

Exon-Capture-Based Phylogeny and Diversification of the Venomous Gastropods (Neogastropoda, Conoidea)

Jawad Abdelkrim,^{1,2} Laetitia Aznar-Cormano,^{1,2} Alexander E. Fedosov,³ Yuri I. Kantor,³ Pierre Lozouet,⁴ Mark A. Phuong,⁵ Paul Zaharias,² and Nicolas Puillandre^{*2}

¹Outils et Méthodes de la Systématique Intégrative (OMSI) UMS 2700, Muséum National d'Histoire Naturelle, Paris, France

²Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 26, 75005 Paris, France

³A.N. Severtzov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninski prospect 33, 119071 Moscow, Russian Federation

⁴Muséum National d'Histoire Naturelle, Direction des Collections, 55, rue Buffon, 75005 Paris, France

⁵Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

*Corresponding author: E-mail: puillandre@mnhn.fr.

Associate editor: Nicolas Vidal

Abstract

Transcriptome-based exon capture methods provide an approach to recover several hundred markers from genomic DNA, allowing for robust phylogenetic estimation at deep timescales. We applied this method to a highly diverse group of venomous marine snails, Conoidea, for which published phylogenetic trees remain mostly unresolved for the deeper nodes. We targeted 850 protein coding genes (678,322 bp) in ca. 120 samples, spanning all (except one) known families of Conoidea and a broad selection of non-Conoidea neogastropods. The capture was successful for most samples, although capture efficiency decreased when DNA libraries were of insufficient quality and/or quantity (dried samples or low starting DNA concentration) and when targeting the most divergent lineages. An average of 75.4% of proteins was recovered, and the resulting tree, reconstructed using both supermatrix (IQ-tree) and supertree (Astral-II, combined with the Weighted Statistical Binning method) approaches, are almost fully supported. A reconstructed fossil-calibrated tree dates the origin of Conoidea to the Lower Cretaceous. We provide descriptions for two new families. The phylogeny revealed in this study provides a robust framework to reinterpret changes in Conoidea anatomy through time. Finally, we used the phylogeny to test the impact of the venom gland and radular type on diversification rates. Our analyses revealed that repeated losses of the venom gland had no effect on diversification rates, while families with a breadth of radula types showed increases in diversification rates, thus suggesting that trophic ecology may have an impact on the evolution of Conoidea.

Key words: exon-capture, Conoidea, phylogeny, supertree.

Introduction

Next-generation sequencing (NGS) techniques have enabled the resolution of deep phylogenetic relationships by facilitating the rapid acquisition of genetic markers across divergent taxa (Lemmon and Lemmon 2013; McCormack et al. 2013). This field of research has been particularly active for non-model organisms, for which genomic and transcriptomic data are scarce (Bi et al. 2012; Cariou et al. 2013; da Fonseca et al. 2016; Yeates et al. 2016). In molluscs, a major bottleneck in applying NGS techniques has been the identification of suitable markers for phylogenetic inference (Schrödl and Stöger 2014). As such, phylogenetic analyses applying NGS techniques have been limited to mitogenomes (Osca et al. 2015; Lee et al. 2016), transcriptome sequencing (Kocot et al. 2011; Smith et al. 2011; Zapata et al. 2014; González et al. 2015; Tanner et al. 2017), or RAD-sequencing (Combosch et al. 2017). While they provided better-resolved phylogenetic trees compared with the traditional Sanger sequencing of a few mitochondrial and nuclear genes, each have their own

limitations. For example, resulting trees can be potentially biased when only mitochondrial data are used (Platt et al. 2017). Further, transcriptome-based approaches need RNA-preserved samples and the associated costs may be prohibitive to laboratories with more limited budgets.

Transcriptome-based exon capture methods provide an approach to recover hundreds of genetic markers from genomic DNA, allowing for robust phylogenetic estimation and does not rely on RNA-preserved or fresh samples (Bi et al. 2012; Hugall et al. 2016; Teasdale et al. 2016). Furthermore, because it targets short DNA fragments (typically between 100 and 400 bp), older specimens that were not preserved in ideal conditions for molecular analysis can also be sequenced. In molluscs, 250 years of taxonomy based on shell collecting have cataloged and preserved an impressive amount of material. Although molluscs are traditionally preserved dried (Bouchet and Strong 2010), recent efforts in the framework of the “La Planète Revisitée” and the “Tropical Deeps-Sea Benthos” expedition programs (<http://expeditions.mnhn.fr>;

last accessed July 29, 2018) have amassed a large collection of alcohol-preserved samples. The high quantity of dry and alcohol-preserved material, together with their huge species diversity, the lack of genomic data and the ease of generating transcriptome data, makes Mollusca a good model to test the applicability of the exon-capture approach to nonmodel organisms.

Conoidea is a group of hyperdiverse venomous gastropods (5,000 described species for a total estimated ~ 10 – $20,000$ —Bouchet et al. 2009), distributed across all oceans, latitudes, and depths. Conoideans, and in particular cone snails, are one of the most studied groups of animals for the toxins they produce. These toxins are highly diversified and specific, and some of them are used—or about to be used—as therapeutics (Prashanth et al. 2014). In a series of articles, several molecular phylogenies were published (Puillandre et al. 2008; Puillandre et al. 2011) followed by subsequent descriptions of new families and the publication of a new classification (Bouchet et al. 2011; Kantor et al. 2012) that were deeply modified compared with the previous classifications. These phylogenies included a representative (although not complete) sampling of the superfamily diversity (WORMS 2018). However, these phylogenies are all based on a few Sanger-sequenced markers, typically the mitochondrial COI (cytochrome *c* oxidase subunit I), 16S rRNA and 12S rRNA genes and the nuclear 28S rRNA, 18S rRNA, and histone H3 genes. While the analyses of these loci recognized clades from the species to family levels (i.e., up to 50 Ma, roughly), their resolving power was limited as they were unable to define older relationships. Researchers working on conoideans now have a phylogenetic classification for the group, but understanding the processes underlying the evolutionary success of the group and of the toxins they produce requires a better resolved phylogenetic tree. Recently, Uribe et al. (2018) proposed a phylogeny based on full mitogenomes, and while it was more resolved compared with previously published trees, sampling of conoidean diversity was still limited.

Here, we aim to generate the most complete and resolved phylogeny of Conoidea published so far, including representatives of most of the known family-level lineages by applying an exon-capture strategy to target thousands of loci. We explored several approaches of tree reconstructions designed for very large data sets. We also identified, for the first time, several fossils that can be attributed to the molecularly defined lineages and used these fossils to reconstruct a time-calibrated tree. The obtained phylogeny is almost fully resolved and supported, reveals new family-level lineages of conoideans and unexpected phylogenetic relationships between the families, and modifies some of the traditionally well-defined groups within the superfamily. Such phylogenies can then be used as a framework to test hypotheses related to the evolutionary success of the group. To illustrate this idea, we analyzed the variability of the diversification rates among conoidean families and tested the relationships between diversification rates and the evolution of two traits (venom gland and radula) related to the venom apparatus, a structure thought to contribute to the evolutionary success of the conoideans.

Results

Exon Capture

We used available Conoidean transcriptome data from the literature and online resources to identify 850 protein coding genes. We tentatively captured 120 samples representing the diversity of the Conoidea, as well as a selection of non-Conoidean neogastropod lineages. Two samples failed at the library preparation step (because of very low quantity of DNA available) and were thus not sequenced: the unique sample of Bouchetispiridae and an unidentified turrid, “HORS_MNHN_Stahlschmidt_9.” Both of them were preserved dried. For the remaining 118 samples, the number of bases recovered, the depth of coverage, the % of reads on target, the % of duplicates, the number of exons, the % of protein-coding genes recovered and the % of missing data in the data set DS2a (see the Materials and Methods for a description of the different data sets) are provided in the [supplementary material 1, Supplementary Material](#) online.

There are significant differences in all capture and sequencing efficiency metrics between the samples processed in the first and in the second batch ([fig. 1](#) and [table 1](#)). The first batch showed systematically better results, except for the % of duplicates. Similarly, a significant difference was detected between samples preserved dried and preserved in alcohol with the alcohol-preserved samples performing better in terms of the % of duplicates and the % of missing data, but the number of bases and the number of exons recovered are only marginally different (P value = 0.044). In contrast, the depth of coverage, the % of reads on target, and the % of proteins recovered were not different. The ingroup versus outgroup samples also present significant differences in the number of bases recovered, the depth of coverage, the number of exons recovered, and the % of missing data (ingroup taxa captured better), as well as the shallow water versus deep water samples for most of the parameters (shallow water samples captured better). The microwaved versus nonmicrowaved samples did not present any significant difference.

The final data set contained 4,376 exons when we applied a threshold of at least four samples per exon alignment and 1,340 exons when we applied a threshold of at least 60 samples per alignment (see Materials and Methods).

Phylogenetic Analyses

In addition to the two samples that were not sequenced (see above), seven other dried preserved samples were included in the data set. Among them, two have been obviously contaminated by another sample from the data set, as revealed by the likelihood analysis of the total data set (DS1ML): the sample of *Kurilohadalia elongata* had almost identical sequences with the two samples of Belomitridae and the sample of *Cruziturrucula arcuata* had almost identical sequences with *Polystira florencae* (Turridae). Additionally, the phylogenetic position of one specimen (IM-2013-55830, 99.5% of missing data) drastically changed from one analysis to another. We thus decided to remove the samples with $>99\%$ of missing

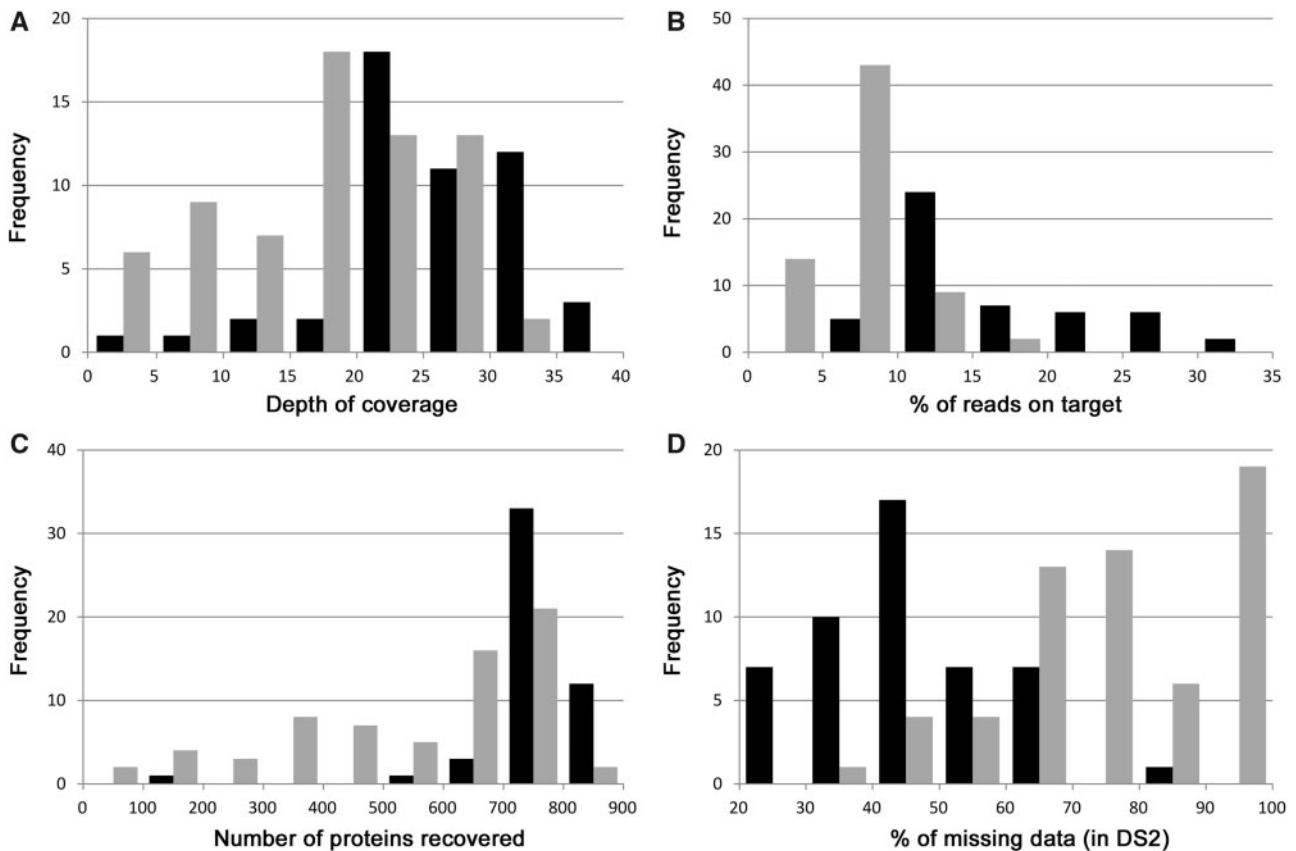


Fig. 1. Distribution of the samples in batch 1 (black bars) and 2 (gray bars) for different parameters of capture and sequencing efficiency. (A) Depth of coverage; (B) % of reads on target; (C) number of proteins recovered; and (D) % of missing data in the DS2.

Table 1. Values of the Parameters Used to Estimate the Capture and Sequencing Efficiency.

	# Bases Recovered	Depth of Coverage	% Reads on Target	% Duplicates	# Exons Recovered	% Proteins Recovered	% of Missing Data (in the DS2a)
Average batch 1	374,333	17	26	75	2,916	89	45
Average batch 2	199,850	7	18	76	1,593	66	76
Student's <i>t</i> -test batch 1/2	6.9^{-16***}	7.8^{-14***}	1^{-07***}	0.6	1.6^{-14***}	5.2^{-10***}	9.2^{-19***}
Average dry samples	107,885	7	12	86	859	46	94
Average alcohol-preserved samples (batch 2)	210,404	8	18	75	1,677	68	75
Student's <i>t</i> -test dry/alcohol	0.044*	0.4	0.2	0.003***	0.044*	0.06	0.003***
Average shallow samples	310,104	13	24	74	2,449	80	56
Average deep samples	245,806	10	19	76	1,927	72	67
Student's <i>t</i> -test shallow/deep	0.006**	0.004***	0.002***	0.2	0.004***	0.024*	0.011*
Average microwaved samples	280,726	12	23	74	2,219	77	62
Average nonmicrowaved samples	266,098	11	20	77	2,082	74	63
Student's <i>t</i> -test MW/non-MW	0.6	0.2	0.2	0.028*	0.5	0.5	0.9
Average ingroup	292,586	12	21	75	2,298	77	58
Average outgroup	213,351	9	21	76	1,689	69	75
Student's <i>t</i> -test ingroup/outgroup	0.003***	0.006***	0.7	0.8	0.003***	0.065	2^{-04***}
Average total	273,784	11.4	21.1	75.6	2,154	75.4	62.5
Variance total	135,566	6.5	8.8	7.9	1,068	22.9	21.8

NOTE.—The parameters for several subsets of samples (the two batches of library preparation and sequencing, the dry and alcohol-preserved samples, the samples collected in shallow and deep waters, the samples microwaved or not in the field, and the ingroup and outgroup samples) are compared with a Student's *t*-test.

p*-value < 0.05; *p*-value < 0.01; ****p*-value < 0.005.

data and/or with very long branches and unstable positions in the phylogenetic trees to constitute the data set 2 (DS2): this resulted in the removal of six samples (see [supplementary material 1, Supplementary Material](#) online—“Too many

missing data”), including one dried preserved sample. The DS2 included 110 samples and the % of missing data in the DS2 is provided in the [supplementary material 1, Supplementary Material](#) online, for each sample.

The DS2aML and DS2bML resulted in trees with most family-level groups monophyletic and highly supported, as well as most relationships between and within families (Bootstraps BS > 95) (supplementary material 2, Supplementary Material online). Two notable exceptions are the family Pseudomelatomidae, with three genera (*Antiplanes*, *Leucosyrinx*, and *Abyssocomitas*, thereafter referred to as the *Antiplanes* clade) clustering together as sister group to Drilliidae+Pseudomelatomidae in DS2b, and the family Borsoniidae, whose representatives do not cluster in a single clade. In particular, the families Conidae and Conorbidae are included in Borsoniidae, and one sample (*Heteroturrus kanacospira*) is more closely related to the clade that includes Mitromorphidae, Mangeliidae, Clathurellidae, and Raphitomidae (MMCR clade) than to the rest of borsoniids. The two analyses DS2aML and DS2bML were highly congruent with only a few unsupported differences (e.g., Pseudomelatomidae are sister group to the *Antiplanes* clade in DS2a and to Drilliidae in DS2b). Overall, the bootstraps are slightly higher with the DS2b, with, for example, the sister group relationships between Conoidea and the Mitridae-Pyramimitridae clade being more supported with the DS2b (100) than with the DS2a (49), and the Terebridae being sister group to the ADP (*Antiplanes* clade, Drilliidae, Pseudomelatomidae) clade, again more supported with the DS2b (94) than with DS2a (65).

Nineteen samples were removed from the data set 2 to create the data set 3. Among these 19, eight are outgroups and eight others were placed in families compatible with their morphological identifications. The last three correspond to dry material whose family attribution are here clarified for the first time: from our results, *Vitjazinella multicostata* is placed in Raphitomidae, *Ptychosyrinx chilensis* to the extended Turridae (results not shown based on COI sequences suggest that other *Ptychosyrinx* species also belong to Turridae, although not closely related to *P. chilensis*) and *Abyssocomitas kurilokamchatica* to the *Antiplanes* clade.

The results obtained with the DS3 are highly congruent with those obtained with the DS2, and the trees obtained with the different methods (ML, AS, and WSB) as well as with the two data sets (with 4 or 60 samples minimum per exon) were also mostly congruent (fig. 2 and table 2). When the AS (Astral) and WSB (Weighted Statistical Binning) analyses were congruent but contradicted the ML trees, the corresponding nodes were rarely supported (see below for an exception within Terebridae); in several cases, the ML and WSB trees are congruent and contradict the AS trees; there is no case where ML and AS trees are congruent and contradicts the WSB trees.

Some parts of the tree remain unsupported, while others were more supported with the DS3 than with the DS2. Except stated otherwise, the support for the described clades were 100 for bootstraps (BS), 1 for posterior probabilities (PP) with AS and 1 for posterior probabilities with WSB ("100/1/1"). Among outgroups, some relationships are highly supported, with, for example, Cancellariidae, Volutidae, and Volutomitridae constituting early-diverging lineages among neogastropods, a clade including Mitridae and

Pyramimitridae being sister group to the Conoidea (100/0.85/0.9), and a clade Melongenidae+Fascioliariidae being sister group to the clade Mitridae+Pyramimitridae+Conoidea (100/.73/.83).

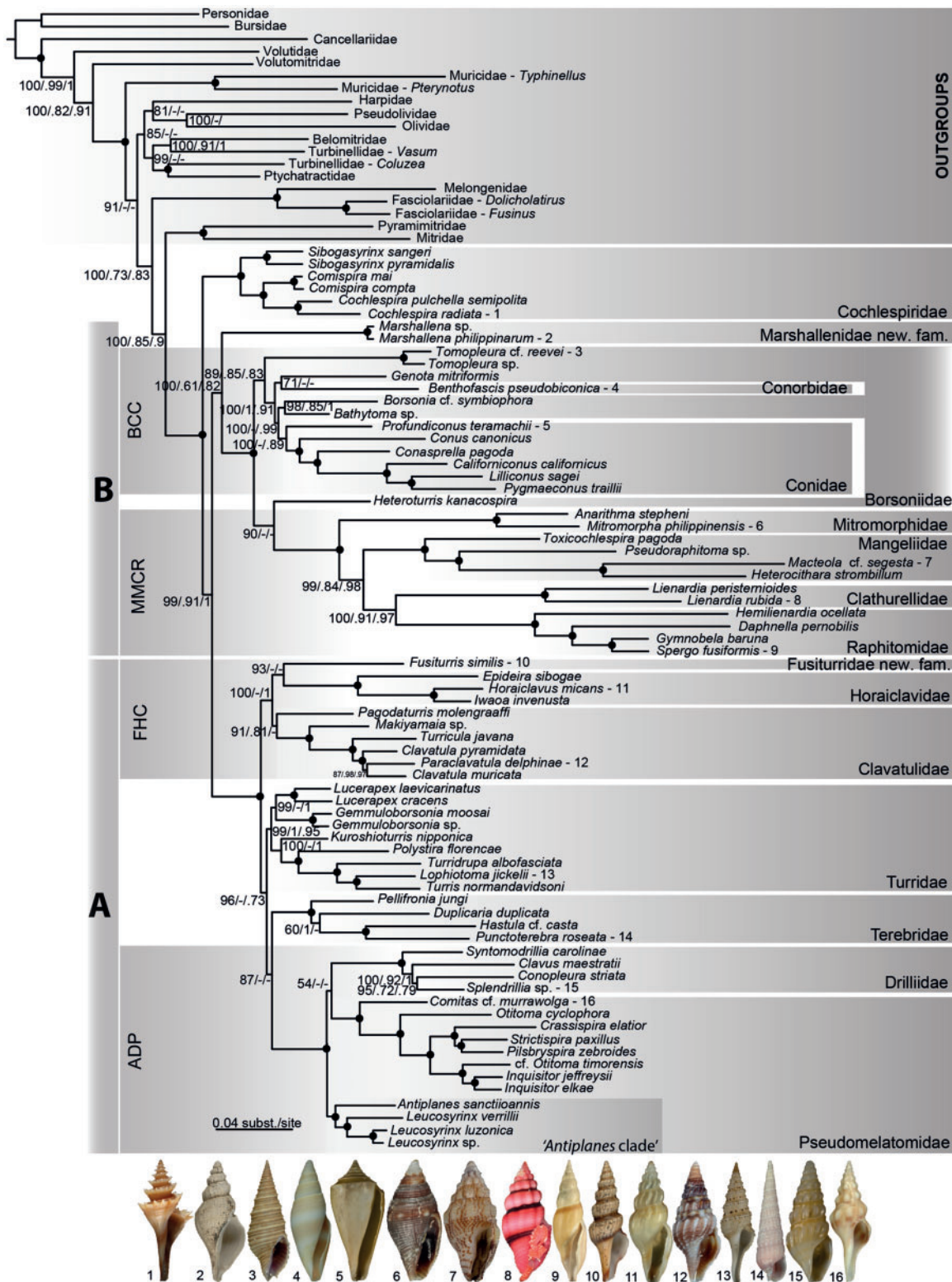
We found Conoidea to be monophyletic, with Cochlespiridae being sister group to the rest of the superfamily (99/.91/1), further divided into two main clades. The first one (100/0.61/0.82) includes the Marshallenidae new fam., sister group to a clade including on one hand the MMCR clade, and on the other hand a poorly supported clade (89/0.85/.83) including Borsoniidae, Conidae, and Conorbidae (BCC clade). The BCC clade constitutes the less resolved part of the Conoidea tree. *Heteroturrus* (Borsoniidae) is either sister group to the MMCR clade in the DS3bML analysis (BS = 90) or sister group to the BCC clade (or to the BCC clade except the *Tomopleura* lineage) in the DS3aML, AS, and WSB analysis (94/1/1). Irrespective of the position of *Heteroturrus*, Borsoniidae was always nonmonophyletic, including Conorbidae and Conidae. Conidae was monophyletic in the ML and WSB trees, but *Profundiconus teramachii* is more closely related to *Borsonia* and *Bathytoma* (Borsoniidae) in the AS trees (PP = 1).

In the last main clade of Conoidea (clade A), most nodes are resolved with ML, but the deepest relationships are variable and unsupported with the AS and WSB analyses. In the ML trees, a clade (100/–/1) including Horaiclavidae, Clavatulidae, and Fusiturridae new fam. is sister group to the rest. Then, we found Turridae to be the sister group (96/–/0.73) to a clade including Terebridae and the ADP clade. A few supported relationships are contradicting between the ML, the AS, and WSB analyses. *Fusiturris similis* is sister group to Horaiclavidae in the ML tree, but sister group to Clavatulidae (or even included in Clavatulidae, sister group to *Pagodaturris*) in the AS and WSB analyses (0.93 < PP < 1). Within families, some relationships are also variable, with, for example, *Kurishioturrus* more closely related to *Gemmuloborsonia* and *Lucerapex* in the AS analyses, while it is sister group to the rest of Turridae in the ML and WSB trees.

Dated Tree and Diversification Analyses

Except for the calibrated nodes, all the ages of the nodes in the dated phylogenetic trees are characterized by a very large confidence interval (fig. 3). The origin of the diversification of the Conoidea would be 138 Ma.

Only the diversification rate of Marshallenidae was consistently found to be significantly lower than expected, regardless of the number of species and the extinction rate considered (fig. 4). Conversely, the diversification rates of Conidae, Raphitomidae, Drilliidae, and Pseudomelatomidae were always significantly higher than expected. Diversification rates of other families were also higher than expected, but only when higher extinction rates were considered and/or only when the estimated total number of species was considered. When considering the described species, MEDUSA detected a decrease in the diversification rates for Cochlespiridae and Marshallenidae; when the total estimated number of species is considered, the diversification rates of



Downloaded from https://academic.oup.com/mbe/article/35/10/2355/5056056 by IFREMER user on 24 March 2022

Fig. 2. IQ-tree phylogenetic tree (DS3b). Fully supported nodes (BS = 100, PP = 1 for both AS and WSB) are in black; for the other nodes, BS and PP values (for both AS and WSB) are provided. Illustrated samples are numbered from 1 to 16.

Marshallenidae, Fusiturridae, and Conorbidae were detected as significantly decreasing. The relationships between the diversification rates and the venom gland was never significant, irrespective of the number of species and the extinction rate considered. Among the five pairs of clades for which one

clade included only species with venom glands, while the other clade included species with or without venom gland, one pair presented similar diversification rates (Drilliidae vs. Pseudomelatomidae), three pairs exhibited a higher diversification rate for the clade with or without venom gland

Table 2. Node Support Obtained with the Different Analyses and for Each Main Clade.

	DS2aML	DS2bML	DS3aML	DS3bML	DS3aAS	DS3bAS	DS3aWSB	DS3bWSB
CONOIDEA+MITRIDAE+PYRAMIMITRIDAE	49	100	100	100	0.89	0.85	0.57	0.9
CONOIDEA	100	100	100	100	1	1	1	1
COCHLESPIRIDAE	100	100	100	100	1	1	1	1
A+B	97	100	100	99	0.94	0.91	1	1
B	100	100	100	100	0.87	0.61	0.99	0.82
MARSHALLENIDAE	100	100	100	100	1	1	1	1
MMCR+BCC	100	100	100	100	1	1	1	1
MMCR	100	100	100	100	1	1	1	1
MITROMORPHIDAE	100	100	100	100	1	1	1	1
MANGELIIDAE	100	100	100	100	1	1	1	1
CLATHURELLIDAE	100	100	100	100	1	1	1	1
RAPHITOMIDAE	100	100	100	100	1	1	1	1
<i>Heteroturris</i> +MMCR	77	81	-	90	-	-	-	-
BCC	-	-	94	-	1	1	1	1
CONORBIDAE	100	100	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CONIDAE	100	100	100	100	-	-	0.87	0.89
<i>Profundiconus</i> + <i>Borsonia</i> + <i>Bathytoma</i>	-	-	-	-	1	1	-	-
A	100	100	100	100	1	1	1	1
FHC	100	100	100	100	0.72	-	1	1
HORAICLAVIDAE+FUSITURRIDAE	100	100	89	93	-	-	-	-
CLAVATULIDAE+FUSITURRIDAE	-	-	-	-	1	1	0.99	0.93
HORAICLAVIDAE	100	100	100	100	1	1	1	1
CLAVATULIDAE	37	100	81	91	0.79	0.81	-	-
ADP+TEREBRIDAE+TURRIDAE	65	98	54	96	-	-	0.7	0.73
TURRIDAE	100	100	100	99	1	1	1	0.95
ADP+TEREBRIDAE	65	94	56	87	-	-	-	-
TEREBRIDAE	100	100	100	100	1	1	1	1
ADP	100	100	100	100	1	1	1	1
DRILLIIDAE	100	100	100	100	1	1	1	1
PSEUDOMELATOMIDAE	100	100	90	-	-	-	-	-

"n.a.", not applicable; "-", node not found in the corresponding data set.

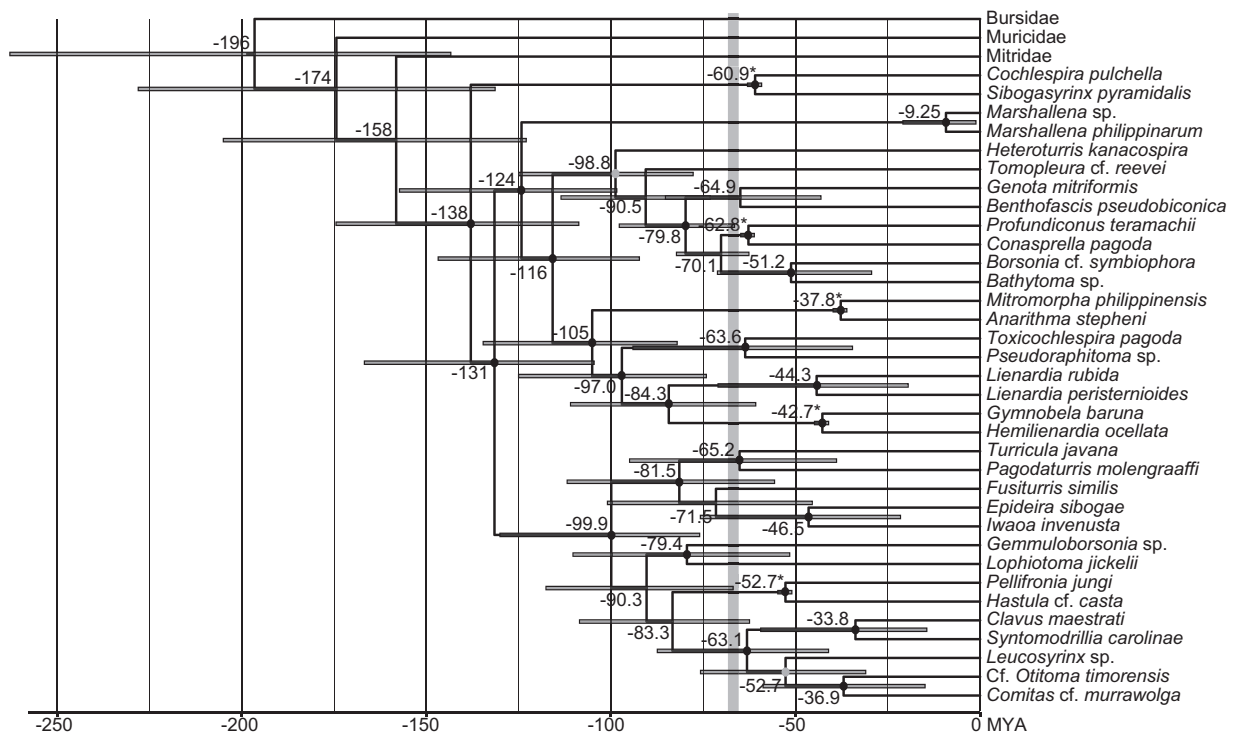


Fig. 3. Dated phylogenetic tree, with representatives of the main lineages as defined with the complete data set (fig. 2). Values on the nodes are the estimated ages; the horizontal bars are the confidence intervals. Black and gray dots represent fully (PP = 1) and highly (PP > 0.98) supported nodes, respectively. Gray vertical bar: C/T limit. *Nodes used for calibration.

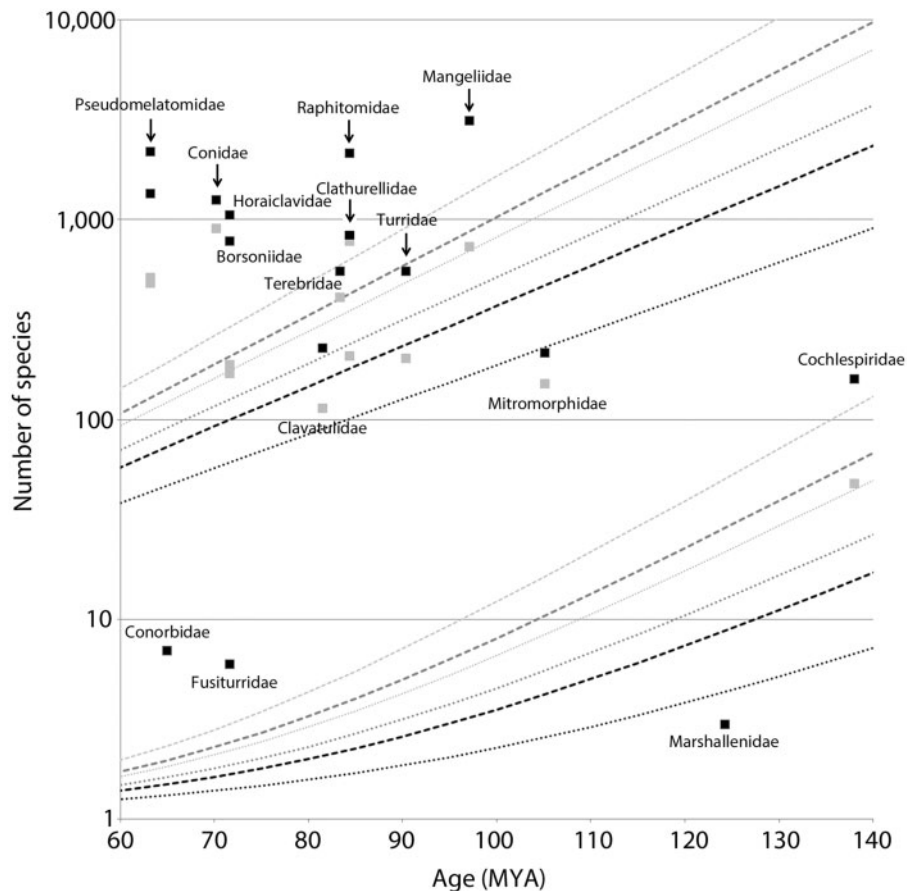


FIG. 4. Number of species relative to the age for each family as black and gray boxes (for the total estimated number of species and the number of described species, respectively), with confidence intervals at 95%: dotted lines for the number of described species, dashed lines for the total estimated number of species, light gray for an extinction rate = 0, dark gray for an extinction rate = 0.5 and black for an extinction rate = 0.9.

(Turridae vs. Terebridae + ADP clade, Clathurellidae vs. Raphitomidae, Horaiclavidae vs. Fusiturridae, or Fusiturridae+Clavatulidae) and one clade displayed either a higher or lower diversification rate for the clade with or without venom gland, depending on the sister-clade considered (Borsoniidae vs. Conidae or Conorbidae). Conversely, there is a significant correlation (P values < 0.05) between the diversification rates and the number of radula types in all cases (irrespective of the number of species and the extinction rate considered), that is, families with more radular types tend to have higher diversification rates (table 3).

Systematics

If most family concepts remain unchanged compared with the previous classification (Bouchet et al. 2011), some transfers of genera in other families are proposed: *Vitjazinella* in Raphitomidae, *Epideira* in Horaiclavidae, and *Gemmuloborsonia* in Turridae. Furthermore, some genera were only tentatively allocated to families in the last classification (Bouchet et al. 2011), either because their relationships remained unresolved or because they were absent from the corresponding molecular phylogeny (Puillandre et al. 2011). Thus, *Lucerapex* was only tentatively included in Turridae, while its position in the family is confidently confirmed

Table 3. Results of the PGLS Analysis, with the Correlation Coefficient and the P Value for Each Comparison.

Comparison	r^2	P Value
DR (described; 0)/VG	0.15	0.1
DR (described; 0.45)/VG	0.14	0.1
DR (described; 0.9)/VG	0.17	0.1
DR (total; 0)/VG	0.2	0.1
DR (total; 0.45)/VG	0.19	0.1
DR (total; 0.9)/VG	0.18	0.1
DR (described; 0)/Radula	0.4	0.008*
DR (described; 0.45)/Radula	0.4	0.008*
DR (described; 0.9)/Radula	0.41	0.007*
DR (total; 0)/Radula	0.36	0.014*
DR (total; 0.45)/Radula	0.36	0.015*
DR (total; 0.9)/Radula	0.35	0.016*

DR, diversification rate; "described", number of described species; "total", estimated total number of species; "VG", venom gland.

*Significant P values.

herein. Two genera, as shown here, correspond to independent lineages and/or are characterized by a very peculiar radula; on the basis of these results, they are described as new families: Fusiturridae new fam. and Marshallenidae new fam. The lineage including the genera *Antiplanes*, *Leucosyrinx*, and *Abyssocomitas* could have been considered a new family, but

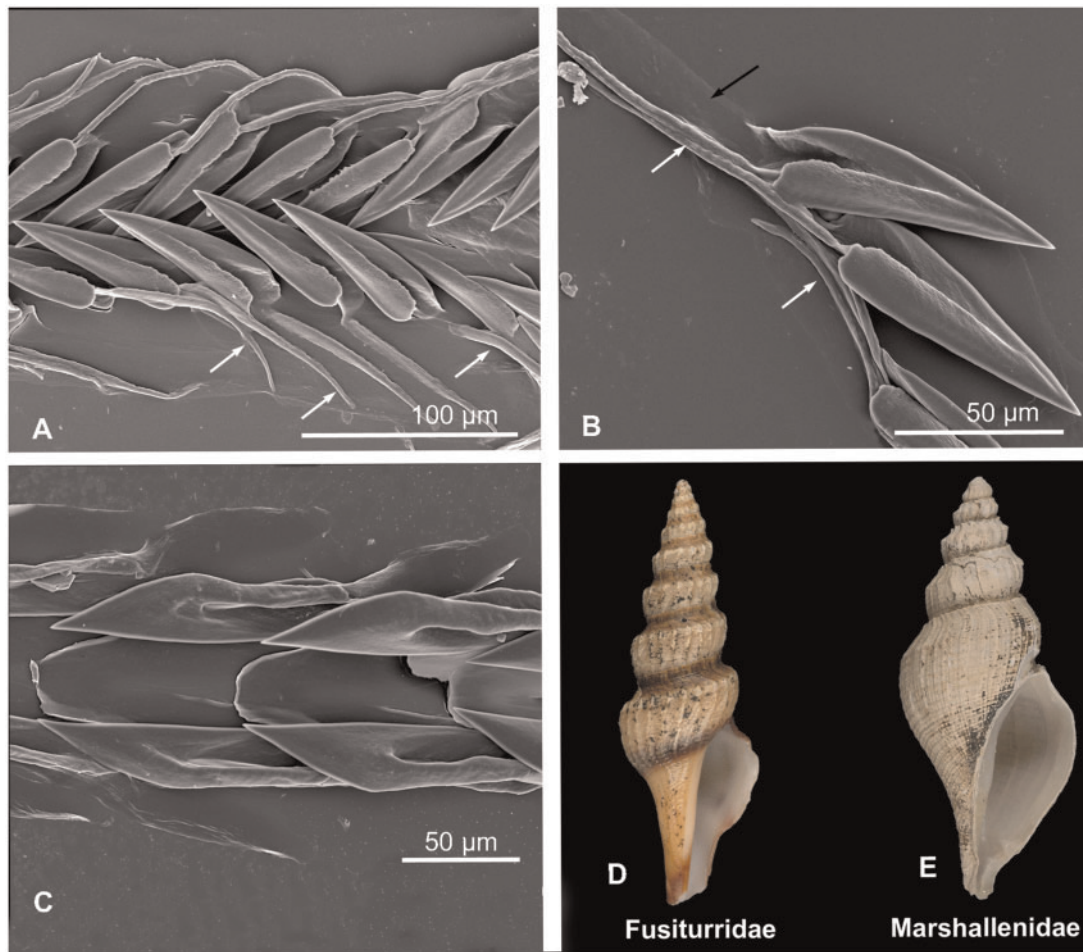


Fig. 5. Radulae and shells of new families of Conoidea. (A, B, and D)—Fusiturridae; (A and B)—radula of *Fusiturris similis* (Bivona Ant. in Bivona And., 1838), MNHN IM-2009-24971 (white arrows indicate posterior sclerotized edge of posterior part of the tooth, black arrow indicates unsclerotized anterior edge); (D)—shell of the same specimen. (C and E)—Marshallenidae, (C)—radula of *Marshallena philippinarum* (Watson, 1882), MNHN IM-2009-17099, (E)—shell of the same specimen.

their shell and radular morphology are not different from Pseudomelatomidae, and we thus decided to retain them in this family, pending on a better resolution of this part of the tree. We retained Borsoniidae as a valid family for the moment, because many other lineages must be included in the data set before eventually splitting it into multiple families. In the process of reconstruction of the tree and detailed examination of the analyzed taxa, we found several new taxa of specific and generic ranks. They were described in two separate papers (Kantor, Fedosov, et al. 2018; Kantor, Horro, et al. 2018): new genera *Comispira* (Cochlespiridae), *Pagodaturris*, and *Paraclavatula* (Clavatulidae), new species: *Comispira compta*, *Sibogasyrinx sangeri* (both Cochlespiridae), *Pagodaturris philippinensis* (Clavatulidae), *Horaiclavus micans*, *Iwaoa invenusta* (both Horaiclavidae), *Lucerapex cracens*, *Lucerapex laeovicarinatus* (Turridae), and *Heteroturris kanacospira* (Borsoniidae).

Marshallenidae new fam.

Type genus—*Marshallena* Finlay, 1926.

Diagnosis. Shell medium-sized, attaining 35 mm, fusiform-biconic, with medium high spire and attenuated medium

long siphonal canal (fig. 5E). Protoconch of 3–5 almost smooth whorls. Subsutural ramp poorly defined, slightly concave. Shoulder weakly angulated. Axial sculpture of distinct narrow axial ribs and thin raised growth lines, forming reticulated sculpture at intersections with thin but distinct spiral cords. Anal sinus weakly pronounced, subsutural. Operculum with terminal nucleus. Radula short, consisting of 12–13 transverse rows of teeth; marginal teeth duplex (fig. 5C), with lanceolate major limb and medium broad accessory limb, which fuses with dorsal side of major limb without a socket. Posterior (basal) half of marginal teeth not sclerotized, membranous, long, and as broad as major limb. Central formation weak, of fused plate-like acuspate lateral teeth without central cusp.

Remarks. Monotypic family, represented by fossil (Eocene to lower Pliocene of New Zealand) and a few Recent bathyal Indo-Pacific species. The genus was provisionally attributed to Horaiclavidae (Bouchet et al. 2011) on the basis of the radula. Nevertheless, *Marshallena* spp. have a central formation of fused lateral teeth, absent in Horaiclavidae. Another character, distinguishing *Marshallena* from studied Horaiclavidae and other Conoidea is the presence of a basal, long, and broad

poorly sclerotized membranaceous part of the marginal teeth.

Fusiturridae new fam.

Type genus—*Fusiturris* Thiele, 1929.

Diagnosis. Shell medium-sized, attaining 50 mm, narrowly fusiform, with high turriculate spire and attenuated long siphonal canal (fig. 5D). Subsutural ramp narrow, well defined, weakly concave. Shoulder obtusely angulated. Axial sculpture of distinct narrow to broad axial ribs. Spiral sculpture of thin cords. Anal sinus moderately deep, peripheral. Protoconch of three smooth whorls, conical. Operculum with terminal nucleus. Radular marginal teeth duplex (fig. 5A–B), differentiated in two parts. Anterior half normally sclerotized, of two subequal limbs, posterior slightly longer part poorly sclerotized, flat. Posterior edge of this poorly sclerotized part is continuous with the accessory limb, thickened, rod-like (marked with white arrows on fig. 5), while the anterior edge continuing major limb is unsclerotized and flat (marked with black arrow on fig. 5B). At junction of these two parts the tooth is narrowing, producing the “waist.” Central formation probably absent.

Remarks. Monotypic family, in shallow waters of Mediterranean and West Africa, fossil—Paleocene to Pleistocene of Europe. The genus was traditionally included in Turridae (Powell 1966) or later in Clavatulidae (Bouchet et al. 2011). The radula, illustrated for the first time by scanning electron micrographs, is very similar in two studied species—*F. similis* (Bivona Ant. in Bivona And., 1838), the sequenced species (fig. 5D), and *F. undatirrugata* (Bivona Ant. in Bivona And., 1838), the type species of the genus. It is unique so far for Conoidea in the presence of basal unsclerotized part of marginal teeth with one thickened rod-like edge.

Discussion

Capture Success

The success of the exon capture experiment was heterogeneous, with 21 samples for which <50% of the proteins were recovered, while >90% were recovered for 38 samples. Similarly, the number of bases, the depth of coverage, the % of reads on target, the number of exons recovered, the % of proteins recovered, and the % of missing data were all characterized by high SDs. One exception is the % of duplicates, which was more homogeneous among samples. These results certainly reflect the heterogeneity of the quality of the DNA preservation, with some samples collected only a few years ago while others were collected >15 years ago. The preservation of the samples (dry or in alcohol) clearly affected the parameters, with alcohol-preserved samples having significantly higher success, but these results are potentially biased by the low number of dried-preserved samples compared with the number of samples preserved in alcohol. Although it is difficult to tell apart the potential effects of the age of the samples and of the microwaves (the microwave oven has been used only for the most recent samples), using the microwave oven to remove tissues from the shell did not have a significant effect on the success of the exon-capture experiment, suggesting that this approach for handling specimens

and extracting tissue can be used as a way to improve efficiency during sample collection while having no impact on downstream applications. The samples collected in deep water (>100 m deep) were also less successful than the shallow-water samples. These results can be explained, for example, by the higher difference in environmental conditions between the living habitat of the samples collected deeper and the surface where they were processed, potentially resulting in samples that were already dead or moribund when they arrive on the deck of the boat, or by differences in tissue components due to adaptation to high pressure, which could interfere with DNA extraction. As expected (Bragg et al. 2016; Portik et al. 2016), samples in the ingroup were captured more effectively than the outgroups. Surprisingly, we were able to successfully recover between 60% and 76% of the proteins for the three members of the Tonnoidea, the supposedly less closely related outgroup to Conoidea even though the baits were not optimized for taxa outside of the Conoideans. Finally, the most significant differences in the parameters used to estimate the capture efficiency were found between the samples processed in the first and second batches, the second batch being less successful. Given that the exact same protocol was used for both batches, the lower quality results for the second batch was probably linked to the fact that most of the low concentration DNA extracts, including all the dry-preserved samples, were processed in the second batch. Thus, lower quality preservation may have led to limited DNA quantities to start our experiment, leading to suboptimal conditions. Moreover, most of the outgroups were also included in the second batch, which could explain a decreased efficiency in hybridization capture because of higher genetic divergence between the targeted DNA loci and the probes, as discussed before.

Despite this high heterogeneity among samples and categories of samples, we consider our capture experiment a success because we were able to enrich the targeted DNA in our samples (average % of proteins recovered = 75%). Our results also suggest that the samples preserved dry can be included in an exon-capture experiment, as well as samples that are phylogenetically distant from the samples used to design the exons (e.g., the Tonnoidea and the Conoidea would have diverged almost 200 Ma). It should also be noted that some of the specimens removed from the DS2 to constitute the DS3 were confidently placed in the phylogenetic tree obtained with the DS2, even if they had >98% of missing data.

Phylogenetic Reconstruction

As detailed in the results section, the different data sets and reconstruction methods used to obtain the phylogenetic trees provided congruent results (fig. 2 and table 2). Between the trees obtained with the ML, AS, and WSB methods, only a few nodes were incongruent (fig. 2). Overall, the ML trees were more resolved and supported, with more nodes in common with the WSB trees than with the AS trees. Concerning the level of missing data, the trees obtained with either 4 or 60 samples minimum per locus were almost identical, with only a few nodes slightly more supported with the

data set with 60 samples minimum per locus, in the ML and WSB trees. The differences in the phylogenetic trees were also limited between the data sets DS2 and DS3, the later (i.e., without the samples with >90% of missing data) resulting in a few better supported nodes (table 2).

The trees obtained using the exon-capture strategy are significantly more resolved and supported than the previously published trees of the Conoidea (Puillandre et al. 2011; Kantor et al. 2012; Uribe et al. 2018) (fig. 6). However, the sequencing approach used (Sanger sequencing of a few genes, mitogenomes, and exon-capture) is not the only difference between these data sets. The sampling is also different, with several lineages added in our study compared with the previous ones (and a few ones present in one of the previously published trees absent here—e.g., Bouchetispiridae, see above). As shown before (Nabhan and Sarkar 2012), the quality of the sampling can impact the inferred phylogenetic relationships. Our sampling strategy was the following: we sought to include in the data set not only all the known families of Conoidea, with several genera represented in each of them (when possible). We also included highly divergent lineages that were recognized as such in a COI tree (results not shown), but never recognized at the family, or even genus level. This strategy was made possible thanks to the active program of expeditions of the MNHN, Paris, that do not target specific taxa when going to the field but maximize the diversity of the samples collected, without any a priori selection in the samples collected. This strategy has been coupled with morphological identification and systematic sequencing of the barcode fragment of the COI gene. Therefore, we identified deeply divergent lineages that would have remained untouched if the sampling was focusing on the known taxa only. This sampling strategy also explains why the MNHN collections were cited in 62% of the mollusc species descriptions published in 2017, and that almost all the recently published molecular phylogenies of neogastropods are based, partly or entirely, on MNHN material (Modica et al. 2011; Fedosov et al. 2015; Couto et al. 2016; Galindo et al. 2016). In the case of the Conoidea, for which all the previously published phylogenies are also based mostly on MNHN material, this strategy led to the inclusion of several recently discovered lineages and to the description of several new families. Furthermore, the exon-capture strategy can take advantage of the vast array of alcohol-preserved and dried specimens housed in natural history collections. On the contrary, a transcriptome-based phylogeny needs samples specifically preserved in RNA-later or in dry ice, and a mitogenome-based phylogeny, needs higher quality DNA when the mitogenomes are obtained through long range PCR. Given the difficulty to resample rare lineages present in the museum collection (Bradley et al. 2014; Wen et al. 2015), but not preserved necessarily in good conditions for high quality DNA or RNA, the exon-capture strategy represents the best option to reconstruct old phylogenetic relationships.

Phylogenetic Relationships

With our exon capture phylogeny, we revealed new relationships among major groups of Conoidea and their relatives

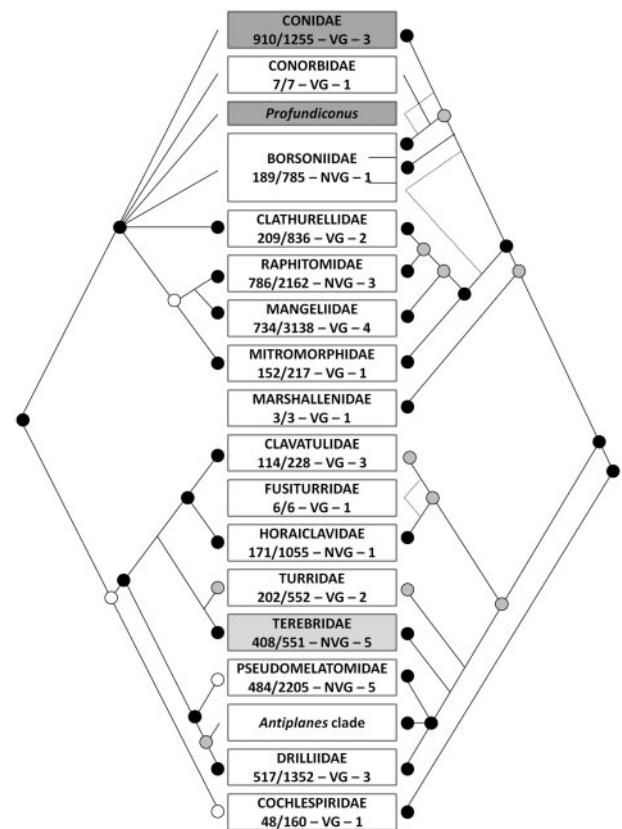


Fig. 6. Right: simplified tree based on mitochondrial genes (Puillandre et al. 2011) (Uribe et al. 2018); for graphical reasons, the supported relationships between Marshallenidae and Cochlespiridae found by Uribe et al. (2018) is not shown. Left: exon-capture-based tree, with alternate branches in dashed lines obtained either with IQ-tree or ASTRAL-II. Dark gray: Conidae, as traditionally conceived (« cone snails »); light gray: Terebridae (“auger snails”); white: “turrids.” For each family are provided the number of described species/the estimated total number of species, the character state for the venom gland (VG = all species has a VG; NVG = some species lost the VG) and the number of radula types. Black circles: fully supported nodes; gray circles: highly supported nodes (Bootstraps > 95 and PP > 0.95); white circles: moderately supported nodes (Bootstraps > 80 and PP > 0.90).

(fig. 6). In particular, Cochlespiridae is here recognized for the first time as the sister group of all other Conoidea, and not sister group of the other members of the clade A (Puillandre et al. 2011). This result makes the Turroidea defined by Tucker and Tenorio (2009) paraphyletic. Furthermore, Cochlespiridae being sister group to all the other Conoidea also explains their unusual for Conoidea venomous apparatus structure. In all members of Cochlespiridae, the venom gland opens into the oesophagus either within the nerve ring or even posterior to it (Medinskaya 1999; Simone 1999; Kantor, Fedosov, et al. 2018). In all the other conoideans, the venom gland passes through the nerve ring and opens in the buccal cavity just posterior to the opening of the radular sac, thus likely increasing the efficiency of the injection of the toxins in the prey (Taylor et al. 1993; Fedosov et al. 2017). The position of opening of the venom gland in Cochlespiridae is thus similar to the opening of the midgut gland (usually called

gland of Leiblein) into the oesophagus in other Neogastropoda. Moreover in Cochlespiridae, like in other neogastropods, the anterior oesophagus forms a more or less long loop, which is straightened during proboscis eversion. Thus, the structure of anterior foregut is rather similar in Cochlespiridae and non-Conoidean neogastropods, and probably represents the plesiomorphic condition for the Conoidea. The homology and evolution of the venom gland of Conoidea was discussed by Ponder (1970) and others (Taylor et al. 1993; Kantor 2003), and it was concluded that the midgut gland and venom gland are homologous structures. This particular opening of the venom gland was supposed to have emerged either twice independently (in the clade A except Cochlespiridae and in the clade B), or once in the ancestor of all the Conoidea but then secondarily lost in Cochlespiridae. The new phylogenetic pattern reduces the number of steps in the evolution of this character, with only one transition from the ancestral state (still found in Cochlespiridae) to the derived state (in all other Conoidea).

The conoidean anatomy is notably specialized due to the emergence of the unique feeding mechanism involving prey envenomation; this fact was always hampering inferences of the conoidean affinities based on the anatomical characters. A whole range of hypotheses as of which taxa within the Neogastropoda would be the most closely related to Conoidea have been proposed (Sheridan et al. 1973; Shimek and Kohn 1981; Taylor and Morris 1988; Ponder and Lindberg 1997). Although the sampling coverage for the non-Conoidean neogastropods is not complete, our analysis would suggest for the first time that the clade Mitridae-Pyramitridae (i.e., the superfamily Mitroidea) would be the sister group of the Conoidea.

The BCC clade probably constitutes the less resolved part of the tree, with Borsoniidae being paraphyletic in all trees (as suggested previously in Puillandre et al. 2011), including Conorbidae and Conidae. Borsoniidae includes species that are morphologically highly divergent, and as a consequence the limits of this group have always been fuzzy in the literature. Our results would tend to show that more taxa are needed to fully resolve this part of the tree. Pending more resolution here, we refrain to revise the classification of Borsoniidae and to eventually create new family names for each independent lineage in the BCC clade. Another striking feature of the BCC clade is the nonmonophyly of Conidae in some trees, with *Profundiconus* being more closely related to some Borsoniidae than to the other Conidae. Given the clearly coniform shells of the *Profundiconus* species, nobody doubted until now that they are Conidae. Our phylogenetic tree also contradicts previously published trees in a few cases (fig. 6). For example, the tree based on full mitogenomes recovered Marshallenidae and Cochlespiridae as sister groups, sister group to the rest of the members of the Clade A. As suggested by the authors (Uribe et al. 2018), this result could actually be a long-branch artefact, and our tree, separating Marshallenidae new fam. and Cochlespiridae, sister groups to the other members of the clade B and to all the other Conoidea, respectively, tend to support this hypothesis. Conversely, a rearrangement in the gene order of the

mitogenome in Cochlespiridae only (Uribe et al. 2018) better aligns with the hypothesis that Cochlespiridae are sister group to all the other Conoidea, than sister group to Marshallenidae new fam. only. The clade Clavatulidae+Horaiclavidae was also characterized by a rearrangement in the mitogenome: this would correspond to the FHC (Fustiturridae new fam., Horaiclavidae, Clavatulidae) clade in our tree, and we can thus predict that the same rearrangement should be found in *Pagodaturris*, a taxon that was absent in the mitogenome tree.

This new phylogeny leads to modifications in the genus and family-level classification of the Conoidea. In most cases, the families as defined in the previously published trees are confirmed, and the modifications concern lineages that were either absent in the previous trees, or those, whose position remained unresolved. In particular, here we establish new families to accommodate the genera *Fusiturris* (Fusiturridae new fam.) and *Marshallena* (Marshallenidae new fam.), previously tentatively placed in Clavatulidae and Horaiclavidae, respectively (Bouchet et al. 2011). The genus *Gemmuloborsonia*, tentatively placed in Clavatulidae by Bouchet et al. (2011) is here included in Turridae. The genus *Strictispira*, elevated at the family level (McLean 1971) because of its particular anatomical feature (the venom apparatus has been lost), is here for the first time shown to be only one of the many lineages of Pseudomelatomidae; the name Strictispiridae is thus considered a synonym of Pseudomelatomidae. Finally, we still consider the genera included in the *Antiplanes* clade as members of Pseudomelatomidae. However, if future analyses reveal that they are in fact sister group to Drilliidae or to the clade Drilliidae+Pseudomelatomidae, and not sister group to other pseudomelatomids, a new family will need to be proposed.

It is remarkable that in the last 25 years, the family-level classification of the Conoidea, and in particular the number of conoidean families, drastically changed, from 3 families until 1993, to 6 (Taylor et al. 1993), 15 (Bouchet et al. 2011), 16 (Kantor et al. 2012) and now 18. Apparently, this trend could be the equivalent of the taxonomic inflation observed at the species level, associated with phylogenetic approaches (Isaac et al. 2004) and sometimes blamed to be purely artefactual (Zachos et al. 2013). If higher taxonomic ranks, contrary to species, are generally considered to be purely artificial (Hedges et al. 2015; Giribet et al. 2016), the increase in the number of families in Conoidea is actually directly related to the phylogenetic relationships, and is not simply an arbitrary and taxonomist-dependent decision. Until 1993, the “turrids,” that is, all of Conoidea except Conidae and Terebridae, were placed in a single family, Turridae. The first phylogeny based on anatomical data (Taylor et al. 1993), then confirmed by DNA-based phylogenies (Puillandre et al. 2008), convincingly demonstrated that Turridae was polyphyletic. One decision could have been to synonymize the three families (Conidae, Terebridae, and Turridae), but the names Conidae and, in a lesser extent, Terebridae, were used in a large taxonomic and toxicological literature since several decades, and it has been decided to take into account the usage to maintain the Conidae and Terebridae taxa at the family

level. Consequently, the different lineages within the “turrids” needed to also be recognized at the family level and the same constraint guided following studies to recognize newly discovered lineages also at the family level, ultimately leading to this inflation in the number of families. Even if these 18 families do not necessarily reflect “naturally” existing groups, as species can be, this partition reflects the phylogenetic tree and the usage of names in the literature, and thus fulfills the objectives of the names: facilitating the communication on the objects they designate. Finally, some lineages included in previous phylogenies but not in the present one (*Cruziturracula*, *Fusiturracula*), could actually correspond to new families. Similarly, potential new lineages probably remain to be discovered and some of them may turn out to be new families: the current classification is probably not the end of the story.

Molecular Dating

The main limitation of our time-calibrated tree reconstructions is the low number of calibrations used (only five). This is the result of the lack of diagnostic shell characters for most of the Conoidea families, as several families of Conoidea are difficult to discriminate morphologically (Bouchet et al. 2011). A striking example is the genera *Leucosyrinx* (Pseudomelatomidae) and *Sibogasyrinx* (Cochlespiridae) pair, impossible to discriminate on conchological grounds (Kantor, Fedosov, et al. 2018), but having diverged >100 Ma. Consequently, to prevent incorrect placement of fossils for node calibration, we preferred to limit the number of calibrations to fossils that corresponded to the oldest known member of a morphologically diagnosable family, such as Conidae, Terebridae, Raphitomidae, Mitromorphidae, and Cochlespiridae. This likely came at the price of having very large confidence intervals for the noncalibrated nodes. It should be noted that other fossils of conoideans have been described from the Paleocene, and in particular members of Borsoniidae (Glibert 1973). Given that this family is not recovered monophyletic, we refrained to use borsoniid fossils to calibrate the tree.

However, even taking into considerations these large intervals, two major features regarding the timing of diversification of the Conoidea can be emphasized: the Conoidea would have emerged during the Cretaceous and most families were already present 50 Ma. These results are in agreement with the fossils record. It is generally accepted that “turrid” (i.e., conoideans that are neither a Conidae nor a Terebridae) fossils can be found in the Cretaceous and that representatives of most families are found in the Paleocene, where some were already diversified (Powell 1942; Powell 1966). Powell (1942) actually suggested that some fossils from the Upper Jurassic may belong to the Conoidea. However, attributing Cretaceous or even Paleocene or Eocene fossils in one of the Recent families is an arduous task, as discussed before, but the hypothesis that these early fossils are conoideans are strengthened by the divergence times estimated here.

Diversification Analyses

Although diversification patterns in Conoideans are thought to be driven by the origin of the venom apparatus leading to increased rates of speciation, our results indicate that diversification rates are highly variable, with some families exhibiting low diversification rates (Marshallidae, and to a lesser extent, Cochlespiridae, Fusiturridae, and Conorbidae), while others were diversifying faster than expected. Among them, the cone snails (Conidae) have long been recognized as one of the most diversified group of marine gastropods (Kohn 1990), but others, less studied, are revealed here for the first time, with even higher diversification rates (Drilliidae, Pseudomelatomidae). However, these results need to be taken cautiously, because the definition of the taxa considered for diversification rates potentially influences the comparisons (Stadler et al. 2014). For example, diversification rates calculated for a whole family may hinder variability among lineages within family. In particular, the cone snails include six genera, three of them represented by <10 species, and one (*Conus*) by 742 species (WORMS 2018). Within-family phylogenies are needed to test this hypothesis. Another potential bias is the use of stem ages (Stadler et al. 2014; Wiens et al. 2015). We choose to not use crown ages as they are likely underestimated because we included a very low number of species in the tree for each family. We believe that the sampling is less biased at the family level, and that the stem ages are overestimated in a much lesser extent.

Surprisingly, the loss of the venom gland does not seem to affect the diversification rates. Again, these results are only preliminary, and within-family phylogenies, with higher fractions of sampled species, are needed to confirm (or not) that the lineages that have secondarily lost the venom gland do not diversify at a lower rate. More expectedly, the number of radular types in a given family influences the diversification rates. One potential explanation is that the origin of new radular types may allow clades to access new preys, or develop a novel feeding mechanism, thus promoting specialization and speciation across different prey species. Again, within-family phylogenies would help to test this hypothesis, in particular in the families Terebridae or Pseudomelatomidae, for which the number of different radular types is the highest.

Conclusion

The use of the exon capture strategy to increase the quantity of molecular data available, and specifically targeting highly conserved markers, clearly represents an improvement in the resolution of the Conoidea phylogeny. The tree is better resolved and supported, several samples conserved dried have been successfully included in the data set, and the cost per sequenced nucleotides is much lower compared with past approaches (rough estimations: 0.01, 0.0017, and 0.00045€ per nucleotide for three genes, full mitogenomes and exon-capture sequencing, respectively). However, a few parts of the tree remain unresolved, even questioning the monophyly of the best known family of the Conoidea, the cone snails. This result, among others, definitely demonstrates that conoidean shells are a poor proxy for systematics, and that any

hypothesis on the evolution of the group should be proposed and tested based on a robust phylogenetic framework. In this context, we propose for the first time that the secondarily loss of the venom gland in some families does not impact diversification rates, contrary to the number of radula types; these hypotheses remain to be tested with more complete phylogenetic trees, that is, including many more species. Producing even more genetic data to clarify the few remaining gray areas in the Conoidea tree, using, for example, more transcriptomic data to identify more conserved exons, or even using transcriptomic data for all the samples, is thus only part of the problem. Even at the family level, it is likely that conoidean lineages are undersampled, as demonstrated by the “exploratory” strategy of the MNHN to sample the Conoidea diversity, that revealed many new genera and families in the last few years. Pursuing the sampling effort, especially in the deep-sea of the Pacific, where most recently described taxa were discovered, will most probably lead to the discovery of new deep lineages in Conoidea. And no doubt either that these new lineages will impact, if not the phylogenetic relationships among the known lineages, at least the diversification patterns that can be inferred from the trees. Exploration of the marine realm, even if not facilitated by the increasing administrative burden associated to permit applications, should thus remain a priority; not doing so would be at the cost of not accurately understanding the evolutionary processes of diversification in the sea.

Materials and Methods

Sampling

Specimens were selected to represent different genera in as many families of Conoidea as possible. Bouchetispiridae is the only currently recognized family missing from our data set. It includes a single species, represented in our collection (and probably worldwide) by only two alcohol-preserved specimens. We were not able to recover enough DNA for the exon-capture library preparation after several attempts. We also selected species that represented highly divergent lineages revealed through COI sanger sequencing (results not shown), potentially corresponding to new family-level taxa.

The outgroups were selected to include representatives of many Neogastropod families. Three representatives of the Tonnoidea superfamily were also used as distant outgroups, being generally considered as one of the most closely related non-Neogastropoda (or even as a Neogastropoda, as revealed in recent phylogenies based on complete mitochondrial genomes—Osca et al. 2015). We included several outgroup taxa to execute the following: 1) to test the monophyly of the Conoidea and identify its potential sister taxon, and 2) to explore the capacity of the exon-capture strategy to recover loci from nontargeted species, as we designed baits using only conoidean transcriptomes (see below).

Most specimens were collected during MNHN expeditions (expeditions.mnhn.fr), except a few ones (see [supplementary material 1, Supplementary Material](#) online). Specimens were either anesthetized in a magnesium chloride solution or microwaved (Galindo et al. 2014) right after sampling to

remove the body from the shell; tissue was then kept in 95° alcohol. Nine specimens were kept dry after sampling ([supplementary material 1, Supplementary Material](#) online). Most vouchers are in the MNHN collections ([supplementary material 1, Supplementary Material](#) online).

Exon Design

We used an exon capture approach to recover genetic markers for phylogenetic inference. For bait design, we chose to generate ancestral sequences, an approach previously used to improve locus recovery across divergent taxa (Hugall et al. 2016) by reducing the genetic distance between the bait sequence and the sample of interest. Generating ancestral sequences was necessary given the expansive phylogenetic breadth of the taxa, we attempted to sequence in this study. At the time, we designed baits for this study, there were only three publicly available non-Conidae conoidean transcriptomes: *Unedogemmula bisaya*, *Gemmula speciosa*, and *Crassispira cerithina* (NCBI Sequence Read Archive [SRA] accession no.'s SRR1574923, SRR1574907, and SRR1574922; Gonzales and Saloma 2014). We chose to pair these species with sequences from a recent exon capture study on Conidae (Phuong and Mahardika 2018), which would allow us to generate ancestral sequences that would represent the ancestor to all Conoidean species. We attempted to target the same loci as in Phuong and Mahardika (2018), where they targeted sequences representing 886 protein coding genes. 404 of the genes were identified through a reciprocal blast approach (Phuong and Mahardika 2018), while the remaining 482 were identified in Teasdale et al. (2016) to be orthologous in Pulmonate gastropods.

To design the bait sequences, we first downloaded the non-Conidae conoidean transcriptomes from the National Center for Biotechnology Information Sequence Read Archive (Leinonen et al. 2011). We used Trimmomatic v.0.32 (Bolger et al. 2014) to remove adapter contamination and filter low quality reads (ILLUMINACLIP option enabled, seed mismatch threshold = 2, palindrome clip threshold = 40, simple clip threshold of 15; SLIDING WINDOW option enabled, window size = 4, quality threshold = 20; MINLEN = 36; LEADING = 3; TRAILING = 3) and assembled transcripts using the version of trinity released on April 13, 2014 (Grabherr et al. 2011) with a minimum contig length of 151 bp. We used blastn v2.2.3 (evalue = 1e-10, word size = 11) to associate the assembled transcripts with sequences from the Conidae baits (Phuong and Mahardika 2018), retaining only sequences with a percent identity > 80%. We aligned sequences per locus using muscle v3.6 (Edgar 2004). We used FastML v3.1 (Ashkenazy et al. 2012) to generate ancestral sequences under a JTT model of sequence evolution using the following pairs of sequences: 1) a Conidae sequence + a *C. cerithina* sequence and 2) a Conidae sequence and a *G. speciosa* or a *U. bisaya* sequence. When both *G. speciosa* or *U. bisaya* were available, we merged both sequences prior to ancestral sequence reconstruction because they are closely related. We chose these two groupings to increase the diversity of bait sequences in our design, as *C. cerithina* is part of the Pseudomelatomidae family and *G. speciosa* and *U. bisaya*

is part of the Turridae family. We were also provided sequences for an additional species, *Mitra badia* (Mitridae, not included in Conoidea), from A. Moussalli and F. Köehler for loci present in the Pulmonate loci set (Teasdale et al. 2016). We generated ancestral sequences using the same methods as described earlier to generate ancestral sequences with *M. badia* and a Conidae sequence.

We defined exon/intron boundaries with EXONERATE v2.2.0 (Slater and Birney 2005) under the est2genome model with the *Lottia gigantea* genome (Simakov et al. 2013) as our reference. We retained all exons that were at least 50 bp and created chimeric baits by merging exons when they were <120 bp (our desired bait length). We performed a self blast using blastn v2.2.3 (evalue = 1e-10) to ensure that the bait sequences from separate proteins did not have percent identity > 80%. We removed loci with GC content < 30% or > 70% because extreme GC content has been shown to decrease hybridization efficiency (Bi et al. 2012) and we used RepeatMasker (Smit et al. 1996) to filter repetitive elements from our bait sequences. We sent these sequences to Mycarray (Ann Arbor, MI) to synthesize a custom MYbaits-1 kit, which allowed for 20,000 bait sequences. The bait length was 120 bp with 2.4X tiling. In summary, our bait design targeted ~678,322 bp, representing 850 protein coding genes.

Library Preparation

DNA was extracted with the NucleoSpin 96 Tissue kit from Macherey-Nagel using the Epmotion 5075 robot (Eppendorf) following the manufacturers' recommendations. The 120 samples were processed in two batches (one first batch of 50 samples followed by a second batch of 70 samples) for three reasons: 1) making our first attempt at this protocol in favorable conditions in the first batch by including whenever possible the best quality samples we had, leaving the lesser condition ones for the second batch; 2) adding in the second round key representatives that did not work in the first round; and 3) adding in the second round more representatives of lineages that were potentially nonmonophyletic, or which position was unexpected, in preliminary phylogenetic analyses performed with the 50 samples of the first round (e.g., Cochlespiridae, Borsoniidae), or, conversely, limiting the number of representatives in families that were already well recovered in the first round.

Library preparation followed Meyer and Kircher (2010) with some modifications. For the first batch, 1 µg of DNA per sample was used as starting material. Due to overall lower samples quality, starting material for the second batch was between 30 ng and 1 µg. Limited availability of specimen tissues made it impossible to get better DNA extracts for several key samples (e.g., Bouchetispiridae, Conorbidae). All purification steps were conducted using homemade SPRI beads (Rohland and Reich 2012; Faircloth and Glenn 2014). DNA was sheared through sonication using a Bioruptor Pico with three cycles of 7 min (30 s ON/30 s OFF). Then, sheared DNA was blunt-end repaired prior to adapter ligation and fill-in. In order to verify the ligation success, an amplified PCR product showing a discrete band of 300 bp also underwent these

steps. A difference in band size after migration in an agarose gel corresponding to the length of the adapters should be observed after ligation. Resulting libraries were quantified and qualified through qPCR and fluorometry (Qubit). Depending on the library concentrations, 5–15 cycles of indexing PCR were conducted. After quantification of the indexing PCR, 120 ng of each library were pooled by groups of 10 samples.

Hybridization Capture

Capture was conducted following MyBaits protocol v3.0 with a few modifications. As recommended, between 100 and 500 ng of each pool was used for the capture. Instead of the blocking oligos provided, xGen Blocking Oligos from Integrated DNA Technologie were used. Capture was conducted for 40 h at 60°C on a BIO-RAD C1000 touch thermal cycler. Postcapture libraries were cleaned-up following MyBaits protocol and quantified using Qubit. Each library was amplified through three independent PCR reactions of 12 cycles. The three PCR products were then pooled in order to increase potential libraries complexity. At this stage, capture success was assessed using two positive and two negative controls. Those controls were genes that could be amplified through PCR and that were supposed to be captured by our baits (positive controls) or not (negative controls). Controls were amplified using specific primers for each library using qPCR prior and after capture. Global success in capture was thus assessed by observation of a gain after capture in the number of cycles needed for the PCR to reach the threshold cycle (Ct or Cq) in the case of the positive controls and a delay or an absence of amplification in the case of the negative controls. Finally, each library of ten samples was quantified again and characterized using an Agilent 2100 Bioanalyzer. The first batch was sequenced on two lanes of Illumina HiSeq 2000 paired-end (100 bp reads). More than 230 M reads were produced. The second batch was sequenced on one lane of Illumina HiSeq 4000 paired-end (100 bp reads) and produced >400 M reads.

Exon Assembly

We trimmed reads for adapter contamination and quality using Trimmomatic v0.36 (ILLUMINACLIP option enabled, seed mismatch threshold = 2, palindrome clip threshold = 40, simple clip threshold of 15; SLIDING WINDOW option enabled, window size = 4, quality threshold = 20; MINLEN = 36; LEADING = 15; TRAILING = 15) and used flash v1.2.11 (Magoč and Salzberg 2011) to merge reads. We generated assemblies using SPAdes v3.8.1 (Bankevich et al. 2012) and used cap3 (Huang and Madan 1999) and cd-hit v4.6.5 (percent identity = 99%) to reduce redundancy in the assemblies. We used blastn v2.2.31 (evalue = 1e-10, word size = 11) to associate contigs with the targeted loci and used EXONERATE vXX under the est2genome model to redefine our target sequences because many of the original predicted exons were actually composed of several smaller exons. To fix misassemblies and estimate average heterozygosity, we mapped reads using bowtie2 v2.2.7 (Langmead and Salzberg 2012) using the very sensitive local and no discordant options, marked duplicates using picard-tools

v2.1.1 (<http://broadinstitute.github.io/picard>; last accessed July 29, 2018), and called single nucleotide polymorphisms (SNPs) using samtools v1.3 and bcftools v1.3 (Li et al. 2009). For each sample, we applied the following filters: 1) we removed sequences if estimated heterozygosity was >2 SDs away from the mean, 2) we removed sequences if they blasted to multiple reference targets, 3) we masked positions $<4\times$ coverage, and 4) removed sequences if there was not a minimum of $4\times$ coverage across 70% of a particular sequence. To assess the capture experiment, we generated the following statistics for each sample: the average depth of coverage, the % of reads mapping to targeted loci, the % of duplication, and the number of exons and protein-coding genes. Potential differences between different categories of samples were assessed with a Student's *t*-test: batch 1 versus batch 2, samples preserved dry versus preserved in alcohol, shallow water (i.e., >100 m deep) versus deep water samples, microwaved versus nonmicrowaved samples, and ingroup versus outgroup samples.

Phylogenetic Analyses

We generated alignments using mafft v7.222 (Katoh et al. 2005). To test the potential effect of missing data, several data sets were constructed. First, RAxML (= "ML" for the data sets 1 and 2—see below) (Stamatakis 2006) analyses were performed on a concatenated alignment of all the loci, including all the samples (data set 1 = "DS1") with a GTRGAMMA model (unpartitioned) and 100 bootstrap replicates. This initial analysis was performed to detect contaminations and/or obvious misplacements. After removal of these problematic samples, a second RAxML analysis was performed, with the same parameters. At this step, we also tested the effect of the proportion of missing data per locus, by analyzing two data sets: one including all the loci shared by at least four samples (the minimum number of taxa per locus needed for the ASTRAL-II analyses—see below) (data set 2a = "DS2a"), and the other including all the exons shared by at least 60 samples, which represent half of the samples included in the initial data set (data set 2b = "DS2b").

To further decrease the level of missing data, we also built an additional data set, keeping only the samples with $<90\%$ missing data in the DS2a data set. This 90% threshold was chosen to ensure that all the main lineages (roughly family-level clades) are represented in this data set (a threshold at 80% would have removed from this data set the two only representatives of Conorbidae). Again, two thresholds were used for the minimum number of samples per locus: 4 (data set 3a = "DS3a") 60 (data set 3b = "DS3b").

The data sets 3a and 3b were analyzed with IQ-tree (= "ML" for the data set 3) (Nguyen et al. 2015). We estimated the best substitution model for each partition (locus) in each concatenated data set with ModelFinder (Kalyaanamoorthy et al. 2017) following the BIC criterion. We then applied 1,000 ultrafast bootstrap (UFBoot) (Hoang et al. 2017) on each data set to obtain branch support. We also applied a supertree approach implemented in the program ASTRAL-II (= "AS"—Mirarab and Warnow 2015). We generated an individual tree per locus with IQ-tree, using the associated best substitution

model. The Weighted Statistical Binning ("WSB"—Bayzid et al. 2015) method was also used to combine the exons in bins, that is, groups of similarly evolving exons, and produce bin-trees that were then combined in a single supertree with ASTRAL-II.

In the remaining text, these analyses will be referred as follow: DS1ML, DS2aML, DS2bML, DS3aML, DS3bML, DS3aAS, DS3bAS, DS3aWSB, DS3bWSB.

Dated Tree

To calibrate the tree, we assigned fossils to Conoidean families based on morphological synapomorphies. The fossils of Terebridae and Conidae were identified on the basis of the unique shape of their shell (elongated and cone-shape shells, respectively). Kohn (1990) and Duda and Kohn (2005) fixed at 55 Ma, the age of the oldest fossil of cone snails. However, a recent revision of the group moved this age to 58 Ma (Tracey et al. 2017) and we used this date to calibrate the Conidae node. The oldest Raphitomidae was identified by its cancelled protoconch (Lozouet 2017), a feature found only in Raphitomidae among conoideans. Similarly, the particular shell shape of *Cochlespira*, with spines extending from the main cord, and of Mitromorphidae, with a small mitriform shell, were also used to attribute fossils to these lineages. The fossils retained to calibrate the tree are listed in table 4.

To reduce computation time, only the 10% of most complete loci (corresponding to 31% of missing data for the less shared locus, and to 11% of missing data in the final data set) were kept from the DS3. This reduced data set included 138 loci, and we used PartitionFinder 2.1.1 (Stamatakis 2014; Lanfear et al. 2016) and the rcluster algorithm (Lanfear et al. 2014) to identify the best substitution model for each locus and to group them when the inferred substitution models were similar, resulting in 10 groups of loci. Furthermore, only three outgroups (chosen among the samples with the lowest amount of missing data, and among closely and distantly related outgroups—Bursidae, Muricidae, and Mitridae) and only two (when possible) samples per Conoidea family (except for the paraphyletic Borsoniidae family and the revised Turridae family—see Results) were retained for the dating analyses (supplementary material 1, Supplementary Material online). MrBayes 3.2.6 (Huelsenbeck et al. 2001) was used to reconstruct a dated tree, with two parallel analyses, each consisting of four Markov chains of 100,000,000 generations with a sampling frequency of one tree every 10,000 generations. We set the number of swaps to three and the chain temperature at 0.02. The branch lengths were set to follow a birthdeath clock, and the clock as an Independent Gamma Rate (clockvarpr=igr) with an IGR parameter (igrvarpr) set to $\exp(10)$. The prior assumptions concerning the base substitution rate of the tree (clockratepr) were set to follow a lognormal distribution, with a mean of -7.1 and a SD of 2.4. The age of the calibrated nodes followed a lognormal distribution (see table 4 for the minimum age, mean age, and SD values applied). Convergence of the analysis was evaluated using Tracer 1.4.1 (Rambaut and Drummond 2014), to ensure that all ESS values were >200 . A consensus tree was then calculated after omitting the first 25% trees as burn-in.

Table 4. Fossils and Parameters used to Calibrate the Tree.

Dated Node	Fossil	Epoch/Stage	Age (MY)	Min. Age	Mean Age	SD	Référence
Terebridae	<i>Mirula plicatula</i> (Lamarck, 1803)	Lower Eocene	56.0–47.8	47.8	52.8	1	Lamarck (1803)
Conidae	<i>Hemiconus leroyi</i> Tracey et al. 2017	Paleocene	58	58	63	1	Tracey et al. (2017)
Raphitomidae	<i>Pleurotomella polycolpa</i> (Cossmann, 1889)	Middle Eocene	47.8–37.8	37.8	42.8	1	Gougerot and Le Renard (1981) and Lozouet (2017)
Mitromorphidae	<i>Mitrolumna bartoniana</i> Boussac, 1911	Upper Eocene	37.8–33.9	33.9	37.9	1	Boussac (1911)
Cochlespiridae	<i>Pseudocochlespira rosenkrantzi</i> Schnetler, 2001	Paleocene	66–56	56	61	1	Schnetler (2001)

Diversification Analyses

Dated trees can be used to infer diversification rates and to propose and test hypotheses related to the processes at the origin of the diversity. However, although some methods may be robust to incomplete sampling of up to 50% of included species in the tree (Rabosky 2014; Wiens et al. 2015), our Conoidea data set only includes c.a. 2% of the described species (and 0.7% of the estimated total number of species—see below). For this reason, we followed the methodology applied by Wiens et al. (2015). First, we used the method-of-moments estimator for stem-group ages (Magallon and Sanderson 2001), implemented in the R package LASER (Rabosky 2006). As extinction rates are unknown for the Conoidea, three extinction rates ($\epsilon = 0, 0.5$ and 0.9 —Magallon and Sanderson 2001) were tested. Because the number of species considered in each clade can impact diversification analyses (Faurby et al. 2016), two species numbers were considered: the number of described species, as provided in WORMS (2018), and a putative total number of species, estimated from the ratio of described/undescribed putative species determined using a data set of ca. 6,000 COI sequences of Conoidea (results not shown) (fig. 5). The diversification rates were then calculated using the stem ages from the dated tree for each family of Conoidea. Because Borsoniidae is not monophyletic, only the lineage that potentially includes most of the borsoniids (i.e., all except *Heteroturris*, *Genota*, and *Tomopleura*) was considered. For the same reason, the *Antiplanes* clade (see Results) was ignored when analyzing the family Pseudomelatomidae. 95% Confidence intervals for the diversification rates were calculated for each data set (described and total number of species, for the three extinction rates). MEDUSA (Alfaro et al. 2009), with default parameters, was also used to identify potential shifts in diversification rates among clades. A backbone tree of Conoidea at the family level was reconstructed by manually deleting all but one branch in each family, and both the number of described species and estimated total number of species were used to estimate diversification rates.

A potential correlation between the venom apparatus, hypothesized to be the key-innovation at the origin of Conoidea, and the diversification rates was assessed using PGLS (Martins and Hansen 1997) as implemented in the R package Caper (Orme 2013). We estimated the value of lambda and used a fixed value of 1 for kappa and delta for these analyses. The same backbone tree created for MEDUSA

was used here. Six series of diversification rates (obtained with three extinction rates and for both the described and total number of species) were considered. It was not possible to test for a relationships between diversification rates and the toxin diversity in each family, because toxins are only well characterized for the cone snails (and in a much lesser extent for three other families—Terebridae, Turridae, and Pseudomelatomidae). We thus focused on the venom apparatus in itself. First, the venom gland has been lost secondarily in several lineages (Terebridae—Miller 1970; Miller 1971; Raphitomidae—Sheridan et al. 1973; Kantor and Sysoev 1989; Kantor and Sysoev 1996; Kantor and Taylor 2002; Fedosov 2007; Pseudomelatomidae—Kantor and Taylor 1994; Horaiclavidae—Fedosov and Kantor 2007; and Borsoniidae—Medinskaya and Sysoev 2003). However, not all the species in each of these family have lost the venom gland, so it was not possible to simply code the venom gland as present or absent. We thus coded the venom gland as either present in all species of the family, or variable (present or absent) in the family (fig. 5), to test whether if the secondary loss of the venom gland has an effect) on the diversification rates. Our hypothesis here is that clades in which all species have venom glands should exhibit higher diversification rates, given that the venom gland is thought to spur evolutionary diversification due to prey specialization (Olivera 2006). Similarly, several radular types can be recognized among the conoidean families, and some families actually developed several types of radula. Similar to the venom gland, a given type of radula is not always fixed in a given family, and we thus did not code the various types of radula but simply the number of radula types found in each family (fig. 5) to test whether families with higher number of radula types diversified more, less, or at the same rate, as the families where all the species have the same radula type. Here, we hypothesize that families with higher number of radula types would have a wider spectrum of preys, and would thus be able to speciate more, showing higher diversification rates.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

The material in this article originates from numerous shore-based expeditions and deep sea cruises, conducted,

respectively, by MNHN and Pro-Natura International (PNI) as part of the Our Planet Reviewed programme (ATIMO VATAE, MAINBAZA, INHACA 2011, GUYANE, PAPUA NIUGINI, KAVIENG 2014) and/or by MNHN and Institut de Recherche pour le Développement (IRD) as part of the Tropical Deep-Sea Benthos programme (AURORA 2007, TAIWAN 2013, DongSha 2014, NANHAI 2014, BIOPAPUA, SALOMONBOA 3, CONCALIS, EXBODI, NORFOLK 2, TERRASSES). Scientific partners included the University of Papua New Guinea (UPNG); National Fisheries College, Kavieng; Institut d'Halieutique et Sciences Marines (IH.SM), Université de Tuléar, Madagascar; Universidade Eduardo Mondlane, Maputo; the Madagascar bureau of the Wildlife Conservation Society (WCS); and Instituto Español de Oceanografía (IOE). Funders and sponsors included the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, Richard Lounsbery Foundation, Vinci Entrepouse Contracting, Fondation EDF, European Regional Development Fund (ERDF), the Philippines Bureau of Fisheries and Aquatic Research (BFAR), the French Ministry of Foreign Affairs, Fonds Pacifique and the Government of New Caledonia. Additional field work included PANGLAO 2004 (a joint project of MNHN and University of San Carlos, Cebu City); KARUBENTHOS 2012 (a joint project of MNHN with Parc National de la Guadeloupe and Université des Antilles); sampling in Western Australia arranged by Hugh Morrison, with support of the Western Australian Museum; and sampling in Congo arranged by Bernard Thomassin. Collection of material in Vietnam was supported by the Russian–Vietnamese Tropical Center. The Office of Naval Research funded the ONR expedition onboard R/V Vidal Gormáz from Chilean Navy. Angelika Brandt was the head of expedition Kurambio. We are thankful to the staff of the Tropical Center for assistance in organization of the field sampling and loan of some laboratory equipment. The CAML-CEAMARC cruises of RSV Aurora Australis and TRV Umitaka Maru (IPY project no.53) were supported by the Australian Antarctic Division, the Japanese Science Foundation, the French polar institute IPEV (ICOTA and REVOLTA programmes), the CNRS, the MNHN and the ANR (White Project ANTFLocks USAR no.07-BLAN-0213-01 directed by Guillaume Lecointre). All expeditions operated under the regulations then in force in the countries in question and satisfy the conditions set by the Nagoya Protocol for access to genetic resources. The authors would like to thank Philippe Maestrati, Virginie Héros, Barbara Buge, and Julien Brisset for their help in curating the vouchers; Javier Sellanes, Enrico Schwabe, Serge Gofas, Juan Horro, Peter Stahlschmidt, and Rafael Zardoya for their help in collecting some samples, Adnan Moussalli, and Frank Köehler for having provided the Mitridae transcriptome, and Gavin Malcolm, Yves Terryn, Manuel Tenorio, and Eric Monnier, among others, for their help in identifying the samples. This work was supported by the Service de Systématique Moléculaire (UMS 2700 CNRS-MNHN), the CONOTAX project funded by the French National Research Agency (grant number ANR-13-JSV7-0013-01), the bilateral cooperation research funding from

the Ministry of Science and Technology, Taiwan (grant number MOST 102-2923-B-002-001-MY3) and the French National Research Agency (grant number ANR 12-ISV7-0005-01) and the Russian Science Foundation (grant number 16-14-10118, PI Yu. Kantor).

Data Accessibility

The data set 2 has been uploaded in FigShare: <https://figshare.com/s/827bbfa2bb5461889912>; last accessed July 29, 2018.

References

- Alfaro ME, Santini F, Brock CD, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci U S A*. 106(32):13410–13414.
- Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T. 2012. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res*. 40(W1):W580–W584.
- Bankevich A, Nurk S, Antipov D, Gurevich A. a, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19(5):455–477.
- Bayzid MS, Mirarab S, Boussau B, Warnow T. 2015. Weighted statistical binning: enabling statistically consistent genome-scale phylogenetic analyses. *PLoS One* 10(6):e0129183.
- Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics* 13:403.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.
- Bouchet P, Kantor Y, Sysoev A, Puillandre N. 2011. A new operational classification of the Conoidea (Gastropoda). *J Molluscan Stud*. 77(3):273–308.
- Bouchet P, Lozouet P, Sysoev AV. 2009. An inordinate fondness for turrids. *Deep-Sea Res II* 56(19–20):1724–1731.
- Bouchet P, Strong E. 2010. Historical name-bearing types in marine molluscs: an impediment to biodiversity studies? In: Polaszek A, editor. *Systema Naturae 250*. London: CRC Press. p. 63–74.
- Boussac J. 1911. Etudes stratigraphiques et paléontologiques sur le Nummulitique de Biarritz. *Annales Hébert*, 5:1–96.
- Bradley RD, Bradley LC, Garner HJ, Baker RJ. 2014. Assessing the value of natural history collections and addressing issues regarding long-term growth and care. *BioScience* 64(12):1150–1158.
- Bragg JG, Potter S, Bi K, Moritz C. 2016. Exon capture phylogenomics: efficacy across scales of divergence. *Mol Ecol Resour*. 16(5):1059–1068.
- Cariou M, Duret L, Charlat S. 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecol Evol*. 3(4):846–852.
- Combosch DJ, Lemer S, Ward PD, Landman NH, Giribet G. 2017. Genomic signatures of evolution in *Nautilus*—An endangered living fossil. *Mol Ecol*. 26(21):5923–5938.
- Couto DR, Bouchet P, Kantor YI, Simone LRL, Giribet G. 2016. A multi-locus molecular phylogeny of Fasciariidae (Neogastropoda: buccinoidea). *Mol Phylogenet Evol*. 99:309–322.
- da Fonseca RR, Albrechtsen A, Themudo GE, Ramos-Madrugal J, Sibbesen JA, Marett L, Zepeda-Mendoza ML, Campos PF, Heller R, Pereira RJ. 2016. Next-generation biology: sequencing and data analysis approaches for non-model organisms. *Mar Genomics* 30:3–13.
- Duda, T. F, Jr, A. J. Kohn. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. *Mol Phylogenet Evol*. 34 (2005): 257–72.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32(5):1792–1797.

- Faircloth B, Glenn T. 2014. Protocol: Preparation of an AMPure XP substitute (AKA Serapure). Web Doc. Doi 10: J9MW2F26.
- Faurby S, Eiserhardt WL, Svenning J. 2016. Strong effects of variation in taxonomic opinion on diversification analyses. *Methods Ecol Evol.* 7(1):4–13.
- Fedosov A, Kantor Y. 2007. Toxoglossan gastropods of the subfamily Crassispirinae (Turridae) lacking a radula, and a discussion of the status of the subfamily Zemaciinae. *J Molluscan Stud.* 74(1):27–35.
- Fedosov A, Puillandre N, Kantor Y, Bouchet P. 2015. Phylogeny and systematics of mitriform gastropods (Mollusca: gastropoda: neogastropoda). *Zool J Linn Soc.* 175(2):336–359.
- Fedosov AE. 2007. Anatomy of accessory rhynchodeal organs of *Vepracula vepratrica* and *Tritonoturris subbrissoides*: new types of foregut morphology in Raphitomininae (Conoidea). *Ruthenica Russ Malacol J.* 17:33–41.
- Fedosov AE, Puillandre N, Herrmann M, Dgebuadze P, Bouchet P. 2017. Phylogeny, systematics, and evolution of the family Costellariidae (Gastropoda: neogastropoda). *Zool J Linn Soc.* 179:541–626.
- Galindo LA, Puillandre N, Strong EE, Bouchet P. 2014. Using microwaves to prepare gastropods for DNA barcoding. *Mol Ecol Resour.* 14(4):700–705.
- Galindo LA, Puillandre N, Utge J, Lozouet P, Bouchet P. 2016. The phylogeny and systematics of the Nassariidae revisited (Gastropoda, Buccinoidea). *Mol Phylogenet Evol.* 99:337–353.
- Giribet G, Hormiga G, Edgecombe GD. 2016. The meaning of categorical ranks in evolutionary biology. *Org Divers Evol.* 16(3):427–424.
- Glibert M. 1973. Révision des Gastropoda du danien et du montien de la Belgique: I-Les Gastropoda du calcaire de Mons. Institut royal des sciences naturelles de Belgique [r. Vautier, 31].
- Gonzales DTT, Saloma CP. 2014. A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. *Toxicon off J Int Soc Toxinol.* 92:66–74.
- González VL, Andrade SC, Bieler R, Collins TM, Dunn CW, Mikkelsen PM, Taylor JD, Giribet G. 2015. A phylogenetic backbone for Bivalvia: an RNA-seq approach. *Proc R Soc B* 282(1801):20142332.
- Gougerot L, Le Renard J. 1981. Clefs de détermination des petites espèces de gastéropodes de l'Eocène du Bassin parisien. 15-Les genres Raphitoma et Mangelia. Cahiers des Naturalistes, Bulletin des Naturalistes Parisiens, 69-82.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29(7):644–652.
- Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015. Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol.* 32(4):835–845.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Le SV. 2017. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 35:518–522.
- Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program. *Genome Res.* 9(9):868–877.
- Huelsenbeck JP, Ronquist F, Hall B. 2001. MrBayes: bayesian inference of phylogeny. *Bioinformatics* 17(8):754–755.
- Hugall AF, Hara TDO, Hunjan S, Nilsen R, Moussalli A. 2016. An exon-capture system for the entire class Ophiuroidea. *Mol Biol Evol.* 33(1):281–294.
- Isaac NJB, Mallet J, Mace GM. 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends Ecol Evol.* 19(9):464–469.
- Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14(6):587.
- Kantor Y, Fedosov A, Puillandre N. 2018. New and unusual deep-water Conoidea revised with shell, radula and DNA characters. *Ruthenica Russ Malacol J.* 28:47–82.
- Kantor Y, Horro J, Rolan E, Puillandre N. 2018. *Paraclavatulula* (Gastropoda; Conoidea; Clavatulidae), a new genus with a distinctive radula type from West Africa. *J Molluscan Stud.* 84:275–284.
- Kantor Y, Sysoev A. 1996. New data of the foregut of Raphitomininae (Conoidei, Conidae). *Ruthenica Russ Malacol J.* 5:155–157.
- Kantor YI. 2003. Morphological prerequisites for understanding neogastropod phylogeny. *Boll Malacol.* 38:161–173.
- Kantor YI, Strong EE, Puillandre N. 2012. A new lineage of Conoidea (Gastropoda: neogastropoda) revealed by morphological and molecular data. *J Molluscan Stud.* 78(3):246–255.
- Kantor YI, Sysoev AV. 1989. The morphology of toxoglossan gastropods lacking a radula, with a description of new species and genus of Turridae. *J Molluscan Stud.* 55(4):537–549.
- Kantor YI, Taylor JD. 1994. The foregut anatomy of *Strictispira paxillus* (Reeve, 1845) (Conoidea: strictispiridae). *J Molluscan Stud.* 60(3):343–346.
- Kantor YI, Taylor JD. 2002. Foregut anatomy and relationships of raphitominine gastropods (Gastropoda: conoidea: raphitomininae). *Boll Malacol.* 38(Suppl 4):83–110.
- Katoh K, Kuma K-I, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33(2):511–518.
- Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Santos SR, Schander C, Moroz LL, Lieb B, et al. 2011. Phylogenomics reveals deep molluscan relationships. *Nature* 477(7365):452–456.
- Kohn AJ. 1990. Tempo and mode of evolution in Conidae. *Malacologia* 32:55–67.
- Lamarck J.-B. 1803. Mémoires sur les fossiles des environs de Paris (suite 1). Annales du Muséum d'Histoire Naturelle, 23:163–169.
- Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol Biol.* 14:82.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol.* 34:772–773.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
- Lee H, Samadi S, Puillandre N, Tsai M-H, Dai C-F, Chen W-J. 2016. Eight new mitogenomes for exploring the phylogeny and classification of Vetigastropoda. *J Molluscan Stud.* 82(4):534–541.
- Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database Collaboration. 2010. The sequence read archive. *Nucleic acids research* 39(suppl_1):D19–D21.
- Lemmon EM, Lemmon AR. 2013. High-throughput genomic data in systematics and phylogenetics. *Annu Rev Ecol Evol Syst.* 44(1):99–121.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
- Lozouet P. 2017. Les Conoidea de l'Oligocène supérieur (Chattien) du bassin de l'Adour (Sud-Ouest de la France). *Cossmanniana* 19:3–180.
- Magallon S, Sanderson MJ. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55(9):1762–1780.
- Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21):2957–2963.
- Martins EP, Hansen TF. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am Nat.* 149(4):646–667.
- McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol.* 66(2):526–538.
- McLean JH. 1971. A revised classification of the family Turridae, with the proposal of new subfamilies, genera and subgenera from the Eastern pacific. *Veliger* 14:114–130.
- Medinskaya AI. 1999. Foregut of the Cochlespirinae (Gastropoda, Conoidea, Turridae). *Zoosystema* 21:171–198.
- Medinskaya AI, Sysoev A. 2003. The anatomy of *Zemacies excelsa*, with a description of a new subfamily of Turridae (Gastropoda, Conoidea). *Ruthenica Russ Malacol J.* 13:81–87.
- Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* 2010(6):pdb.prot5448.

- Miller BA. 1970. Studies on the biology of Indo-Pacific Terebra. Durham: University of New Hampshire.
- Miller BA. 1971. Feeding mechanisms of the family Terebridae. *Annu Rep Am Malacol Union* 72:74.
- Mirarab S, Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31(12):i44–i52.
- Modica MV, Bouchet P, Cruaud C, Utge J, Oliverio M. 2011. Molecular phylogeny of the nutmeg shells (Neogastropoda, Cancellariidae). *Mol Phylogenet Evol.* 59(3):685–697.
- Nabhan AR, Sarkar IN. 2012. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Brief Bioinform.* 13(1):122–134.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Olivera BM. 2006. Conus peptides: biodiversity-based discovery and exogenomics. *J Biol Chem.* 281(42):31173–31177.
- Orme D. 2013. The caper package: comparative analysis of phylogenetics and evolution in R. *R Package Version* 5:1–36.
- Osca D, Templado J, Zardoya R. 2015. Caenogastropod mitogenomics. *Mol Phylogenet Evol.* 93:118–128.
- Phuong M, Mahardika G. 2018. Targeted sequencing of venom genes from cone snail genomes reveals coupling between dietary breadth and venom diversity. *Mol Biol Evol.* 35(5):1210–1224.
- Platt RN, Faircloth BC, Sullivan KA, Kieran TJ, Glenn TC, Vandeweyer MW, Lee TE, Baker RJ, Stevens RD, Ray DA. 2017. Conflicting evolutionary histories of the mitochondrial and nuclear genomes in New World *Myotis* Bats. *Syst Biol* 67(2): 236–249.
- Ponder W. 1970. Some aspects of the morphology of four species of the neogastropod family Marginellidae with a discussion on the evolution of the toxoglossan poison gland. *J Malacol Soc Aust.* 2:55–81.
- Ponder WF, Lindberg DR. 1997. Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. *Zool J Linn Soc.* 119(2):83–265.
- Portik D, Smith L, Bi K. 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: amphibia, Order: anura). *Mol Ecol Resour.* 16(5):1069–1083.
- Powell AWB. 1942. The New-Zealand Recent and fossil Mollusca of the family Turridae. With general notes on turrid nomenclature and systematics. *Bull Auckl Inst Mus.* 2:1–192.
- Powell AWB. 1966. The molluscan families Speightiidae and Turridae. An evaluation of the valid taxa, both Recent and fossil, with lists of characteristics species. *Bull Auckl Inst Mus.* 5:5–184.
- Prashanth JR, Brust A, Jin A-H, Alewood PF, Dutertre S, Lewis RJ. 2014. Cone snail venomics: from novel biology to novel therapeutics. *Future Med Chem.* 6(15):1659–1675.
- Puillandre N, Kantor Y, Sysoev A, Couloux A, Meyer C, Rawlings T, Todd JA, Bouchet P. 2011. The dragon tamed? A molecular phylogeny of the Conoidea (Mollusca, Gastropoda). *J Molluscan Stud.* 77(3):259–272.
- Puillandre N, Samadi S, Boisselier MC, Sysoev AV, Kantor YI, Cruaud C, Couloux A, Bouchet P. 2008. Starting to unravel the toxoglossan knot: molecular phylogeny of the “turrids” (Neogastropoda: conoidea). *Mol Phylogenet Evol.* 47(3):1122–1134.
- Rabosky DL. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol Bioinformatics* 2:117693430600200–117693430600250.
- Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9(2):e89543.
- Rambault, A, Drummond A. 2003. 2005. Tracer. MCMC Trace Analysis Tool version 1.3. Available from: <http://beast.bio.ed.ac.uk/Tracer>, last accessed July 29, 2018.
- Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22(5):939–946.
- Schrödl M, Stöger I. 2014. A review on deep molluscan phylogeny: old markers, integrative approaches, persistent problems. *J Nat Hist.* 48(45-48):2773–2804.
- Schnetler KI. 2001. The Selandian (Paleocene) mollusc fauna from Copenhagen, Denmark: the Poul Harder 1920 collection. *Geology of Denmark Survey Bulletin*, 37:1-85.
- Sheridan R, van Mol J, Bouillon J. 1973. Etude morphologique du tube digestif de quelques Turridae (Mollusca, Gastropoda, Prosobranchia, Toxoglossa) de la région de Roscoff. *Cah Biol Mar.* 14:159–188.
- Shimek RL, Kohn AJ. 1981. Functional morphology and evolution of the toxoglossan radula. *Malacologia* 20:423–438.
- Simakov O, Marletaz F, Cho S-J, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo D-H, Larsson T, Lv J, Arendt D, et al. 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* 493(7433):526–531.
- Simone LRL. 1999. The anatomy of *Cochlespira* Conrad (Gastropoda, Conoidea, Turridae) with a description of a new species from the Southeastern coast of Brazil. *Rev Bras Zool.* 16(1):103–115.
- Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31.
- Smit AF, Hubley R, Green P. 1996. RepeatMasker.
- Smith SA, Wilson NG, Goetz FE, Feehery C, Andrade SCS, Rouse GW, Giribet G, Dunn CW. 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature* 480(7377):364–369.
- Stadler T, Rabosky DL, Ricklefs RE, Bokma F. 2014. On age and species richness of higher taxa. *Am Nat.* 184(4):447–455.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
- Tanner AR, Fuchs D, Winkelmann IE, Gilbert MTP, Pankey MS, Ribeiro AM, Kocot KM, Halanych KM, Oakley TH, da Fonseca RR, et al. 2017. Molecular clocks indicate turnover and diversification of modern coleoid cephalopods during the Mesozoic Marine Revolution. *Proc R Soc B* 284(1850):20162818.
- Taylor JD, Kantor YI, Sysoev AV. 1993. Foregut anatomy, feedings mechanisms and classification of the Conoidea (= Toxoglossa)(Gastropoda). *Bull Nat Hist Mus Lond Zool.* 59:125–170.
- Taylor JD, Morris NJ. 1988. Relationships of neogastropods. *Malacol Rev.* 4:167–179.
- Teasdale LC, Köhler F, Murray KD, O’Hara T, Moussalli A. 2016. Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. *Mol Ecol Resour.* 16(5):1107–1123.
- Tracey S, Craig B, Belliard L, Gain O. 2017. One, four or forty species? – early Conidae (Mollusca, Gastropoda) that led to a radiation and biodiversity peak in the late Lutetian Eocene of the Cotentin, NW France. *Carnets Voyag Paléontol Dans Bassin Anglo-Paris* 3:1–38.
- Tucker JK, Tenorio MJ. 2009. Systematic classification of Recent and fossil conoidean gastropods. Hackenheim (Germany): Conchbooks.
- Uribe JE, Zardoya R, Puillandre N. 2018. Phylogenetic relationships of the conoidean snails (Gastropoda: caenogastropoda) based on mitochondrial genomes. *Mol Phylogenet Evol.* in press.
- Wen J, Ickert-Bond SM, Appelhans MS, Dorr LJ, Funk VA. 2015. Collections-based systematics: opportunities and outlook for 2050. *J Syst Evol.* 53(6):477–488.
- Wiens JJ, Lapoint RT, Whiteman NK. 2015. Herbivory increases diversification across insect clades. *Nat Commun.* 6:8370.
- WORMS. 2018. World Register of Marine Species. Available from: <http://www.marinespecies.org>. Appeltans W, Bouchet P, Boxshall GA, Fauchald K, Gordon DP, Hoeksema BW, Poore GCB, van Soest RWM, Stöhr S, Walter TC, Costello MJ.

- Yeates DK, Meusemann K, Trautwein M, Wiegmann B, Zwick A. 2016. Power, resolution and bias: recent advances in insect phylogeny driven by the genomic revolution. *Curr Opin Insect Sci.* 13:16–23.
- Zachos FE, Apollonio M, Bärman EV, Festa-Bianchet M, Göhlich U, Habel JC, Haring E, Kruckenhauser L, Lovari S, McDevitt AD, et al. 2013. Species inflation and taxonomic artefacts—A critical comment on recent trends in mammalian classification. *Mamm Biol-Z Für Säugetierkd* 78(1):1–6.
- Zapata F, Wilson NG, Howison M, Andrade SC, Jörger KM, Schrödl M, Goetz FE, Giribet G, Dunn CW. 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc R Soc B* 281(1794):20141739.