capture. Previous analyses of conotoxin gene-family evolution have shown that conotoxin loci have evolved adaptively in molluscivorous and vermivorous *Conus* species (Duda & Palumbi 1999, 2000; Conticello *et al.* 2001) and that the expression of different gene combinations fine-tunes venom composition among related species (Duda & Palumbi 2000). Such patterns suggest that selection operates to develop and maintain a venom that is most effective against particular prey.

Phylogenetic analyses of *Conus* based on a region of the mitochondrial 16S gene and a nuclear intron of a calmodulin locus have indicated that molluscivorous *Conus* species comprise a monophyletic clade, but were inconclusive with regard to the number of times piscivory has arisen in this genus (Duda *et al.* 2001; Espiritu *et al.* 2001). Although these phylogenies suggest that piscivory has evolved more than once, owing to a lack of resolution a hypothesis of the monophyly of piscivores cannot be rejected. We specifically sought to determine whether piscivory has evolved more than once by improving the ability to resolve phylogenetic relationships of piscivorous species. To accomplish this, we sequenced an additional nuclear intron in six piscivores, a molluscivore and 13 vermivores and combined these data with 16S and calmodulin intron sequences to reconstruct the phylogeny of these taxa and test whether piscivores are monophyletic.

Conotoxins are intricately related to the ability to subdue prey and so the origin(s) of piscivory and the evolution of conotoxin gene families in *Conus* are likely to be linked. To investigate the evolution of conotoxins of piscivorous species of *Conus*, we sequenced four-loop conotoxin mRNA transcripts in five piscivores and combined these with published four-loop conotoxin sequences from other fish and worm-eating species; four-loop conotoxins are one of the dominant and best-known classes of conotoxins found in the venoms of *Conus* (Olivera *et al.* 1991). We investigated the molecular evolution of these conotoxins by analysing the relationships of sequences through reconstructing gene trees and examining the patterns of nucleotide substitution among sequences.

We expect phylograms constructed from conotoxin mRNA sequences to be similar to those assembled from other loci (e.g. the mitochondrial and nuclear gene regions that were used for phylogenetic reconstruction). If piscivores are monophyletic, their conotoxins should cluster within one or multiple clades distinct from clades of vermivore conotoxins. However, if piscivores arose numerous times from vermivorous lineages, the relationships of conotoxins of piscivores and vermivores should reflect the phylogenetic relationships of the taxa from which they were derived. Failure to comply with these expectations could be the result of convergent evolution of conotoxins for use on similar prey or expression of discrete sets of conotoxin loci among species with different feeding modes. A hypothesis of the convergent evolution of the conotoxins of piscivores can be tested by determining whether any amino acid altering (i.e. non-synonymous) substitutions differentiate conotoxins of piscivores from

*Author and address for correspondence: Department of Biology, University of Washington, Seattle, WA 98195, USA (tfduda@u.washington.edu).

© 2004 The Royal Society

conotoxins of vermivores and by examining conotoxin phylograms built from levels of synonymous divergence.

To understand the association between the origin(s) of piscivory and the evolution of conotoxin gene families, we first tested the hypothesis that piscivores are monophyletic through analyses of the phylogenetic relationships of 20 *Conus* species. Next, we examined a phylogram reconstructed from 65 mRNA sequences of four-loop conotoxins from six piscivorous and six vermivorous *Conus* to assess whether the relationships of these sequences are congruent with the phylogenetic relationships of the species from which they were obtained. We also tested a hypothesis that conotoxins of piscivores evolved convergently by examining the molecular evolution of the conotoxins of piscivores and through reconstruction of conotoxin gene trees with levels of synonymous divergence. Finally, we determined whether the conotoxins of piscivores have evolved adaptively by comparing proportions of synonymous and non-synonymous substitutions among supposed recently duplicated and orthologous loci and alleles.

2. MATERIAL AND METHODS

(a) *Phylogenetic analyses*

To determine the number of times piscivory has evolved in *Conus*, we sequenced a nuclear intron located within a tubulin locus of 20 *Conus* species, including six piscivores, a molluscivore and 13 vermivores. Tubulin sequences were amplified with 5′ primers TUB3.1 (GATTTGGAGCCG GGTACCATGGA), TUBC3 (CCGGAGCCGGCAAYAAYT GGGC) or TUBCI1 (CTGCGACTGTCTGCAAGGTATGG) and 3′ primers TUB4.1 (ATACGGTCTGGGTACTCCTCG CG) or TUBCI2 (GAATGCGTCAGCTGGAAACCTGC) (TUBCI1 and TUBCI2 bind within the intron–exon boundary; all others bind within exon regions) and determined as described previously (Duda *et al.* 2001). The 16S and calmodulin sequences were determined as described previously (Duda *et al.* 2001).

We analysed separate and combined datasets with MODELTEST 3.06 (Posada & Crandall 1998) to identify the best-fitting models of nucleotide substitution for each dataset. Phylogenies were reconstructed with PAUP* 4.0 (Swofford 2000) with these datasets, and the parameters were determined with MODELTEST. Levels of support for the branches in the trees were estimated with bootstrapping methods, as implemented in PAUP*. We tested the hypothesis of the monophyly of piscivores by comparing the log-likelihood scores of trees constructed with piscivores constrained to be monophyletic and unconstrained trees as built with combined sequence data.

(b) *Specimens*

Adult animals used for conotoxin analyses were collected in the field: specimens of *C. catus* and *C. striatus* were collected in June 1998 in Oahu (Hawaii, USA); specimens of *C. ermineus* were collected in Cape Verde in July 2003; specimens of *C. purpurascens* were collected in September 2000 in Las Islas Perlas, Panama; and specimens of *C. tulipa* were collected in November 2000 in American Samoa. The venom ducts of animals from American Samoa and Panama were stored in RNAlater (Ambion Incorporated) following the manufacturer's recommendations; mRNA of animals from Hawaii was extracted directly from the venom ducts of fresh specimens. Adult specimens used for phylogenetic analyses include the specimens listed above plus specimens collected from throughout the Pacific.

(c) *Conotoxin sequences*

We constructed cDNA libraries from mRNA extracted from the venom ducts of two specimens of each of the piscivores *C. catus*, *C. ermineus*, *C. purpurascens*, *C. striatus* and *C. tulipa* as described previously (Duda & Palumbi 1999). We also included in our analyses published sequences of 52 other unique four-loop conotoxin transcripts, representing known and presumed κ- and ω-conotoxins, previously described from eight piscivorous and vermivorous species (see Colledge *et al.* 1992; Shon *et al.* 1998; Duda & Palumbi 1999, 2000; Lu *et al.* 1999; Lewis *et al.* 2000; Conticello *et al.* 2001). Because they are quite divergent from κ- and ω-conotoxins, δ-conotoxins, another class of four-loop conotoxins, were not included in our analyses.

We amplified four-loop conotoxin sequences from cDNA libraries as described previously (Duda & Palumbi 1999). The primers we used (TOX-1 and TOX-2; see Duda & Palumbi 1999) amplify, on average, *ca*. 230 bp of conotoxin transcript sequence: *ca*. 120 bp of prepro region coding sequence, 81 bp of mature-toxin coding sequence and 30 bp of 3′ untranslated region sequence. Because amplifications failed on cDNA libraries of *C. tulipa*, we used 3′ RACE (rapid amplification of cDNA ends; Frohman *et al.* 1988) with two rounds of amplifications using the TOX-1 primer (Duda & Palumbi 1999) with MOBY (AAGGATCCGTCGACATCGATAATACGACTC ACTATAGGGATTTTTTTTTTTTTTTT) and an internal 5′ primer, TOX-1.1 (CGCCGTGCTGTTCTTGACGGC), with IN2 (CGATAATACGACTCACTATAG); this method was used on the cDNA of only one specimen. Amplification products were cloned as described previously (Duda & Palumbi 1999). Approximately 50 conotoxin transcripts were sequenced from individuals of *C. catus* and *C. striatus*; 20 transcripts were sequenced from individuals of *C. ermineus*, *C. purpurascens* and *C. tulipa*. We also amplified conotoxin loci directly from the genomic DNA of the piscivore *C. striatus* with the TOX-2 primer (Duda & Palumbi 1999), which is homologous with a section of the 3′ untranslated region of four-loop conotoxins, and a primer designed within the prepro–toxin junction (GTOX-1 = CCA AGAAGTGCACGCAGACCAAT).

(d) *Conotoxin phylogram*

We aligned all four-loop conotoxin sequences by eye. We built a phylogram of four-loop conotoxins with Tamura–Nei (Tamura & Nei 1993) distances using PAUP* 4.0 (Swofford 2000). Levels of support for the branches in the phylogram were measured through bootstrapping methods as implemented in PAUP*.

(e) *Tests for convergent evolution*

To test whether the conotoxin loci of piscivores evolved convergently, we examined the types of substitution that occur on branches leading to these sequences and the similarity of predicted amino acid sequences. We also used synonymous divergence within the prepro region, a region that appears to be under selective constraints and typically shows the lowest levels of divergence among conotoxin transcript sequences (Duda & Palumbi 1999, 2000; Conticello *et al.* 2001), to reconstruct the conotoxin phylogram with MEGA2 (Kumar *et al.* 2001). We assume that a conotoxin phylogram built from synonymous divergence should reflect the relationships of sequences and not be skewed by selection (i.e. non-synonymous substitutions).

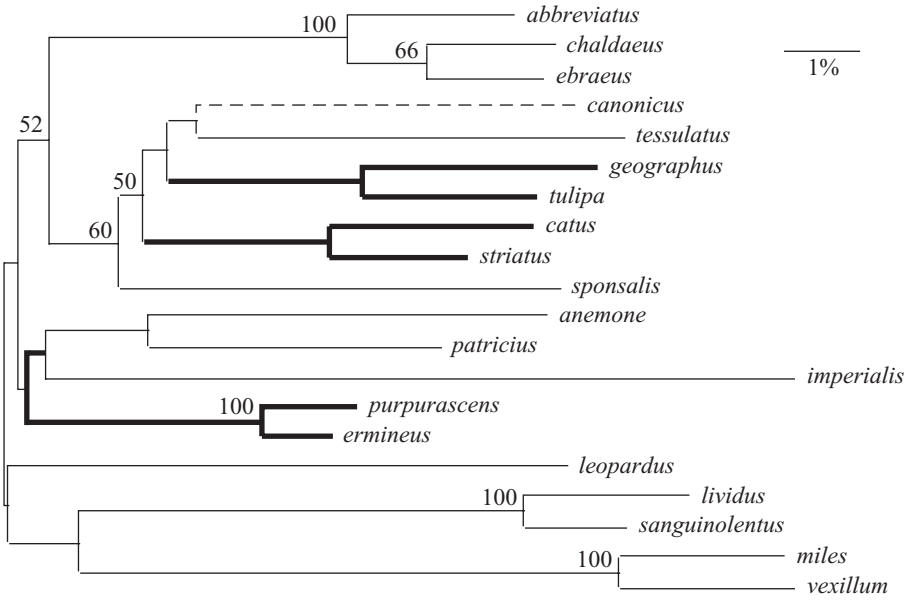

Figure 1. Phylogeny of 20 *Conus* species reconstructed by neighbour joining of HKY distances (Hasegawa *et al.* 1985). Molluscivore, dashed line; piscivore, thick black line; vermivore, thin black line.

(f) *Molecular evolution of conotoxins*

To determine whether the conotoxins of piscivores have evolved adaptively, we calculated the proportions of non-synonymous (d_N) and synonymous (d_S) substitutions per respective sites within the toxin coding regions and the ratio of d_N to d_S (ω) among four-loop conotoxins of piscivorous species of *Conus* using the maximum-likelihood pairwise methods of Goldman & Yang (1994) with PAML (Yang 1997). Comparisons were made only among sequences of piscivores that are similar, cluster together with strong bootstrap support and possibly represent recently duplicated or orthologous loci or alleles.

3. RESULTS

(a) Conus *phylogeny*

We obtained tubulin intron sequences (a maximum of 240 bp of exon sequence and, on average, *ca*. 550 bp of intron sequence) of 20 *Conus* species, including one molluscivore, six piscivores and 13 vermivores; 16S (sequence-length averages of *ca*. 450 bp) and calmodulin sequences (55 bp of exon sequence and, on average, *ca*. 225 bp of intron sequence) were also determined for some taxa or obtained from published data (Appendix A).

The *Conus* phylogeny was reconstructed by neighbour joining of HKY distances (Hasegawa *et al.* 1985) using an estimated proportion of invariable sites of 0.2275 and a gamma correction (the shape parameter of the gamma distribution) of 0.8798 for the combined 16S, calmodulin and tubulin sequence data (figure 1). The phylogram and bootstrap values from 1000 replicates were determined with PAUP* 4.0 (Swofford 2000); the phylogeny is midpoint rooted. Other distance algorithms, maximum-parsimony analyses and separate examinations of each dataset gave similar topologies.

Our results cause rejection of the hypothesis of a single origin of piscivory in *Conus* because the negative log-likelihood score of a tree constrained to make all piscivorous species monophyletic is significantly greater ($p < 0.001$) than that of the unconstrained phylogeny (figure 1). Thus, piscivory evolved at least twice in *Conus*—in the lineage that gave rise to the piscivores *C. ermineus* and *C. purpurascens* and in the lineage that gave rise to the other piscivores (figure 1).

(b) *Conotoxin sequences*

We identified 13 unique four-loop conotoxin sequences that appear to represent conotoxins or alleles from sequencing over 300 mRNA transcripts of five piscivorous species of *Conus* (Appendix A). These 13 transcripts were common (i.e. comprised a high percentage of the sequences determined through cloning) and were typically observed in both individuals analysed for each species. We also obtained several rare sequences (i.e. comprising a low percentage of those sequences determined through cloning) that were observed in only one of the two individuals analysed for each species. These sequences had one or two nucleotide differences separating them from one of the common transcripts or were mosaic sequences composed of two halves of common sequences determined for that taxon. We assumed these variants to be the result of polymerase error or amplification-induced recombination (see Duda & Palumbi 2000) and excluded them from further analyses. Because these sequences are very similar to the sequences that were analysed, their exclusion does not affect the general topology of the conotoxin phylogram or our interpretations.

In most cases, identical sequences were found from both individuals in taxa where two individuals were analysed. However, conotoxins *purpurascens*-P2a and *purpurascens*-P2b were determined from one individual; *purpurascens*-P2c was determined from the other. The similarity of these sequences and their segregation patterns in the two individuals analysed suggest that they are alleles of a single conotoxin locus (see Duda & Palumbi 2000). Conotoxins *catus*-C1a and *catus*-C1b were found in both individuals analysed, but, based on their similarity, also appear to represent alleles.

Some of the four-loop conotoxin sequences we recovered were similar to others previously obtained for the same species. Conotoxins *purpurascens*-P1 and PVIIA (as described by Shon *et al.* 1998) and *striatus*-S2 and SO5 (as described by Lu *et al.* 1999) sequence pairs are identical within the toxin coding region; the only differences between each of these sequence pairs are nucleotide substitutions within the prepro region. Conotoxins *striatus*-S1 and SO3 (as described by Lu *et al.* 1999) show seven nucleotide differences between them (two synonymous and three non-synonymous substitutions in the prepro region and one synonymous and one non-synonymous substitution in the toxin region) and possibly represent alleles; whereas *striatus*-S1 was recovered from both individuals we examined, *striatus*-SO3 was reported from analyses of the pooled mRNA of 25 specimens of *C. striatus* from China (Lu *et al.* 1999).

(c) *Conotoxin phylogram*

The phylogram of the conotoxin transcripts of six piscivorous and six vermivorous *Conus* species was reconstructed by neighbour joining of Tamura–Nei (Tamura & Nei 1993) distances with a gamma correction (gamma-distribution shape parameter) of 0.6574 among sequences (figure 2). The phylogram and bootstrap values from 1000 replicates were determined with PAUP* (Swofford 2000); the phylogram is midpoint rooted. There are several well-resolved clusters of sequences in the conotoxin phylograms. One of these clades contains all the conotoxins of piscivores and is supported by a bootstrap value of 94%. The remaining clades contain only conotoxins of vermivorous species.

(d) *Tests for convergent evolution*

To test for convergence, we examined the types of substitution that are unique to the clade of conotoxins of piscivores. If these conotoxin sequences evolved convergently, then we would expect to find several non-synonymous substitutions within the toxin coding region occurring on the branch leading to this clade that are responsible for similar amino acid sequences in the conotoxin peptides of piscivores. On the contrary, the amino acid sequences of the conotoxins of piscivores are highly divergent (figure 3), and most of the signal that is responsible for the grouping of the conotoxins of piscivores occurs within the prepro region of these transcripts. Moreover, the conotoxins of piscivores still cluster strongly together when the conotoxin phylogram is reconstructed from synonymous divergence within the prepro region.

(e) *Conotoxin sequence recovered from the genome of the piscivore* Conus striatus

We recovered one sequence from the genomic DNA of *C. striatus* (GenBank accession number AF480336; see Appendix A) that is most similar to a conotoxin from the vermivore *C. abbreviatus* (locus *abbreviatus*-A5; figure 2). These sequences differ at 14 nucleotide positions within the toxin coding region, 12–14 of which are non-synonymous (depending on the pathway of substitutions) and none of which affects cysteine codons, and no frameshift mutations.

(f) *Molecular evolution of the conotoxins of piscivores*

We estimated the proportions of non-synonymous (d_N) and synonymous (d_S) substitutions per respective site among 15 pairs of sequences. Based on the relationships of the species from which they were derived (figure 1) and the similarity of these conotoxins (figures 2 and 3), we suggest that these sequences may represent recently duplicated or orthologous loci or alleles. Conotoxins *purpurascens*-P1 and PVIIA and *striatus*-S2 and SO5 were not compared because there are no substitutions in the toxin regions of these sequences; because the toxin regions of conotoxins *striatus*-S2 and SO5 are identical, only *striatus*-S2 was compared with *striatus*-SO4. In all cases except one, ratios of d_N to d_S (ω) were greater than one (table 1). The value of ω was 0.8 for the comparison between conotoxins *striatus*-S1 and SO3 (table 1), sequences that may represent alleles in *C. striatus*. The toxin regions of these sequences differed at only two positions: a non-synonymous substitution (see figure 3; amino acid position 56 for the affected amino acid) and a synonymous substitution within the codon that encodes the final cysteine residue.

4. DISCUSSION

Phylogenetic analyses show that, although piscivory is rare among molluscs, this novel feeding mode arose more than once in the venomous gastropod genus *Conus*. Moreover, analyses of 65 distinct four-loop conotoxin transcripts from six piscivorous and six vermivorous *Conus* species reveal that conotoxin expression patterns differ among species with different feeding modes and that members of evolutionarily distinct clades of piscivores express conotoxins that are exclusive to species with this diet. Also, as has been shown for conotoxins of vermivorous and molluscivorous species of *Conus* (Duda & Palumbi 1999, 2000; Conticello *et al.* 2001), conotoxins of piscivores have also evolved adaptively.

(a) *Evolution of piscivory in* Conus

The combined-sequence phylogenetic reconstruction shows that piscivory arose at least twice in this genus, with a definite independent origin of piscivory in the lineage that gave rise to *C. ermineus* and *C. purpurascens* relative to other piscivores (figure 1). Whereas Duda *et al.* (2001) identified three clades of piscivores, Espiritu *et al.* (2001) identified four; two of the clades identified by Duda *et al.* (2001) and three of the clades identified by Espiritu *et al.* (2001) (the first contains *C. catus* and *C. striatus*, the second contains *C. geographus* and *C. tulipa* and the third of Espiritu *et al.* (2001) contains *C. ermineus* and *C. purpurascens*) have members that were included in the phylogenetic analyses we present here. Although the resolution of our phylogenetic reconstructions was powerful enough to determine that the lineage that gave rise to *C. ermineus* and *C. purpurascens* evolved piscivory independently of the evolution of piscivory in the other clades, it does not show whether piscivory also evolved independently (i.e. more than twice) in the lineages that gave rise to these other clades.

The fourth clade of piscivores identified by Espiritu *et al.* (2001) (the third identified by Duda *et al.* 2001)

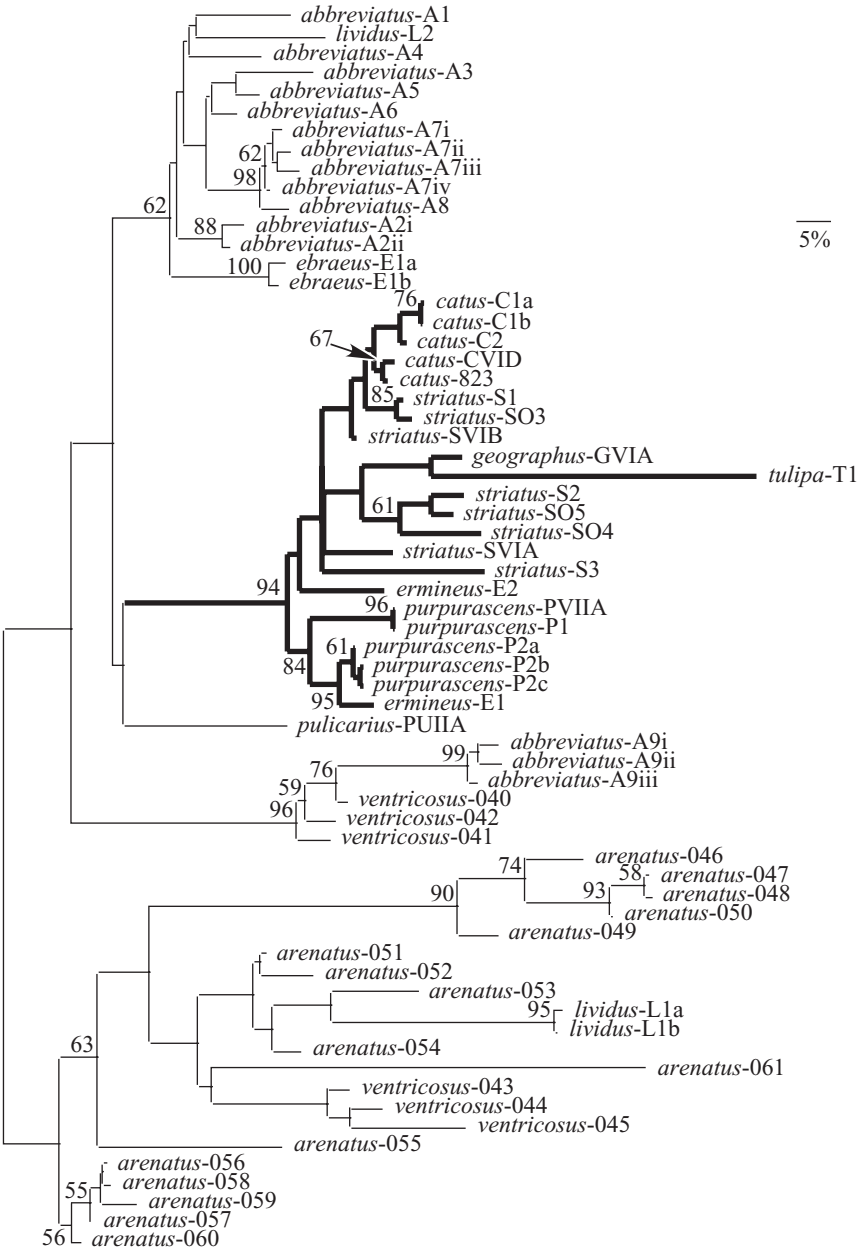

Figure 2. Molecular phylogram of 65 four-loop conotoxin transcripts from six piscivorous (thick black lines) and six vermivorous (thin black lines) *Conus* species reconstructed with neighbour joining of Tamura–Nei (Tamura & Nei 1993) distances. Sequence names include the specific epithet of the species from which they were described and (except those mentioned below) capital letters denoting the specific epithet appended with numbers to denote putative loci and lowercase letters to denote putative alleles. Published sequences of conotoxins of *C. arenatus* and *C. ventricosus* and one of *C. catus* are appended with the last three digits of their GenBank accession numbers; other published sequences of conotoxins of *C. abbreviatus*, *C. catus*, *C. ebraeus*, *C. geographus*, *C. lividus*, *C. pulicarius* and *C. striatus* are appended with names that were previously assigned.

included species not represented here because tubulin sequence data were not available for them. Separate phylogenetic analyses of mitochondrial 16S and calmodulin intron sequence data show that, while members of this clade cluster separately from other clades of piscivores, the resolution of the phylogenies is too poor to determine whether piscivory evolved in the lineage that gave rise to this clade independently of the evolution of piscivory in other lineages. Nonetheless, further phylogenetic investigations of this genus could identify as many as four origins of piscivory in *Conus*.

(b) *Molecular evolution of the conotoxins of piscivores*

As has been observed for other conotoxin sequences (Duda & Palumbi 1999, 2000; Conticello *et al.* 2001), diversifying selection has operated within the toxin coding regions of the conotoxins of piscivores, as evidenced by larger values of d_N (the proportion of non-synonymous substitutions per non-synonymous site) than d_S (the proportion of synonymous substitutions per synonymous site) in nearly all comparisons of potentially orthologous and recently duplicated loci (table 1). In addition, putative

```
                                  1111111111222222222233333333334444444444555555555566666666667777777777888888
                         1234567890123456789012345678901234567890123456789012345678901234567890123456789012345

catus-C1a                -----vvivavllltacqlitandsrgtqkhralrsdtkls-m--str-CKGKGASCRRT-SYDCCTG-SC--RSG-RCG-----
catus-C1b                -----....................................-.--...-..ST.......-.......-..--...-...-----
catus-C2                 -----....................................-.--...-.Q.R.....K.-M.N..S.-..N-.-.-...-----
catus-CVID               mkltc.................d..................-.--...-..S...K.SKL-M....S.-..SGTVV-...-----
catus-823                mkltc....................................-.--...-..S...K.SKL-M....S.-..SGTV.-...-----
ermineus-E1              -----.......f.........d...r.........t..ra-t--.n.P..P..RK.FPH-QK...N-KT.T-..--K.P-----
ermineus-E2              -----.......f.........d...r.........t....-.--l..Q.TPH.G..GLV-.T-..-.-R.S-VPRNK.E-----
geographus-GVIA          mkltc.................d...........g.t.e..-l--...-..SP.S..SP.-..N..R--..N-PYTK..YG----
purpurascens-P1          -----.......f.........d...r.........t....-l--...-.RIPNQK.FQH-LD...S-RK.N-.F-NK.V-----
purpurascens-P2a         -----.......f.........d...r.........t..ga-t--.n.P..TP.RK.FPH-QK...-.RA.I-IT--I.P-----
purpurascens-P2b         -----.......f.........d...r.........t..ra-r--.n.P..KS.RK.FPH-QK...-.RA.I-IT--I.P-----
purpurascens-P2c         -----.......f.........d...r.........t..ra-t--.n.P..KT.RK.FPH-QK...-.RA.I-IT--I.P-----
purpurascens-PVIIA       mkltc....v..f.........d...r.........t....-l--...-.RIPNQK.FQH-LD...S-RK.N-.F-NK.V-----
striatus-S1              -----.................d..................-.--...-..AA.K..S.I-A.N....-.--...-K..-----
striatus-S2              -----.................d.........s...t..v.-k--..s-.MEA.SY.GS.-TRI..-.-Y.A-YF.KK.IDYPSN
striatus-S3              -----.................e......e...........-.--l.lr.ESY.KP.GIY--N...N--A.D-PAKKT.T-----
striatus-SO3             mkltcm................d.........t...k....-.--...-..AA.KP.S.I-A.N....-.--...-K..-----
striatus-SO4             mkltc.m...............d.........s...t..v.-k--..s-.MEA.SY.GS.-TRI..-.-Y.A-YF.KK.IDYPSN
striatus-SO5             mkltcm................d.........s...t..v.-k--a.d-.IEA.NY.GP.VMKI..-.-F.S-PYSKI.MNYPKN
striatus-SVIA            mkltc.m...............e.........t...tarr.-kse...-.RSS.SP.GV.-.I-..-.-R.Y-.-.-K.T-----
striatus-SVIB            mkltc.................d.................p-.--...-..L..Q...K.-.....S.-..-G...-K..-----
tulipa-T1                ---------------.......l..k.me.r.....t....klrqvsw-.GKPE.T.GKL-YLK..S.-M.N-KANWK.LR----
```

Figure 3. Alignment of predicted amino acid sequences of the four-loop conotoxins of piscivorous species of *Conus*. Conotoxin sequence names as in figure 2. Single-letter amino acid codes are used; dots indicate identity of amino acids to those in the reference sequence above (*catus*-C1a). Amino acid positions are indicated above all sequences. Amino acids within the prepro region are in lowercase letters; those within the toxin region are in uppercase. Within the mature-toxin region only the cysteine backbone that defines four-loop conotoxins is conserved across all sequences.

Table 1. Proportions of non-synonymous (d_N) and synonymous (d_S) substitutions per respective sites (± s.e.) and the ratio of d_N to d_S (ω) among four-loop conotoxins of piscivorous species of *Conus* as calculated within the toxin coding regions.

comparison	d_S	d_N	ω
intraspecific			
catus-C1a and C1b	0.0 ± 0.2	4.4 ± 3.3	> 1
catus-C1a and C2	0.2 ± 0.1	17.4 ± 7.2	87.0
catus-C1b and C2	0.2 ± 0.4	18.5 ± 7.9	92.5
catus-CVID and 823	0.0 ± 0.7	1.6 ± 1.6	> 1
purpurascens-P2a and P2b	0.0 ± 0.3	3.9 ± 2.8	> 1
purpurascens-P2a and P2c	0.0 ± 0.0	3.5 ± 2.4	> 1
purpurascens-P2b and P2c	0.0 ± 0.0	1.7 ± 1.6	> 1
striatus-S1 and SO3	3.1 ± 3.1	2.4 ± 2.5	0.8
striatus-S2 and SO4	25.6 ± 12.1	32.0 ± 8.3	1.3
interspecific			
catus-C1a and *striatus*-S1	7.4 ± 4.9	31.7 ± 12.4	4.3
catus-C1b and *striatus*-S1	6.9 ± 5.2	35.1 ± 15.8	5.1
catus-C2 and *striatus*-S1	7.5 ± 4.8	48.9 ± 20.7	6.5
ermineus-E1 and *purpurascens*-P2a	9.5 ± 6.9	27.7 ± 8.1	2.9
ermineus-E1 and *purpurascens*-P2b	8.7 ± 4.6	30.5 ± 8.5	3.5
ermineus-E1 and *purpurascens*-P2c	9.5 ± 6.5	27.9 ± 8.2	2.9

alleles differing only at non-synonymous sites within the toxin coding region were found among loci of the piscivores *C. catus* and *C. purpurascens* (table 1; figure 3) as has been reported for the vermivore *C. ebraeus* (Duda & Palumbi 2000). Thus, the conotoxins of piscivores appear to be under the same selective pressures as those of molluscivores and vermivores and have evolved adaptively.

(c) *Relationships of conotoxins of piscivorous and vermivorous* Conus

Conotoxin gene trees (or the clades of sequences contained within them) should be similar to gene trees constructed from other loci. Observations contrary to this could result from the differential expression of conotoxin loci, though other explanations are possible (e.g.

convergent evolution, methodological or other biases associated with the identification of conotoxin transcript sequences, and horizontal transfer). The phylogram of four-loop conotoxin transcripts shows that, out of the sequences we analysed, conotoxins of piscivorous species are more similar to each other than they are to any of several distinct lineages of conotoxins of vermivorous species (figure 2). This result contrasts strongly with the proposed phylogenetic relationships of *Conus* (figure 1) and differs from results obtained from molluscivorous *Conus* species (Conticello *et al.* 2001). Unlike piscivores, molluscivores are monophyletic (Duda *et al.* 2001), yet conotoxins from mollusc-eating species do not cluster completely separately from those of vermivorous taxa (Conticello *et al.* 2001). Based on the evolutionary relationships of *Conus* species proposed from other gene sequences, it is not only unusual that the sequences of four-loop conotoxins of piscivores form a single well-supported monophyletic clade, but also peculiar that clades of conotoxins of vermivores do not contain conotoxins of piscivores (figure 2).

Although we included conotoxins from a broad range of *Conus* species in terms of their phylogenetic relationships (figure 1), conotoxins of some of the vermivorous taxa that occur on the intervening nodes between the independently derived clades of piscivores were not available (e.g. *C. anemone*, *C. imperialis*, *C. patricius* and *C. sponsalis*; figure 1). Although sequences from *C. abbreviatus* and *C. ebraeus*, two vermivorous species that also occur on intervening nodes, were included, addition of conotoxin sequences from the above taxa as well as from other members of the genus may alter the pattern of unique clustering of the conotoxins of piscivores separately from the conotoxins of vermivores. However, no new conotoxin transcript-sequence data from these species should dispel the lack of congruence between the conotoxin phylogram and the phylogenetic relationships of *Conus* that we detected with the currently available data.

(**d**) *Convergent evolution of the conotoxins of piscivores?*

Is the similarity of the conotoxin transcript sequences of piscivores the result of the convergence of conotoxins for use on similar prey? We tested for convergence by examining the nature of the substitutions responsible for the monophyly of the conotoxins of piscivores (figure 2). Convergence would be manifested by non-synonymous substitutions in the toxin coding region of the gene being grouped along the branch leading to the piscivore conotoxin clade, resulting in the conotoxins of piscivores sharing particular amino acid motifs. Instead, predicted amino acid sequences of the mature conotoxins of piscivores are identical at cysteine residues only in extreme cases (figure 3). Moreover, analyses show that the substitutions along the branch leading to the conotoxins of piscivores occur mostly in the prepro region of the transcript. Little is known about the function of the translated prepro peptide, but these results suggest that the conotoxins of piscivores were selected long ago to have particular prepro regions. A hypothesis of convergence of prepro regions is also rejected because a phylogram built from synonymous divergence of this region shows a similar topology to that presented in figure 2.

Bulaj *et al.* (2001) and Espiritu *et al.* (2001) investigated the diversity of δ-conotoxins, a different class of four-loop conotoxins from those we analysed, expressed by molluscivorous and piscivorous *Conus*. Contrary to what we observed among the four-loop conotoxins of piscivores that we analysed, these authors found that some amino acids were conserved among the δ-conotoxins of piscivorous *Conus*. Because neither group of authors specifically determined whether the transcript sequences were analogous (i.e. similar owing to convergent evolution) or homologous (i.e. similar owing to descent from a common ancestral locus), we constructed phylograms of 31 known and suspected δ-conotoxin sequences from 11 species (seven piscivores, two molluscivores and two vermivores; sequences from molluscivores and piscivores were from Bulaj *et al.* (2001) and Espiritu *et al.* (2001); data from vermivores were from Conticello *et al.* (2001)) and could not reject either of these hypotheses owing to the lack of resolution among the deeper nodes of the δ-conotoxin phylograms (trees not shown). Four-loop conotoxins are probably only a small subset of those expressed in the venoms of *Conus*. Further analyses will reveal whether the pattern we observed for the known and presumed κ- and ω-conotoxins we analysed is similar to that of other venom components.

(**e**) *Methodological or other biases?*

In this and previous studies that identified the sequences of four-loop conotoxin transcripts of piscivores and vermivores that we analysed, several features may have biased the recovery of particular conotoxin transcript sequences from different individuals and species of *Conus*. Because of the competitive and selective nature of the PCR, using this method may bias the detection of expressed members of large gene families and identify different sets of sequences from different individuals. Moreover, the identification of conotoxin transcripts with amplifications from cDNA libraries probably detects only a proportion of the four-loop conotoxin loci that are expressed by *Conus* owing to the affinity of primers for particular transcripts. *Conus* may also differentially express conotoxin loci on geographical or temporal scales. Although the above factors could cause the retrieval of an arbitrary set of conotoxins among species and in some cases among individuals of species, they should not be responsible for the observed non-random sorting of conotoxins based on the feeding modes of the species from which these sequences were derived (figure 2). Furthermore, the four-loop conotoxins of piscivores and vermivores that were included in our analyses were described by a variety of workers who used a variety of techniques (i.e. not just amplifications from cDNA libraries) or different sets of primers to detect them.

(**f**) *Evolution of conotoxin gene families: differential expression among feeding modes*

The incongruence of the conotoxin gene tree and the *Conus* phylogeny suggests that conotoxin expression patterns differ between piscivorous and non-piscivorous species of *Conus* and that, despite being polyphyletic, piscivores express a set of loci that are unique to this feeding mode. We interpret the different clusters of sequences in the phylogram (figure 2), including that which contains

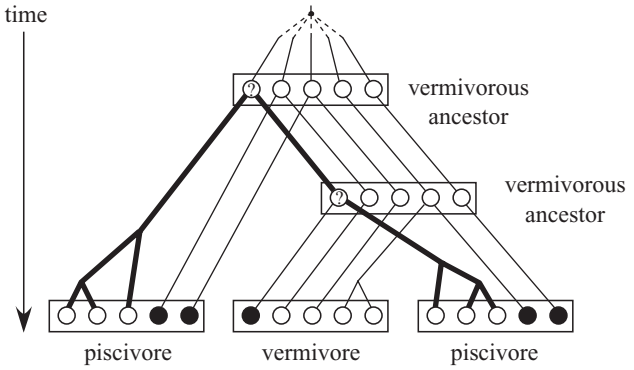

Figure 4. Proposed model of the evolution of conotoxin gene families of *Conus*: conotoxin loci in ancestral vermivorous lineages gave rise to the diversity of loci exclusively expressed by modern piscivores (open circles, expressed; filled circles, unexpressed or lost; circle with enclosed question mark, unknown expression). Circles are used to denote conotoxin loci present in modern and ancestral *Conus* lineages. Branches leading to loci expressed by modern piscivores are drawn thicker than the other branches.

the conotoxins of piscivores, as representing distinct subsets of the four-loop conotoxin gene family. As illustrated in figure 4, the conotoxins expressed by piscivorous species were probably derived from a set of loci that arose before the origination of piscivores and that was present in the ancestral vermivorous lineages that gave rise to the different clades of piscivores. Apparently, modern vermivores do not express these loci; this could be because their gene products no longer play an important role in prey capture owing to the specialized evolution of other conotoxins. Nonetheless, the model we present is corroborated by levels of divergence among conotoxins and other nuclear loci. The maximum synonymous divergence among the conotoxin sequences of piscivores (33.3% as calculated within the prepro region—the region of the transcript that shows the lowest levels of divergence) is greater than the maximum divergence among the combined calmodulin and tubulin intron sequences of all species analysed (24.7%), and suggests that the divergence of these conotoxin loci predates the divergence of all species shown in figure 1.

Although we assume that piscivores evolved from vermivorous ancestors, the specialized diets of modern *Conus*, including piscivory, may have arisen from ancestors with a broad diet, and these ancestral lineages may have possessed a diversity of conotoxins that could paralyse a wide range of prey. Indeed, *C. californicus*, a species that occurs on a basal branch of *Conus* phylogenies (see Duda *et al.* 2001; Espiritu *et al.* 2001), preys on fishes, molluscs and worms though presumably it is a vermivore (Kohn 1966). If piscivory arose more than once from lineages of generalists rather than vermivores, the similarity of conotoxin transcripts among independently derived piscivores is still best explained by the expression of distinct subsets of conotoxin gene families in species with different feeding modes as our model illustrates (figure 4). Nonetheless, the most parsimonious explanation for the evolution of diets of *Conus* is that vermivory is the ancestral feeding mode from which all specialized diets in this group arose and the apparently broad diet of *C. californicus* is a derived character (Duda *et al.* 2001).

Because the prepro and toxin regions of four-loop conotoxins are separated by an intron (see Olivera *et al.* 1999; Conticello *et al.* 2001), piscivores may express similar well-conserved prepro sequences to which are joined a diverse set of distantly related toxin sequences. Such a mechanism could in fact be responsible for the different rates of evolution that have been observed between prepro and toxin regions (see Duda & Palumbi 1999; Olivera *et al.* 1999; Conticello *et al.* 2001). However, although phylograms reconstructed with sequence data of the toxin region are not completely resolved, their topologies are similar to those obtained with prepro region and complete transcript sequences (figure 2), and the conotoxins of piscivores still cluster separately from the conotoxins of vermivores in these phylograms.

According to our model of the evolution of conotoxin expression in piscivores (figure 4), *Conus* possess a repertoire of conotoxin loci whose expression products have distinct functions, and patterns of expression are related to the specificity of conotoxins for particular prey. As predicted by this model and the phylogenetic relationships of *Conus* species (figure 1), there ought to be 'vermivore-like' conotoxin gene sequences in the genomes of piscivores as well as 'piscivore-like' conotoxins in the genomes of vermivores. However, if these loci are non-functional and thus not under selective constraints, they may have accumulated numerous substitutions and so may be difficult to identify. Although we did not detect a piscivore-like conotoxin from a vermivore, we determined a partial sequence of a vermivore-like conotoxin from the genome of the piscivore *C. striatus*. This sequence is most similar to a conotoxin transcript from the vermivore *C. abbreviatus* (*abbreviatus*-A5; figure 2) and the similarity of these sequences is congruent with phylogenetic expectations. This result corroborates our viewpoint that the incongruence of the conotoxin gene tree and *Conus* phylogeny is the result of differential expression of conotoxin loci among different feeding modes and supports the model of convergent expression patterns among distinct polyphyletic piscivorous *Conus* lineages (figure 4).

Conticello *et al.* (2001) recently proposed that functionally unimportant conotoxins may be expressed at low levels and that their expression is enhanced only when their functions become important (e.g. with shifts to new prey types or the development of resistance in prey). Our model of conotoxin gene-family evolution and expression in piscivores (figure 4) is congruent with this hypothesis. However, gene products that are not used are not under selective constraints, and the loci that encode them should not be able to evade deletion or the accumulation of mutations that would render their products non-functional. Although the vermivore-like conotoxin sequence recovered from the genomic DNA of *C. striatus* does not appear to be expressed based on the absence of this sequence in analyses of mRNA of this species, if it was expressed then its product may be functional based on the lack of substitutions that affect cysteine codons, insert a premature stop codon or cause changes in the reading frame of the transcript. This observation supports the hypothesis of Conticello *et al.* (2001). Nonetheless, because some piscivorous *Conus* consume worms as juveniles (Nybakken & Perron 1988), conotoxin expression may change during the ontogeny of piscivores, and the

vermivore-like locus of *C. striatus* may be expressed at earlier stages of development in this species and constrained by selection. Our results stress the need for characterization of additional unexpressed members of conotoxin gene families; future studies directed at this goal will further elucidate the dynamics of conotoxin gene-family evolution and the importance of differential expression in ecological adaptations.

The authors acknowledge the discussions and assistance from their colleagues, especially H. A. Lessios, A. J. Kohn, K. Zigler, L. Cortés Ortiz and F. Cipriano. Comments from five anonymous referees were very constructive and insightful. They also thank those who aided in the collection of specimens, including Z. Evora, C. G. Fiedler, D. Strang, D. Barclay, A. Calderón, K. A. del Carmen and P. S. Armstrong. They are also grateful to the Instituto Nacional de Desenvolvimento das Pescas of Cape Verde for permission to study specimens from Cape Verde. T.F.D. was supported by a Tupper Fellowship from the Smithsonian Tropical Research Institute. Partial support for some of this work also came from NSF grants to S.R.P.

APPENDIX A: DATA DEPOSITION

GenBank accession numbers (in alphabetical or numerical succession).

16S sequences: AF174140, AF480306, AF174152, AF174154, AF174155, AF174164, AY236860, AF174171, AF174173, AF174175, AF174178, AF174182, AY382021, AF480308, AF174198, AF174199, AF174202, AF174205, AF174207 and AF174209.

Calmodulin intron sequences: AF113252, AF480309, AF113261, AF113260, AF113262, AF113272, AY236861, AF113280, AF113282, AF113284, AF113287, AF113291, AY382054, AF480311, AF113308, AF113309, AF113311, AF113315, AF113317 and AF113320.

Tubulin intron sequences: AF480317–AF480322, AY236864, AF480323–AF480327, AY382063 and AF480329–AF480335.

Conotoxin sequences: *C. abbreviatus*—AF090041, AF090035, AF090008, AF089995, AF090007, AF089988, AF089997, AF089983, AF090075, AF089985, AF090074, AF090006, AF090055, AF090063 and AF090064; *C. arenatus*—AF215046–AF215061; *C. catus*—AF174214, AF174225, AF174230 and BD241823; *C. ebraeus*—AF174268 and AF174281; *C. ermineus*—AY236862 and AY236863; *C. geographus*—M84612; *C. lividus*—AF089913, AF089965 and AF089977; *C. pulicarius*—AF132130; *C. purpurascens*—AF480312—AF480315; *C. striatus*—AF174240, AF174248, AF174251, AF146348–AF146350, AF146346, AF146347 and AF480336; *C. tulipa*—AF480316; and *C. ventricosus*—AF215040–AF215045.

REFERENCES

Bulaj, G., DeLaCruz, R., Azimi-Zonooz, A., West, P., Watkins, M., Yoshikami, D. & Olivera, B. M. 2001 δ-conotoxin structure/function through a cladistic analysis. *Biochemistry* **40**, 13 201–13 208.

Colledge, C. J., Hunsperger, J. P., Imperial, J. S. & Hillyard, D. R. 1992 Precursor structure of omega-conotoxin GVIA determined from a cDNA clone. *Toxicon* **30**, 1111–1116.

Conticello, S. G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z. & Fainzilber, M. 2001 Mechanisms for evolving hypervariability: the case of conopeptides. *Mol. Biol. Evol.* **18**, 120–131.

Duda Jr, T. F. & Palumbi, S. R. 1999 Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc. Natl Acad. Sci. USA* **96**, 6820–6823.

Duda Jr, T. F. & Palumbi, S. R. 2000 Evolutionary diversification of multi-gene families: allelic selection of toxins in predatory cone snails. *Mol. Biol. Evol.* **17**, 1286–1293.

Duda Jr, T. F., Kohn, A. J. & Palumbi, S. R. 2001 Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. *Biol. J. Linn. Soc.* **73**, 391–409.

Endean, R. & Rudkin, C. 1965 Further studies of the venoms of Conidae. *Toxicon* **2**, 225–249.

Espiritu, D. J. D., Watkins, M., Monje, V. D., Cartier, G. E., Cruz, L. J. & Olivera, B. M. 2001 Venomous cone snails: molecular phylogeny and the generation of toxin diversity. *Toxicon* **39**, 1899–1916.

Frohman, M. A., Dush, M. K. & Martin, G. R. 1988 Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA* **85**, 8998–9002.

Goldman, N. & Yang, Z. 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736.

Hasegawa, M., Kishino, H. & Yano, T. 1985 Dating the human–ape split by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174.

Hughes, A. L. 1994 The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond.* B **256**, 119–124.

Hughes, A. L. 1999 *Adaptive evolution of genes and genomes*. Oxford University Press.

Kohn, A. J. 1959 The ecology of *Conus* in Hawaii. *Ecol. Monogr.* **29**, 47–90.

Kohn, A. J. 1966 Food specialization in *Conus* in Hawaii and California. *Ecology* **47**, 1041–1043.

Kohn, A. J. 1968 Microhabitats, abundance and food of *Conus* on atoll reefs in the Maldive and Chagos Islands. *Ecology* **49**, 1046–1062.

Kohn, A. J. 1981 Abundance, diversity, and resource use in an assemblage of *Conus* species in Enewetak Lagoon. *Pacific Sci.* **34**, 359–369.

Kohn, A. J. & Nybakken, J. 1975 Ecology of *Conus* on eastern Indian Ocean fringing reefs: diversity of species and resource utilization. *Mar. Biol.* **29**, 211–234.

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

Lewis, R. J. (and 14 others) 2000 Novel ω-conotoxins from *Conus catus* discriminate among neuronal calcium channel subtypes. *J. Biol. Chem.* **275**, 35 335–35 344.

Li, W.-H. 1997 *Molecular evolution*. Sunderland MA: Sinauer.

Lu, B. S., Yu, F., Zhao, D., Huang, P. T. & Huang, C. F. 1999 Conopeptides from *Conus striatus* and *Conus textile* by cDNA cloning. *Peptides* **20**, 1139–1144.

Marsh, H. 1971 Observations on the food and feeding of some vermivorous *Conus* on the Great Barrier Reef. *Veliger* **14**, 45–53.

Nybakken, J. 1979 Population characteristics and food resource utilization of *Conus* in the Sea of Cortez and West Mexico. *J. Molluscan Stud.* **45**, 82–97.

Nybakken, J. & Perron, F. 1988 Ontogenetic change in the radula of *Conus magus* (Gastropoda). *Mar. Biol.* **98**, 239–242.

Ohno, S. 1970 *Evolution by gene duplication*. New York: Springer.

Ohno, S. 1973 Ancient linkage groups and frozen accidents. *Nature* **244**, 259–262.

Ohta, T. 1991 Multigene families and the evolution of complexity. *J. Mol. Evol.* **33**, 34–41.

Ohta, T. 1994 Further examples of evolution by gene duplication revealed through DNA sequence comparisons. *Genetics* **138**, 1331–1337.

Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., De Santos, V. & Cruz, L. J. 1985 Peptide neurotoxins from fish-hunting cone snails. *Science* **230**, 1338–1343.

Olivera, B. M., Rivier, J., Clark, C., Ramilo, C. A., Corpuz, G. P., Abogadie, F. C., Mena, E. E., Woodward, S. R., Hillyard, D. R. & Cruz, L. J. 1990 Diversity of *Conus* neuropeptides. *Science* **249**, 257–263.

Olivera, B. M., Rivier, J., Scott, J. K., Hillyard, D. R. & Cruz, L. J. 1991 Conotoxins. *J. Biol. Chem.* **266**, 22 067–22 070.

Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R., Shetty, R., Watkins, M., Bandyopadhyay, P. & Hillyard, D. R. 1999 Speciation of cone snails and interspecific hyperdivergence of their venom peptides. *Ann. N. Y. Acad. Sci.* **870**, 223–237.

Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Reichelt, R. E. & Kohn, A. J. 1985 Feeding and distribution of predatory gastropods on some Great Barrier reef platforms. In *Proc. 5th Int. Coral Reef Congr.* vol. 5, pp. 191–196. Moorea, French Polynesia: Antenne Museum—EPHE.

Shon, K. J. (and 10 others) 1998 κ-Conotoxin PVIIA is a peptide inhibiting the shaker K+ channel. *J. Biol. Chem.* **273**, 33–38.

Swofford, D. L. 2000 *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*, v. 4. Sunderland: Sinauer.

Tamura, K. & Nei, M. 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526.

Yang, Z. 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.