Published in final edited form as:
Mol Phylogenet Evol. 2012 October ; 65(1): 335–338. doi:10.1016/j.ympev.2012.06.014.

A very short, functionally constrained sequence diagnoses cone snails in several Conasprella clades

Nicole J. Kraus, **Maren Watkins**, **Pradip K. Bandyopadhyay**, **Jon Seger**, **Baldomero M. Olivera**, and **Patrice Showers Corneli***
Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112

Abstract

The traditional taxonomy of ca. 700 cone snails assigns all species to a single genus, *Conus* Linnaeus, 1758. However an increasing body of evidence suggest that some belong to a genetically distinct clade alternatively referred to as the *Conasprella* (Thiele, 1929), Previously we showed that a short (259 bp) conserved intronic sequence (CIS) of the γ-glutamyl carboxylase gene (intron 9) is surprisingly informative for delineating deep phylogenetic relationships among other *Conus* snails (Kraus, et al. 2011). In this work, we once again use intron 9 (338 bp) to easily resolve problemaric relationships among the *Conasprellans*. Counterintuitively, we show that these relationships can be inferred from just 39 synapomorphic isites. The sequence is so well conserved that conflicting sites do not obscure the few informative sites that provide clear phylogenetic signal.

Unexpectedly we also found that intron 9 unambiguously distinguishes *Conasprella* species from the *Conus* species studied earlier. The respective alignments are so different from one another that the sequences from the two groups cannot be aligned and thus a phylogeny describing the genetic relationship between the formerly desginated congeners cannot be inferred from these data alone. This lack of homology between the intronic sequences belonging to each group clearly shows that they are separated by considerable evolutionary history.

Keywords

nuclear genes; conserved intron; *Conus*; *Conasprella*; evolution

Introduction

The venomous cone snails have traditionally been assigned to a single large genus, *Conus*, Linnaeus 1758 *c*omprising about 700 species. Attempts based on shell morphology to clarify the phylogeny of the genus have left the evolutionary history of *Conus* species largely unresolved (for an overview, see Rockel et al., 1995). However, recent molecular data divides the species in this large genus into two major clades, separated by considerable evolutionary distance. A substantial number of species, clearly divergent from the bulk of the *Conus* clade has been referred to as the "*Conasprella* clade" (Bandyopadhyay et al.,

*Corresponding author: Patrice Showers Corneli, 257 South 1400 East, Salt Lake City, UT 84112, 801-581-8370, Fax: 801-585-5010, corneli@biology.utah.edu.

2008), the “small major clade” (Williams & Duda, 2008), or “Clade 21” (Puillandre et al., 2008). A recent taxonomic proposal separates these species into a separate family from Conidae, Conolithidae (Tucker and Tenorio, 2009) that also comprises some other cone snail species likely to be only distantly related to the *Conasprella* clade (e.g., “*Conus*” *californicus, “Conus” profundorum*). An even more recent phylogenetic proposal (Puillandre et al., 2011b; Bouchet et al., 2011) assigns these species to the genus *Conasprella*, in the family Conidae (with the majority of cone snails remaining in the genus *Conus*). This is the taxonomic framework that we have implicitly adopted in the discussion of our data.

The first indication that this group (*Conasprella*) was distinct from the bulk of *Conus* species was the discovery that the deep-water IndoPacific species *Conus memiae* is genetically distant from other *Conus* species (Espiritu, et al., 2001). Establishing this as a widespread group, it was later discovered that based on molecular criteria, other deep-water IndoPacific species and a significant complement of Panamic and Caribbean species are similarly unrelated to most *Conus* species, and closely related to *Conus memiae* (Duda and Kohn, 2005; Puillandre et al., 2008; Williams and Duda, 2008; Bandyopadhyay et al., 2008; Olivera and Biggs, 2010).

The complete sequence of the mitochondrial genome of *Conus textile* revealed an intergenic region between the coxI and coxII genes (Bandyopadhyay et al., 2008). In most neogastropods, including other conoidean groups (Turridae and others), this region ranges in length from 0-40 base pairs. However, among cone snails, this region is significantly longer (from 130-170 base pairs in most species). The discovery that within *Conasprella*, the length is also much more variable than in other *Conus* species provided an additional criterion for assigning species to the *Conasprella* clade. Thus, in *Conus jaspideus*, for example, the intergenic region is over 500 nucleotides in length.

Here we demonstrate that a remarkably small molecular sequence can be used to unequivocally separate species in *Conasprella* from standard *Conus*. We previously reported the unusually strong phylogenetic signal for standard *Conus* species from a very short region of the intron 9 from the gamma-glutamyl carboxylase gene sequences (Imperial et. al., 2007; Krause et al., 2010). Here we evaluate the feasibility of using this intron as a marker for species in *Conasprella*; these data provide further evidence for the marked divergence of species in *Conasprella* from other *Conidae*. Despite its short length, the conserved intronic sequence interval is shown to be useful for clarifying molecular phylogenetic relationships within the *Conasprella* at the species level.

Methods

We analyzed a short portion of “intron 9” of the γ-glutamyl carboxylase gene (Table 1) of 15 *Conasprella* species using methods described for *Conus* species in Krause, et al., (2011). DNA sequence accession numbers for intron 9 are listed in Supplementary Table 1 along with GenBank accession numbers for each of the other genes studied.

Sequences alignment, model estimation, maximum likelihood, maximum parsimony and Bayesian inference of phylogenies for Intron 9, 12S rRNA and COI sequences are as described in Kraus, et al. 2009

Results

Analysis of Intron 9 sequences

Compared to the Conus species studied earlier (Krause, et al. 2011), the analysis of intron 9 sequences from species in the *Conasprella* clade reveals a similarly short (338 bp) conserved interval (Supplementary Table 2). There is, however, no apparent homology with *Conus* species as the two groups align to separate and non-overlapping regions of the full *Conus textile* sequence (Supplementary Figure 1). Hence although *Conus* and *Conasprella* species are clearly highly divergent, inferring the evolutionary history of the two major clades from the intron 9 sequences alone is not possible.

Within the *Conasprella*, however, the short intron 9 region is highly informative phylogenetically yielding a remarkably well resolved tree (Figure 1). With the exception of the uncertain placement of *Conus arcuatus* and *Conus ichnoseana*, these species fall into three well-supported subgenera: Conasprella (*Conasprella memiae/pagodus*) Fusiconus (*Conasprella longurionis/orbignyi*) and Ximeniconus (*Conasprella mahogani/tornatus*). In contrast the trees inferred from the commonly used 12SrRNA sequence leaves the relationship among the three subgenera unresolved (Supplementary Figure 2). The COI alignment (Supplementary Figure 3) is useful for distinguishing among the three clades but otherwise leaves the relationships within the *Conasprella* species unresolved with very little support from either Bayesian or Maximum Likelihood analysis.

These observations are not surprising given that the homology is higher and the resolution is far better for the tree inferred from the short Intron 9 CIS alone (338 base pairs, 59 informative sites, rescaled consistency index=0.65) than for the longer sequences from 12S rRNA (556 base pairs, 99 informative sites, RI =0.37) and COI (709 base pairs, 100 informative sites, RI =0.33).

Maximum resolution is inferred from the long concatenated sequences of 12S, COI and intron 9 (Supplementarry Figure 4) although concatenation increases conflict among the sites and so increases phylogenetic noise (sampling error) and the level of homology (RI=0.27).

Discussion

We demonstrate that the *Conasprella* γ-glutamyl carboxylase intron 9 CIS sequences are easily aligned along a conserved interval so that a consensus sequence can be defined (“the *Conasprella* CIS consensus”). This very short sequence is sufficiently variable to define most evolutionary relationships among *Conasprella* species. Furthermore the variable sites appear to have changed systematically instead of randomly: most of the sequence is highly conserved so that very few sites conflict with the 39 synapomorphic sites (Figure 1b) that define the optimal tree.

There is some correlation between shell morphology (Figure 1a) and biogeography in the different branches of *Conasprella* that have been defined using Intron 9 and other molecular markers. Two of the well-defined subgenera, Conasprella and Fusiconus are primarily Indo-Pacific and can readily be distinguished on the basis of shell morphology; most species in the Conasprella subgenera are biconic, with a fairly broad body whorl, while Fusiconus has an extremely slender body whorl (these are known to shell collectors as the “needle cones”). The third group, Ximeniconus, is exclusively a new world group, absent in the Indo-Pacific.

Most species in the *Conasprella* clade (broadly defined) occur offshore, in relatively deep water in the IndoPacific. The shallow-water IndoPacific fauna of cone snails comprise exclusively species from the large clade of *Conus*, with the total absence of any *Conasprella*. At increasing depths *Conasprella* species begin to play a more and more significant role. This is quite apparent in the material from the Aurora Expedition carried out a few years ago off the eastern coast of Luzon (Puillandre et al., 2011a). However, in the new world (both in the Panamic and the Western Atlantic fauna), species in the *Conasprella* clade do make a major contribution to the shallow-water Conidae; some species in the *Conasprella* clade (e.g., *mahogani*, *jaspideus*) can be found in large numbers in the intertidal zone.

In the previous work on Intron 9 CIS of *Conus* (Kraus et al., 2010), we noted an unusual pattern of evolution with implications for functionality. For an intronic sequence, the CIS is remarkably conserved within (but not across) major clades, suggesting that the exact CIS sequence is subject to strong selection. However, the episodic nature of base changes is another notable feature, with an enrichment of mutations during periods that define clade divergence. Interestingly, much earlier Allan Wilson (Irwin and Wilson, 1989) and later others (Gillespie, 1993, Messier and Stewart, 1997) had noted this episodic pattern of sequence evolution.

Given the large divergence between *Conasprella* and the major clade of *Conus* (as estimated by the standard genetic markers), it is not surprising that their CIS sequences in intron 9 of the gamma-glutamyl carboxylase gene cannot be aligned. Within *Conasprella*, this sequence is nevertheless conserved. Thus, despite the complete lack of any apparent sequence homology between the *Conus* and *Conasprella* CIS sequences, it seems quite possible that the Intron 9 CIS in *Conus* and *Conasprella* have homologous functions. As we discussed previously (Kraus, et al., 2011) if proteins that regulate posttranscriptional events recognize the CIS sequences, the episodic nature of base changes observed within otherwise conserved CIS sequences can be explained. The divergence of species clades may involve changes in posttranscriptional regulation; however, once a clade is established and a pattern of post-transcriptional regulation is in place, then purifying CIS selection could conserve the sequence over long periods of time.

As has been noted previously (Olivera and Biggs, 2010), despite the great evolutionary distance between the major clade of *Conus* and the *Conasprella* clade, it can often be quite a challenge to unequivocally assign species to either *Conus*, or *Conasprella*. In the analysis of *Conus praecellens* and morphologically similar forms, it was pointed out that based on shell morphology, some *Conasprella* species look closely similar to *Conus praecellens* and its

relatives, but based on molecular phylogeny, these are highly divergent from each other. In fact the very short phylogenetcally informative intron CIS sequences provide a diagnostic tool for distinguishing *Conasprella* from *Conus* species. This may, in turn, give insight into which subtle morphological features really do distinguish the two groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the National Institutes of Health, GM 48677. The field collection for some of the specimens was done in collaboration with a team supported by the Philippine Mollusk Symbiont International Cooperative Biodiversity Grant (NIH 1U01TW008163-01) from the Fogarty International Center, National Institutes of Health. We are grateful to fine critiques of the manuscript from two reviewers and the editor.

References

Bandyopadhyay PK, Stevenson BJ, Ownby JP, Cady MT, Watkins M, Olivera BM. The mitochondrial genome of *Conus* textile, coxI-coxII intergenic sequences and Conoidean evolution. Molecular Phylogenetic Evolution. 2008; 46:215–223.

Bouchet P, Kantor YI, Sysoev A, Puillandre N. A new operational classification of the Conoidea (Gastropoda). Jour of Mollus Stud. 2011; 77:273–308.

Duda TF Jr, Kohn AJ. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Mol Phylogenet Evol. 2005; 34:257–272. [PubMed: 15619440]

Espiritu DJD, Watkins M, Dia-Monje V, Cartier GE, Cruz LJ, Olivera BM. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon. 2001; 39:1899–1916. [PubMed: 11600154]

Gillespie J. Episodic Evolution of RNA Viruses. PNAS. 1993; 90:10411–10412. [PubMed: 8248122]

Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology. 2003; 52:696–704. [PubMed: 14530136]

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. Bayesian inference of phylogeny and its impact on evolutionary biology. Science. 2001; 294:2310–2314. [PubMed: 11743192]

Imperial J, Silverton N, Olivera BM, Bandyopadhyay P, Sporning A, Ferber M, et al. Using chemistry to reconstruct evolution : On the origins of fish-hunting in venomous cone snails. Proceedings of the American Philosophical Society. 2007; 151:185–200.

Imperial JS, Kantor Y, Watkins M, Heralde FM 3rd, Stevenson B, Chen P, et al. Venomous auger snail Hastula (Impages) hectica (Linnaeus, 1758): molecular phylogeny, foregut anatomy and comparative toxinology. Journal of Experimenta Zoology Part B: Molecular and Developmental Evolution. 2007; 308:744–756.

Irwin DM, Wilson AC. Multiple cDNA sequences and the Evolution of Bovine Stomach Lysozyme. J Biol Chem. 1989; 264(19):11367–8. [PubMed: 2500440]

Kraus NJ, Corneli PS, Watkins M, Bandyopadhyay PK, Seger J, Olivera BM. Against expectation: A short sequence with high signal elucidates cone snail phylogeny. Mol Phylogenet Evol. 2011; 58(2):383–389. [PubMed: 21147238]

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

Maddison DR, Maddison WP. MacClade 4.08. 2005

Messier W, Stewart CB. Episodic Adaptive Evolution of Primate Lysozymes. Nature. 1997; 385:151–154. [PubMed: 8990116]

Moretti S, Wilm A, Higgins DG, Xenarios I, Notredame C. R-coffee a web server for accurately aligning noncoding RNA sequences. Nucleic Acids Res. 2008 (July);36

Nam HH, Corneli PS, Watkins M, Olivera B, Bandyopadhyay P. Multiple genes elucidate the evolution of venomous snail-hunting *Conus* species. Mol Phylogenet Evol. 2009; 53:645–652. [PubMed: 19616106]

Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000; 302:205–217. [PubMed: 10964570]

Olivera BM, Biggs JS. Defining a clade by morphological, molecular, and toxinological criteria: distinctive forms related to *Conus* praecellens A. Adams, 1854 (Gastopoda: Conidae). The Nautilus. 2010; 124(1):1–19. [PubMed: 20835370]

Puillandre N, Samadi S, Boisselier MC, Sysoev AV, Kantor YI, Cruaud C, et al. Starting to unravel the toxoglossan knot: molecular phylogeny of the “turrids” (Neogastropoda: Conoidea). Molecular Phylogenetic Evolution. 2008; 47:1122–1134.

Puillandre N, Meyer CP, Bouchet P, Olivera B. Genetic Divergence and Geographic Variation in the deep-water *Conus* Orbignyi Complex (Mollusca: Conoidea). Zoological scripta. 2011a; 40(4): 350–363.

Puillandre N, Kantor I, Sysoev I, Couloux A, Meyer C, Rawlings T, Todd JA, Bouchet P. The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). Journal of Molluscan Studies. 2011b; 77:259–272.

Röckel, D.; Korn, W.; Kohn, AJ. Manual of the Living Conidae(Vol I: Indo-Pacific Region). Wiesbaden, Germany: Verlag Christa Hemmen; 1995.

Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19:1572–1574. [PubMed: 12912839]

Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 2008; 57:758–771. [PubMed: 18853362]

Swofford, DL. PAUP/. Phylogenetic Analysis Using Parsimony (/and Other Methods) Version 4.10. Sinauer Associates; Sunderland, Massachusetts: 2002.

Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. Am Math Soc: Lect Math Life Sci. 1986; 17:57–86.

Thiele. 1929

Tucker, JK.; Tenorio, MJ. Systematic classification of Recent and fossil conoidean gastropods. Conchbooks; 2009.

Williams ST, Duda TF Jr. Did tectonic activity stimulate oligo-miocene speciation in the Indo-West Pacific? Evolution. 2008; 62:1618–1634. [PubMed: 18410535]

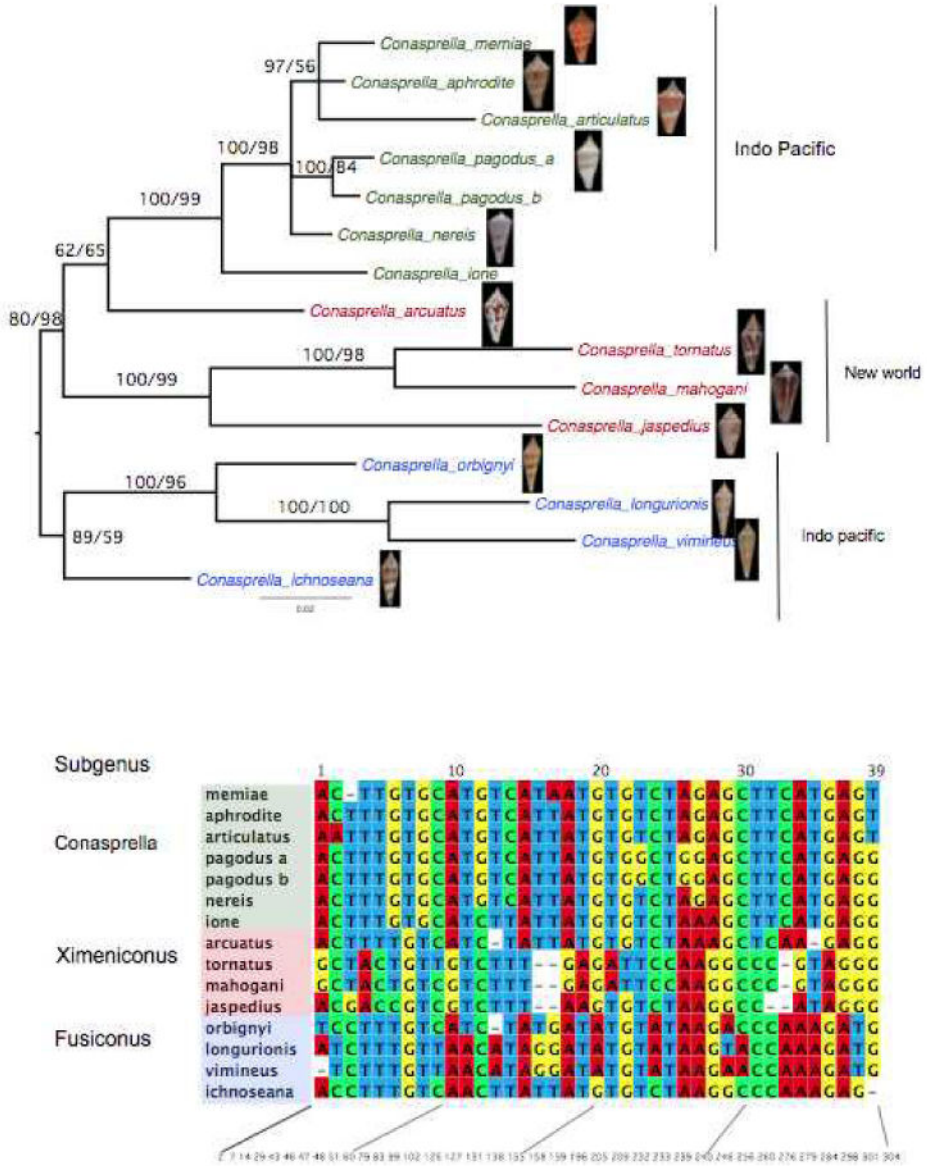

Figure 1.

A. Bayesian phylogenetic tree inferred from a conserved region of intron9 from the γ-glutamyl carboxylase gene of 15 *Conasprella* species. To indicate clade support, branches are labeled with Bayes posterior probabilities (left) and with maximum likelihood bootstrap values (right). Branches with support values that fall below 50% are not labeled. Species are colored by subgenera defined by morphology and geography: Conasprella (green), Ximeniconus (red) and Fusiconus (blue). B. The 39 informative sites from a 338 bp portion of the intron 9 of the gamma-glutamyl carboxylase gene that define the evolutionary history of the *Conasprella* species within their subgenera.

Table1

Aligned exons 9 and 10 within the γ-glutamyl carboxylase gene used in this study of *Conasprella* species and our previous study of *Conus* species (Kraus, et al. 2011).