



**HAL**  
open science

# Taxonomie intégrative, phylogénie et diversification des Turridae (Gastropoda, Conoidea)

Paul Zaharias

► **To cite this version:**

Paul Zaharias. Taxonomie intégrative, phylogénie et diversification des Turridae (Gastropoda, Conoidea). Systématique, phylogénie et taxonomie. Muséum national d'Histoire Naturelle (Paris, France), 2019. Français. NNT: . tel-03902778

**HAL Id: tel-03902778**

**<https://hal.science/tel-03902778v1>**

Submitted on 16 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License



**MUSEUM NATIONAL D'HISTOIRE NATURELLE**  
Ecole Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2019

N°attribué par la bibliothèque

□□□□□□□□□□□□□□

**THESE**

Pour obtenir le grade de

**DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE  
NATURELLE**

Spécialité : Systématique, Evolution

Présentée et soutenue publiquement par

**Paul Zaharias**

Le 13 Décembre 2019

---

**Taxonomie intégrative, phylogénie et diversification  
des Turridae (Gastropoda, Conoidea)**

---

Sous la direction de : **Nicolas PUILLANDRE**, Maître de Conférence HDR, MNHN

**JURY :**

Mme. Williams, Suzanne	Head of Division, Natural History Museum, London, United-Kingdom	Rapporteuse
M. Giribet, Gonzalo	Professor, Harvard University, Cambridge, USA	Rapporteur
M. Achaz, Guillaume	Professeur, Muséum national d'Histoire naturelle, Paris, France	Examineur
Mme. Cruaud, Astrid	Chargée de Recherche, INRA, Montferrier-sur-Lez, France	Examinatrice
Mme. Modica, Maria-Vittoria	Assistant Professor, Stazione zoologica Anton Dohrn, Roma, Italia	Examinatrice
M. Puillandre, Nicolas	HDR, Muséum national d'Histoire naturelle, Paris, France	Directeur de thèse





*A mes parents, systématiciens à leur façon*



## Remerciements

Avant toute chose, y compris les remerciements plus officiels, mes premières pensées vont évidemment à Nicolas. N'être en relation qu'avec un seul encadrant pouvait laisser présager des conflits et des frustrations sur le long terme, et ce à tous les niveaux : intellectuel, moral, matériel... Il n'en a rien été. Nicolas s'est montré d'une patience (ceux qui me connaissent savent) et d'une disponibilité à tout égard. Mais il en faut plus pour faire un bon encadrant, et je garde une mémoire très nette des innombrables échanges, débats et autres explorations intellectuelles stimulantes que nous avons pu avoir depuis bientôt 4 ans. Si bien que ce n'est pas sans nostalgie (si, déjà !) que je rends ce manuscrit, symbole de la fin de l'encadrement, mais début – je l'espère – d'une complicité pérenne.

Ensuite, je souhaiterais remercier l'Institut de Systématique, Évolution, Biodiversité et son directeur, Philippe Grandcolas, pour son accueil au sein de cette UMR. Je tiens également à remercier les membres du secrétariat de l'ISYEB, composante essentielle des unités de recherche. Je tiens également à remercier Sarah Samadi à double titre : celui de responsable de mon équipe d'accueil (3E), ainsi qu'en tant qu'HDR durant les premiers mois de la thèse, car sans cela cette thèse n'aurait tout simplement pas eu lieu.

Je souhaite également remercier les membres de mon jury : Suzanne Williams, Gonzalo Giribet, Guillaume Achaz, Astrid Cruaud et Maria-Vittoria Modica. Leur travail est une source d'inspiration scientifique, et m'a apporté ses lumières pendant la rédaction de ce manuscrit. J'en profite également pour remercier le comité de thèse, en les personnes de Yuri Kantor et Eric Pante, également sources d'inspiration scientifique, ainsi que les représentants de l'ED Vincent Debat et Violaine Llaurens qui ont participé à mes comités.

Merci chaleureusement à tous les membres de l'équipe 3E physiquement proches, en particulier ceux qui ont réussi à me supporter au quotidien pendant 3 ans : Nico, Sarah, Dario, Julien, Magalie, Laure, et ceux qui ont échappés aux 3 ans : Malcolm, Perrine, Sophie et Juliette (team aquarium !), ainsi que Didier. Je souhaite également remercier le groupe polyphylétique, très éloigné géographique (à l'autre bout du jardin), mais que j'ai quand même envahi ponctuellement : Philippe, Virginie, Gilberto, Abde, Alice, Manu. Un très grand merci particulièrement à Barbara pour sa gentillesse et sa disponibilité, et à Philippe Bouchet, Philippe Maestrati et Pierre Lozouet pour nos échanges malacologiques, leur expertise et leur expérience. Enfin, je souhaite remercier Laure et Sarah, responsables des missions d'explorations BIOMAGLO et KANADEEP (respectivement). Ces missions de terrains ont été extraordinairement stimulantes et précieuses pour former un début de connaissance naturaliste. Je remercie donc également les programmes Tropical Deep-Sea Benthos et La Planète Revisitée.

Merci à l'école doctorale 227 de m'avoir accordé ces trois années très enrichissantes de monitorat. Les interactions autour de ces missions ont été nombreuses, merci à Emilie Detouillon, Natacha Ouvrié, Nathalie Machon, Christine Russello, Anaïs Ranera, Alexandra Moreira, Guillaume Lecointre, Charlène, Véro, René, Nico, Kévin, Perle, Margot & Thomas (il marche bien cet atelier !).

Merci à tous les malacologues du monde entier avec qui j'ai pu échanger du matériel, qui font voyager les photos, les coquilles et les bouts de barbaques par monts et par vaux : Jon Todd, Andreia Salvador, Yasunori Kano, Anders Hallan & Francesco Cricione, Jose Rosado,

Severine Hannam, Gonzalo Giribet & Adam Baldinger, et bien évidemment Sasha (« this isn't exactly chicken soup »). Merci à un malacologue en particulier, Yuri, maître des radulas, qui m'a formé à l'anatomie des Turridae et initié aux secrets de la dissection (« girlfriends eyebrows are the best »).

Merci à tous les molécularistes, notamment Laetitia Aznar et Agnès Dettai pour leurs conseils, et merci surtout au Service de Systématique Moléculaire et aux membres passés ou actuels : Delphine, Céline, Carole, Régis, Josie, Thomas, Isabelle. Merci tout spécialement à Delphine pour son accompagnement dans la longue épreuve de la préparation de banques. Je voudrais également remercier Jawad Abdelkrim (prêt à aider tout le temps, même pour un produit en croix !) Julio Pedraza, et Mark Phuong pour m'avoir initié aux joies de la bioinformatique et au côté obscure de Python.

Je tiens à remercier également la SFS, et son conseil, duquel j'ai pu avoir la chance d'être membre pendant 4 ans. Merci particulièrement à Marc, pour l'organisation de ces journées à Banyuls, et à Véronique, pour l'organisation de cette autre journée. La systématique est en marche !

Merci à tous les méthodologistes qui aiment se prendre la tête avec de grands concepts ou de tout petits détails : Agnès, Artémis, Bouziane, Damien, Eric, Fabien, Gilberto, Giulia, Jawad, Jérémie, Jorge, Laetitia, Laetitia, Lou, Lucas, Malcolm, Maria-Vittoria, Nico, Paul, Perrine, René, Sasha, Stéphane, Valentin. A ce titre, je remercie particulièrement les membres du SPECTRE pour leur talent, et pour créer un groupe de discussion alternatif, où flottent les vapeurs d'esprits échauffés. Merci à Valentin, catalyseur de nouvelles idées, et à Jérémie, catalyseur. Merci à Laetitia de toujours tout savoir, mais de douter en même temps que moi. Merci à Sasha, pour me montrer qu'on peut être bon en tout.

Merci à tous les amis, Mits, Valentin, Laetitia, Malcolm, Sandra, Zoe, Haris, Nina, Chris, Claudia, Mathilde, Donald, Lucas, Jérémie, Anaïs, Perrine, Marc, Mark, Sophie, Morgane, Charlène, Aurélie, Paul, Nico, Philippe, Ninon, Dimitri, Alexis, Laura, les collègues du CR2P, les FTs, les habitués de TS et tous ceux que j'oublie au passage.

Enfin merci à ma famille, Maman, Papa, Lisa, Claude et Doris, tous les cousins, tantes, oncles, grands-parents. Et merci Isa, tout est mieux avec toi que sans.

La présente thèse ne constitue pas une publication au sens  
du Code International de Nomenclature Zoologique.  
[Code, Recommandation 8E]







## SOMMAIRE

INTRODUCTION.....	13
A. INTRODUCTION GENERALE .....	14
B. INTRODUCTION A L'HISTOIRE NATURELLE DES TURRIDAE .....	16
1. Bref historique d'une taxonomie chaotique .....	16
2. Classification moderne & premières phylogénies .....	23
3. Biologie.....	28
4. Conclusion .....	49
5. Planches .....	49
CHAPITRE 1 : TAXONOMIE INTEGRATIVE .....	85
1. La réalité de l'espèce et son importance dans l'histoire de la taxonomie .....	87
2. Ontologie et épistémologie .....	87
3. « Définition » et « critères » d'espèce .....	89
4. Opérationnalité et hiérarchisation des critères chez les Turridae .....	90
5. Critique .....	95
6. Un exemple de l'application du pipeline de taxonomie intégrative chez le genre <i>Cryptogemma</i> .....	97
7. Résultat du pipeline de taxonomie intégrative sur l'ensemble des Turridae.....	140
8. Arbre phylogénétique des MOTUs.....	144
9. Conclusion .....	147
CHAPITRE 2 : UTILISATION EXPERIMENTALE DES TRANSCRIPTOMES EN SYSTEMATIQUE .....	149
1. Données, Temps, Argent : à la recherche du meilleur compromis pour construire des phylogénies moléculaires .....	150
2. A la recherche de nouveaux marqueurs de type « barcode » .....	163
3. Conclusion .....	169

CHAPITRE 3 : PHYLOGENIE ET DIVERSIFICATION DES TURRIDAE.....	171
1. Introduction .....	173
2. Material and Methods.....	175
3. Results .....	182
4. Discussion.....	189
CONCLUSIONS & PERSPECTIVES.....	201
A. Une stratégie opérationnelle de systématique exploratoire .....	202
B. Les Turridae ou comment être un clade en déclin et faire partie d'un groupe hyperdiversifié.....	205
C. Perspectives.....	207
Bibliographie.....	211
ANNEXE 1 : .....	231
ANNEXE 2 : .....	261
ANNEXE 3 : .....	282

## Articles réalisés pendant la thèse ou en cours de réalisation

(\* : inclus dans la thèse)

### **Publiés :**

\* **Zaharias, P.**, Pante, E., Gey, D., Fedosov, A. E. & Puillandre, N. (2019). Data, Time and Money: Evaluating the Best Compromise for Inferring Molecular Phylogenies of Non-Model Animal Taxa. *Molecular phylogenetics and evolution*, 106660.

Fassio, G., Modica, M. V., Mary, L., **Zaharias, P.**, Fedosov, A. E., Gorson, J., Kantor, Y. I., Holford, M. & Puillandre, N. (2019). Venom Diversity and Evolution in the Most Divergent Cone Snail Genus *Profundiconus*. *Toxins*, 11(11), 623.

\* Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., **Zaharias, P.**, & Puillandre, N. (2018b). Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Molecular ecology*, 27(22), 4591-4611.

\* Abdelkrim J., Aznar-Cormano L., Fedosov A.E., Kantor Y.I., Lozouet P., Phuong M.A., **Zaharias P.**, Puillandre N. (2018a). Exon-capture based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Mol. Biol. Evol.* 35:2355–2374.

\* Puillandre, N., Fedosov, A. E., **Zaharias, P.**, Aznar-Cormano, L., & Kantor, Y. I. (2017). A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess. *Zoological Journal of the Linnean Society*, 181(2), 243-271.

### **Soumis :**

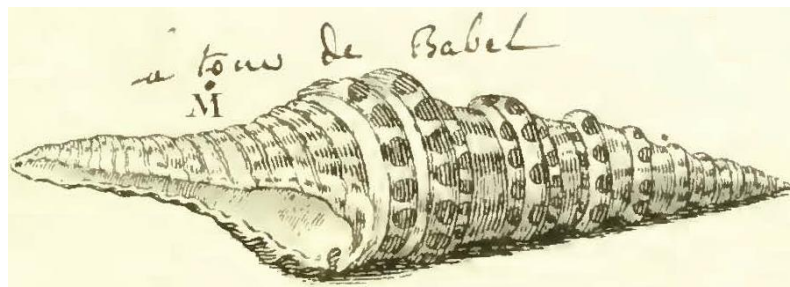
\* **Zaharias, P.**, Kantor, Y. I., Fedosov, A. E., Criscione, F., Hallan, A., Kano, Y., Bardin, J. & Puillandre, N. Just the once will not hurt: DNA suggests species lumping over two oceans in deep-sea snails. *Zoological Journal of the Linnean Society*.

### **En préparation :**

\* **Zaharias, P.**, Kantor, Y. I., Fedosov, A. E. & Puillandre, N. Phylogeny and diversification of Turridae (Gastropoda, Conoidea).

**Zaharias, P.**, Kantor, Y. I., Fedosov, A. E. & Puillandre, N. A new classification of the Turridae.

# INTRODUCTION



## **A. INTRODUCTION GENERALE**

Dans le livre blanc de la systématique, on trouve le programme de cette discipline défini tel que :

*« ... l'étude et la description de la diversité des êtres vivants, la recherche de la nature et des causes de leurs différences et de leurs ressemblances, la mise en évidence des relations de parenté existant entre eux et l'élaboration de classifications traduisant ces relations de parenté ».* – Société Française de Systématique, 1987

Malgré les bouleversements majeurs de biologie ces dernières décennies (annoncés dans ce même livre blanc), cette définition est toujours d'actualité, et ce en dépit de l'utilisation timorée du terme « systématique » dans les projets acutels de recherche. Les différents modules qui composent cette définition sont d'ailleurs représentés par les mots-clefs du sujet de cette thèse. « Taxonomie intégrative » fait référence à « l'étude et la description de la diversité des êtres vivants », tandis que « diversification » fait allusion à « la recherche de la nature et des causes de leurs différences et ressemblances », c'est-à-dire tout simplement à l'étude de leur évolution dans l'espace et le temps. Enfin « phylogénie » fait doublement référence à « la mise en évidence des relations de parenté existant entre eux et l'élaboration de classifications traduisant ces relations de parenté », car les phylogénies constituent à la fois des hypothèses évolutives et la base d'un système classificatoire. En somme, cette thèse pourrait sobrement s'intituler « Systématique des Turridae ».

Il est par ailleurs tout aussi intéressant de constater que ces différents modules disciplinaires ne sont pas vraiment indépendants les uns des autres. En effet, il est tout à fait possible d'étudier la phylogénie d'un groupe sans étudier sa diversification, mais – dans le paradigme actuel de la théorie de l'évolution – la réciproque n'est pas vraie. De même, il est possible d'étudier et de décrire des unités d'êtres vivants sans s'intéresser à leurs liens de parenté, mais il est inconcevable d'établir une phylogénie sans avoir caractérisé au préalable les unités taxonomiques pour lesquelles on souhaite trouver les relations de parenté. Ainsi, l'étude de la taxonomie est une condition nécessaire à l'étude de la phylogénie, elle-même condition nécessaire à l'étude de la diversification. Une étude systématique complète sur un groupe devra donc suivre cette séquence logique.

Les Turridae H. Adams & A. Adams, 1853 (1838) sont un groupe de mollusques gastéropodes marins que l'on retrouve exclusivement dans la zone intertropicale. Ils appartiennent aux Conoidea, un groupe de prédateurs possédant une glande à venin, et qui se servent de leur radula (une structure chitineuse en forme de langue possédant plusieurs dents)

pour injecter un cocktail de toxines dans leurs proies. Parce qu'ils présentent cette combinaison de caractères uniques et un très grand nombre d'espèces (~5500 espèces actuelles décrites, ~10 à 20000 estimées), les Conoidea sont souvent qualifiés de taxon « à succès évolutif ». Des études de diversification publiées cette année sur les familles Conidae et Terebridae (appartenant également aux Conoidea) semblent aller dans ce sens, mais le lien entre la diversification (au sein de ces groupes) et leur nature venimeuse n'a pas pu être établie, suggérant l'implication d'autres traits pouvant expliquer ce « succès évolutif ». L'idée de cette thèse était de continuer ce travail d'investigation sur les Conoidea en choisissant cette fois-ci les Turridae. Ces derniers sont caractérisés, relativement aux autres familles de Conoidea, par une diversité plus faible (~200 espèces actuelles décrites, contre ~950 pour les Conidae par exemple), une plus grande abondance et une plus grande taille que la moyenne des Conoidea, facilitant la documentation de traits morpho-anatomiques, géographiques, écologiques et toxinologiques.

En accord avec la séquence logique exposée plus haut, les objectifs de cette thèse sont de (i) réviser la taxonomie alpha des Turridae, (ii) proposer une phylogénie des Turridae et (iii) étudier la diversification des Turridae, notamment en relation avec certains traits. Etant donné les résultats extrêmement prolifiques de l'utilisation de l'ADN pour répondre à ces problématiques ces 15 dernières années, l'outil moléculaire tiendra une place prépondérante dans cette étude. Pour cela j'ai pu m'appuyer sur les collections exceptionnelles du Muséum national d'Histoire naturelle (MNHN), et plus particulièrement sur les résultats des programmes d'exploration *Tropical Deep-Sea Benthos* (TDSB) et *La Planète Revisitée* (LPR). La méconnaissance de la systématique du groupe a conduit à l'utilisation d'approches pluralistiques et exploratoires, proposant une "stratégie générale d'étude systématique". A terme, je l'espère, cette stratégie pourra être appliquée à grande échelle et utilisées dans les nombreux autres groupes dont la systématique demande à être clarifiée.

Je présenterai d'abord, à la suite de cette introduction, un historique puis un état des connaissances du groupe. Puis, dans le premier chapitre, j'exposerai la stratégie taxonomique utilisée pour délimiter les espèces chez les Turridae. Les unités délimitées dans cette première partie serviront d'unités opérationnelles pour reconstruire la phylogénie. Je proposerai ensuite dans le chapitre 2 d'illustrer quelques utilisations possibles de données transcriptomiques, pouvant avoir des applications en taxonomie et en phylogénie. Enfin, le chapitre 3 sera consacré à l'étude de la phylogénie et de la diversification des Turridae. A terme, cette nouvelle systématique offrira un nouveau cadre de réflexion sur l'évolution du groupe.

## **B. INTRODUCTION A L'HISTOIRE NATURELLE DES TURRIDAE**

Dans cette partie, je propose de récapituler l'état des connaissances actuelles sur les Turridae, allant de l'histoire de leur classification à l'écologie, en passant par des considérations sur la morpho-anatomie du groupe. Les noms « Turridae » et « Conoidea » (et les espèces comprises au sein de cette famille) sont renvoyés à leur définition actuelle et non celle utilisée par les auteurs des différentes époques. S'il faut faire référence aux Turridae dans un contexte donné, l'expression « Turridae *s.l.* » est utilisée, par opposition à l'expression « Turridae *s.s.* » qui ne concerne que les Turridae dans leur définition actuelle. Pour accompagner le lecteur, des planches de coquilles (notées « C ») et de radulas (notées « R ») ont été réalisées et sont consultables à la fin de cette introduction.

### **1. Bref historique d'une taxonomie chaotique**

#### 1.1. La Tour de Babel

Le premier spécimen de Turridae à être figuré dans l'histoire des sciences est certainement représenté par Martin Lister en 1688 dans son *Historiae Conchyliorum*, sous l'appellation *Buccinum R[ostratum] maculatum, Striatum asperum fissurâ ad imum labro* (Fig. 1). Cette même espèce sera plus tard représentée par Rumphius dans son *Thesaurus imaginum* (1711) sous une appellation binominale qui, à un accord latin près, est celle encore utilisée actuellement : *Turris Babylonica*, la Tour de Babel. Dans la littérature pré-linnéenne, c'est le nom vernaculaire « Tour de Babel » qui est retenu pour désigner cette espèce, bien que certains auteurs la regrouperont avec d'autres espèces tantôt sous le nom *Strombus* (Lange, 1722 ; Gualtieri, 1742), tantôt sous le nom *Buccinum* (D'Argenville, 1742 ; Hebenstreit, 1743). Finalement Linné lui donnera le nom de

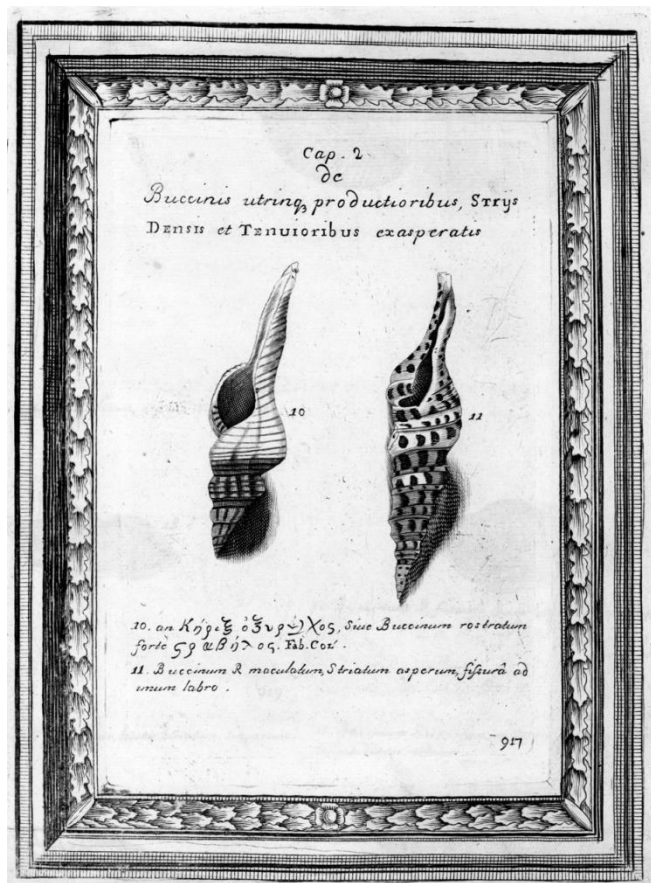


Figure 1 : Page 917 du *Historiae Conchyliorum* de Lister (1688) où figure un spécimen de *Turris*, possiblement *Turris babylonica*, à droite.

*Murex babylonius* dans le Systema Naturae (1758), fixant (sans le savoir) l'espèce-type de la famille Turridae. Le lectotype de *M. babylonius* a récemment été désigné par Kilburn *et al.* (2012) et a été numérisé, encore plus récemment, par la Linnean Society of London (Fig. 2). Curieusement, malgré l'existence d'espèces côtières et abondantes dans cette famille, il faudra attendre les travaux d'Helbling en 1779 – soit près d'un siècle après la première apparition de la Tour de Babel dans la littérature – pour que de nouvelles espèces de Turridae soient décrites.



Figure 2 : Lectotype de *Turris babylonica*.  
Echelle : 1cm ©Linnean Society of London.

### 1.2. *Turris* contre *Pleurotoma*

Après l'imposition – notamment par Linné (1758) – de la nomenclature binominale, nombre de naturalistes vont proposer des nouveaux noms de genre, en plus de décrire de nouvelles espèces. Cependant les règles de nomenclature ne sont pas très contraignantes, notamment en ce qui concerne la priorité des noms, et nombre de naturalistes vont créer et user des noms différents, souvent sans se référer aux travaux précédents, pour désigner la même chose. Le *Murex babylonius* de Linné, échappe raisonnablement à cette tendance ; seuls deux noms de genre alternatifs à *Murex* vont être créés pour désigner l'espèce, *Turris* et *Pleurotoma*. Comme vu plus haut dans le 1.1, le nom *Turris* est proposé par Rumphius dès 1711, mais également par Klein (1753) et, après Linné, par Stadius Müller (1766) et Chemnitz (1780) de façon non valide par rapport au système binominal émergent. Une description d'un *Turris* cette fois-ci correcte mais anonyme, qui s'avère en réalité l'œuvre d'Humphrey, est publiée dans le *Museum Calonnianum* (1797), œuvre considérée aujourd'hui comme invalide nomenclaturalement aux yeux du code. C'est pourtant le *Turris* d'Humphrey qui sera désigné genre porte-nom de la famille Turridae décrite en 1853 par les frères Adams. Enfin, en 1789 et en 1798 sont proposés respectivement (de façon indépendante) le *Turris* de Batsch et le *Turris* de Röding. Mais ce dernier, issu d'un catalogue de vente assez rare, le *Museum Boltenianum*, ne sera reconnu d'intérêt nomenclatural qu'en 1906 par Dall, et ne sera appliqué véritablement qu'à partir du début du 20<sup>e</sup> siècle. Quant au *Turris* Batsch 1789, il ne sera redécouvert qu'en 2010 dans l'article de Dubois & Bour et constitue actuellement l'assignation de genre correcte à l'espèce décrite par Linné, nominativement *Turris babylonica*



Batsch 1789. C'est donc un autre nom de genre pour désigner les Turridae qui va prédominer dans l'usage pendant tout le 19<sup>e</sup> siècle : *Pleurotoma* Lamarck 1799.

Si l'on peut supposer que Lamarck n'a pas eu accès au *Museum Boltenianum*, on peut en revanche se demander pourquoi il a choisi délibérément d'ignorer un siècle de tradition et d'usage du nom *Turris*. C'est en tout cas ce que déplore Dall (1909). Toujours est-il que le genre *Pleurotoma* sera utilisé par la quasi-totalité des naturalistes du 19<sup>ème</sup> et du début du 20<sup>ème</sup> siècle. Or, c'est pendant cette même période qu'ont lieu les grandes expéditions naturalistes, qui rapporteront des dizaines de milliers de coquilles de mollusques en Europe et aux Etats-Unis. Le nombre d'espèces de *Pleurotoma* ne va pas tarder à exploser : on compte 100 espèces décrites en 1834, puis plus de 400 en 1850 et finalement plus de 1000 en 1903 (Fig. 3). Dans une étude plus exhaustive, incluant les fossiles, Tucker (2004)

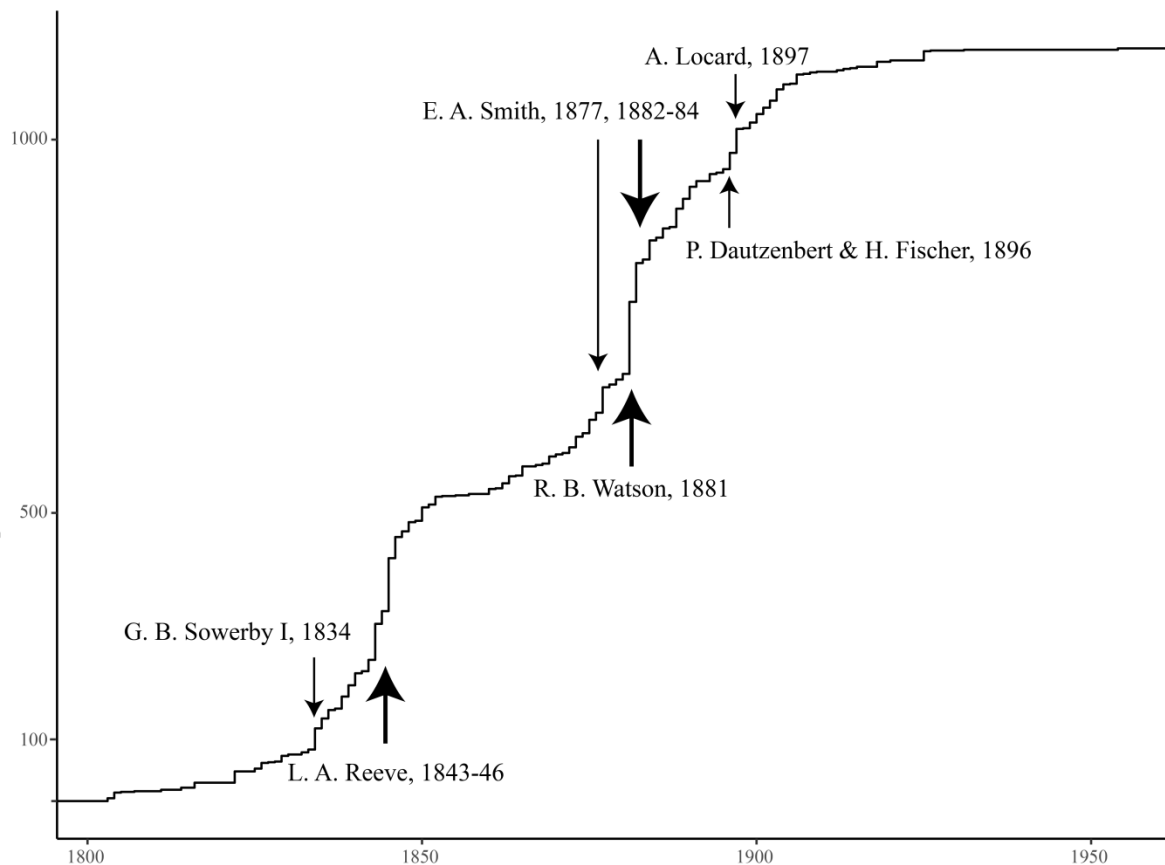


Figure 3 : Courbe cumulative du nombre d'espèces de *Pleurotoma* décrites depuis le 18<sup>ème</sup> siècle. Chiffres d'après Worms. En ordonnée, le nombre d'espèces, en bascisse les dates.

comptabilise pas moins de 3214 espèces décrites originellement sous le nom *Pleurotoma*, dont 1342 d'actuelles. Ces résultats en font un des genres de mollusques les plus riche, aux côtés de *Conus* Linné 1758 et *Murex* Linné 1758, par exemple. Chez les Turridae, la quasi-totalité des espèces actuelles décrites pendant la période 1811-1925 ont pour nom de genre *Pleurotoma* : on n'en compte pas moins de 80, soit environ 40% des espèces actuelles de

Turridae. Les pics de description sur la courbe cumulative des descriptions de *Pleurotoma* coïncident parfaitement avec des travaux de monographie réalisés à la suite de grandes expéditions (Fig. 3). On citera notamment la collection de H. Cuming, constituée majoritairement lors de l'expédition de la goélette *Discover*, qui a servi de socle pour les travaux des Sowerby (par exemple en 1834), et surtout pour la monographie de Reeve du genre *Pleurotoma* (1843-46), ainsi que l'expédition *Challenger* qui alimentera les travaux de R. B. Watson (1881) et de E. A. Smith (1877, 1882-84), ce dernier profitant également des récoltes de l'*Investigator*. Dans une moindre mesure, les travaux de A. Adams & L. Reeve (1850 ; *Samarang*), Dall (1881 ; *Blake*), P. Dautzenberg & H. Fischer (1896 ; l'*Hirondelle*, *Princess-Alice*), Locard (1897 ; *Travailleur*, *Talisman*), Schepman (1913 ; *Siboga*) et Thiele (1925 ; *Valdivia*) ont contribué à la description de nouvelles espèces de Turridae.

Dans ce contexte, la famille Turridae (« Turritidae » dans la première version) proposée à raison (car le nom *Turris* est plus ancien) par Adams & Adams (1853) passe inaperçue, et c'est la famille Pleurotomidae (-inae Gray, 1838 ; -aceae Hinds, 1844) qui est en usage, même s'il existe quelques exceptions, principalement dans des études paléontologiques (par exemple Meek, 1876).

### 1.3 De Pleurotomidae à Turridae s.l. à Turridae s.s.

Avec l'augmentation du nombre d'espèces de *Pleurotoma*, mais aussi d'autres genres de Conoidea tels que *Cytherea* Lamarck, 1805, *Mangelia* Risso, 1826, *Drillia* Gray, 1838, *Daphnella* Hinds, 1844, *Bela* Leach, 1847 ou *Clathurella* Carpenter, 1857, le besoin d'une classification des Pleurotomidae se fait sentir. Powell (1942) indique le chiffre de 362 noms génériques proposés pour la famille, dont énormément de synonymes. On citera notamment trois tentatives d'établissement de systèmes vers la fin du 19<sup>ème</sup> siècle: celui de Tryon (1883), celui de Cossman (1896) et celui de Casey (1904). Ces systèmes ne sont que très peu (voir pas) suivi d'un auteur à un autre, ni même par leurs successeurs, et les modifications consistent principalement en des remaniements des statuts des tribus, genres, sous-genres ou sections de façon arbitraire, avec peu de références à des caractères morphologiques ou anatomiques.

En 1904 a lieu un accord international finalisant les règles de nomenclature zoologique, qui donnera lieu en 1905 à la publication des *Règles internationales de la Nomenclature zoologique*, ancêtre du *Code international de nomenclature zoologique* (CINZ) contemporain. Les nouvelles règles renforcent la règle de priorité, et de nombreux débats

entre zoologistes ont lieu pour savoir si certains ouvrages comme par exemples les catalogues de ventes doivent être admis et adoptés par les zoologistes. C'est le cas du *Museum Boltenianum*, dont W. H. Dall sera un fervent partisan de l'acceptation de l'ouvrage comme valide aux yeux du code (Jukes-Browne, 1904). En 1906, Dall publie un article indiquant *Pleurotoma* comme synonyme junior de *Turris*, et en 1908 un second article indiquant « Turritidae » comme famille du genre *Turris*. A partir de cette date, très peu d'espèces seront décrites sous *Pleurotoma* à part par quelques paléontologues ou d'irréductibles allemands (par exemple Thiele), même si plusieurs auteurs déplorent la réhabilitation du *Museum Boltenianum* (par exemple Berry, 1920). Il est à noter qu'un débat s'engage dans les pages du *Nautilus* (vol. 23, 33, 34), entre W. S. Dall (Conservateur Honoraire) et S. S. Berry (22 ans), sur quelle appellation utiliser entre « Turritidae » et « Turridae » – l'histoire donnera raison à Berry, même si certains collègues proches de Dall (par exemple Paul Bartsch) continueront d'utiliser les termes « Turritidae » ou « turritid ».

Ces changements en profondeur dans la nomenclature ne vont toutefois pas permettre un renouveau dans la façon de classer les Turridae, et les systèmes de Dall (1918) ou de Hedley (1922) sont tout aussi arbitraires que ceux de leurs prédécesseurs, basés encore une fois sur un nombre limité de caractères morphologiques. Il faudra attendre Powell (1942) pour que la radula soit prise en compte dans sa première classification des Turridae *s.l.*, même si Thiele (1925, 1929) avait déjà commencé à la prendre en compte dans ce groupe. Powell sera

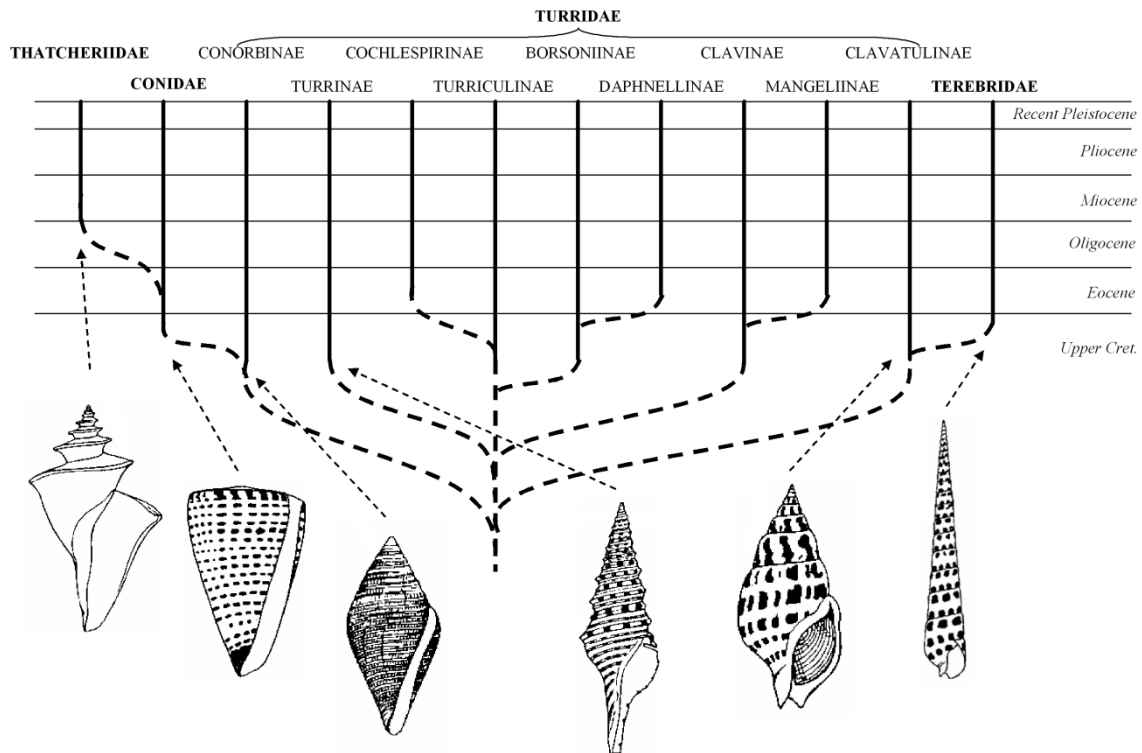


Figure 4 : Phylogénie des Conoidea d'après Powell, 1942 (modifié d'après Puillandre, 2008).

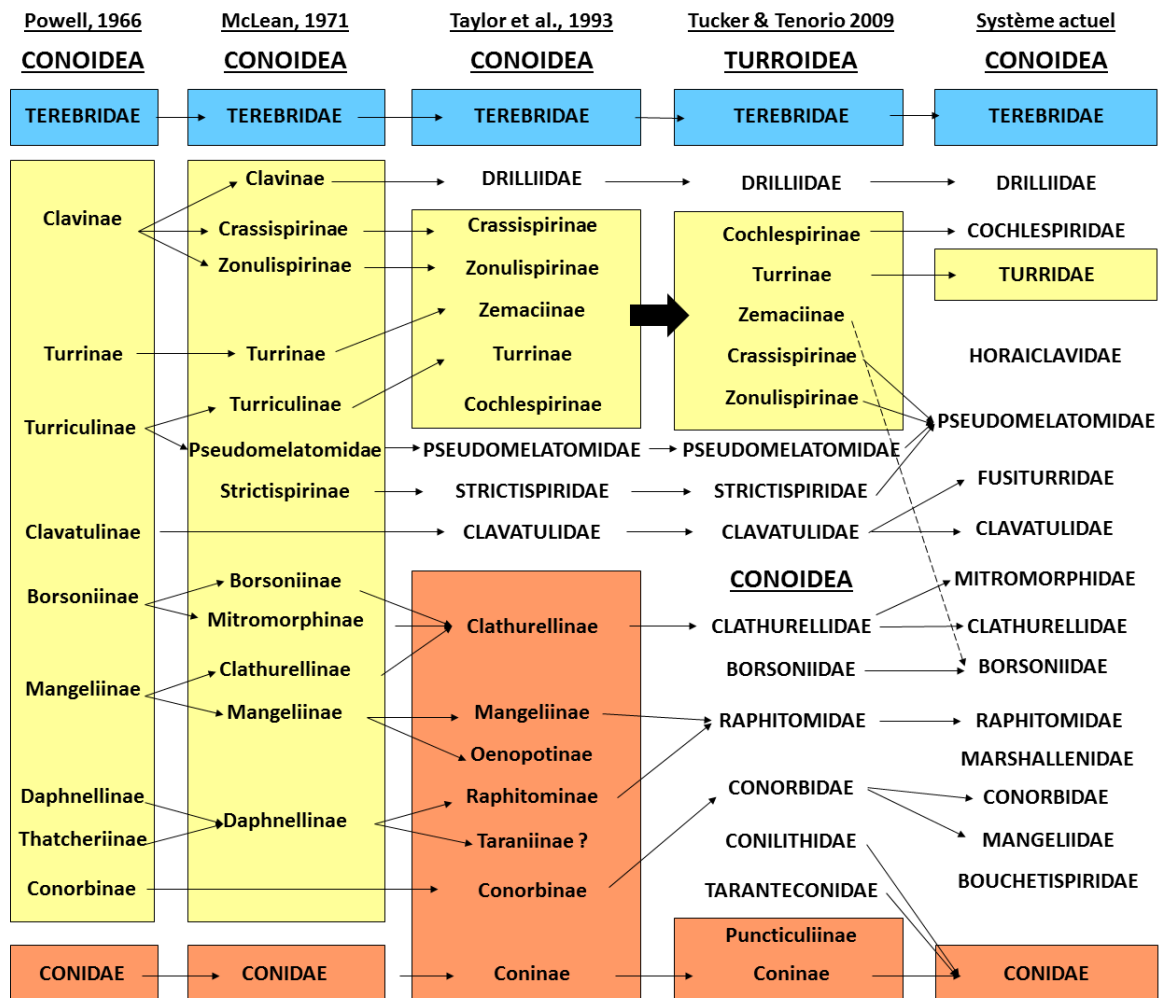


Figure 5 : Evolution de la classification des Conoidea, modifié d'après Puillandre, 2008.

également le premier à parler de « phylogénie » et à proposer une histoire globale du groupe). Fig. 4), même si cette histoire paraît désuète à la lumière de nos connaissances actuelles. Néanmoins, sa classification de 1964-1967 des sous-familles se rapproche sensiblement du système actuel des familles. L'histoire de la classification des Turridae (au sens large) de 1966 à aujourd'hui a déjà été adéquatement synthétisée dans la thèse de Puillandre (2008; Fig. 5), en comparant les systèmes de Powell (1964-1967), McLean (1971) et Taylor *et al.* (1993). Le système de Taylor *et al.* constitue notamment le premier système établi à partir d'une analyse cladistique des caractères morpho-anatomiques des Conoidea, intégrant le concept de monophylie des familles, mêmes si certaines d'entre elles sont finalement redéfinies sur la base de groupe paraphylétiques. Ces travaux ont pu mettre en évidence le caractère homoplasique de la plupart des traits morpho-anatomiques traditionnellement utilisés pour classer les Conoidea, et plus largement les Mollusques. Ils ont également résulté en une réduction drastique de la famille Turridae, auparavant incluant tous les Conoidea (sauf les Conidae et les Terebridae) et en élevant certaines sous-familles au rang de famille. Il faudra

encore attendre les articles issus de la thèse de Puillandre (Puillandre *et al.*, 2008) pour qu'une phylogénie moléculaire – basée sur quelques marqueurs – (quasi-)résolue et stable des Conoidea voit le jour, suivie d'une classification (Bouchet *et al.*, 2011). A noter qu'une classification alternative existe (Tucker & Tenorio, 2009), mais qu'elle ne sera pas prise en compte dans cette thèse. En effet, la classification de Tucker & Tenorio se base principalement sur les caractères de la radula, même si elle prend également en compte d'autres caractères comme la morphologie et les quelques phylogénies moléculaires publiées à l'époque. De plus, le travail des auteurs se concentre principalement sur les Conidae. Enfin, la dernière phylogénie en date des Conoidea (Abdelkrim *et al.*, 2018a – Annexe 2), basée sur une capture d'exons, semble donner raison à la classification de Bouchet *et al.* plutôt qu'à celle de Tucker & Tenorio.

Ce bref historique a pour intérêt d'illustrer la confusion qui a régné sur plus de trois siècles de tentatives d'établissement d'une classification stable chez les Turridae. Durant la majorité de cette période, les travaux se sont concentrés sur la description de coquilles vides rapportées d'expéditions et les débats sur des considérations abstraites de classification, axées principalement sur des remaniements nomenclaturaux. Il faut attendre la monographie de Powell (1964-1967) pour qu'une dimension intégrative soit apportée, permettant un regard critique tant sur les espèces qui ont été décrites (Powell va mettre en synonymie beaucoup de noms) que sur les regroupements de celles-ci en groupes naturels. Cette dimension intégrative va prendre une nouvelle forme au milieu des années 2000 avec la thèse de Puillandre et l'apport des données ADN, permettant l'établissement fiable et rapide d'hypothèses d'espèces. Actuellement l'un des plus grands obstacles à une avancée rapide d'un système classificatoire linnéen stable des Turridae est la prise en compte et le traitement de ces 300 années d'histoire chaotique, où les spécimens-types sont parfois perdus, souvent en mauvais états et surtout non exploitables moléculairement – en tout cas pour le moment. Les études effectuées au cours de cette thèse sur *Lophiotoma* (Puillandre *et al.*, 2017 – Annexe 1) ou sur *Cryptogemma* (voir : article 1 - Chapitre 1) donnent une idée de l'ampleur de la tâche qui attend les taxinomistes.

## **2. Classification moderne & premières phylogénies**

Avant que les Turridae – en tant que famille – ne soient redéfinis dans la classification de Bouchet *et al.* (2011), les différents auteurs faisaient référence à la sous-famille Turrinae pour parler du groupe étudié dans cette thèse. Le groupe Turrinae/Turridae est à peu près stable depuis 50 ans: les genres actuellement compris dans la famille Turridae sont déjà tous listés dans la monographie de Powell (1964-1967), mis à part ceux décrits pendant ou après la monographie comme *Decollidrillia* Habe & Ito, 1965, *Gemmuloborsonia* Shuto, 1989 et *Iotyrris* Medinskaya & Sysoev, 2001. On trouve également chez Powell quelques autres genres qui ont été transférés par la suite dans d'autres familles de Conoidea comme par exemple *Carinoturris* Bartsch, 1944 (Pseudomelatomidae) ou *Viridoturris* Powell, 1964 (Borsoniidae).

Jusqu'aux années 2000, les réflexions phylogénétiques sur les Turridae se résument soit – comme tout systématicien post-darwinien – dans les assignations taxinomiques (les espèces d'un même genre partagent un ancêtre commun), soit – comme Powell (1942, 1964-1967) – dans la recherche d'ancêtres dans le registre fossile ou soit encore – comme McLean (1971), Sheridan *et al.*, (1973) et Kilburn (1983) – dans des réflexions sur la polarité des caractères (ancestraux ou dérivés), notamment concernant la radula, le type de protoconque ou bien de la partie antérieur du tube digestif (très modifié chez les Conoidea). Etrangement, ni la révolution hennigienne (Hennig, 1966), ni la « guerre des taxinomistes » qui suivit dans les années 60-70 (cf. Hull, 1988) ne semblent avoir motivé les malacologues de l'ère pré-moléculaire à établir une phylogénie des Conoidea qui soit issue d'une analyse méthodologique. Seules exceptions au tableau, les analyses cladistiques et phénétiques faites « à la main » de Shimek & Kohn (1981) et l'analyse cladistique de Taylor *et al.* (1993), réalisée avec le logiciel PAUP 3.0. La première étude présente un cladogramme basé sur quatre caractères de la radula, qui sera mis en perspective avec deux phénogrammes basée sur cinq paramètres continus de la coquille. Les différents dendrogrammes de cette étude ne montrent aucune corrélation entre eux, comme le soulignent les auteurs, mais il s'agit là des premiers arbres phylogénétiques de relations entre sous-familles de Conoidea basés sur des méthodes analytiques et répétables. La deuxième étude est basée sur 43 caractères morpho-anatomiques, mais l'arbre présenté est un consensus majoritaire à 50%, n'indiquant pas les valeurs de soutien à chaque nœud. Cette deuxième étude sera d'ailleurs remise en question par Rosenberg (1997), notamment sur sa reproductibilité. L'approche anatomique sera également utilisée par Medinskaya (2002a) pour reconstruire un cladogramme cette fois des Turridae

mais les résultats ne sont pas convaincants. L'auteure elle-même conclut l'étude en affirmant que chez chaque famille de Conoidea il est possible de trouver des « types anatomiques » censés être caractéristiques d'autres familles de Conoidea. Loin d'être satisfaisantes, ces premières analyses phylogénétiques ont toutefois le mérite de synthétiser les connaissances morpho-anatomiques des différentes familles de Conoidea, notamment avec des réflexions sur la polarité des caractères, et servent toujours de socle pour discuter d'hypothèses évolutives (par exemple Page & Hookham, 2017 ; Holförd *et al.*, 2008).

Une explication possible à la quasi-absence d'étude phylogénétique chez les Turridae, et qui s'étend à nombre de taxons de mollusques, est que la plupart des taxinomistes, confortables dans la description pure (notamment d'espèces), ont parfaitement conscience de la faiblesse des caractères morpho(-anatomiques) pour résoudre les relations de parenté. Une phrase de Kilburn illustre le propos:

*« If any conclusion is possible, it is that protoconch characters probably contain valuable phylogenetic information, but that parallelism and character-reversals [...] may provide confusing data. »* (Kilburn, 1983)

Bien entendu il est possible de trouver des synapomorphies morpho-anatomiques, mais ces synapomorphies sont tellement rares et en sous-nombre par rapport aux autres caractères qu'il est inenvisageable d'espérer trouver un arbre phylogénétique résolu à l'issue d'une analyse cladistique (cf. Rosenberg, 1997) – et même si c'est le cas, de trouver des arguments pour le défendre (cf. Medinskaya, 2002a). Beaucoup de caractères, comme le dit Kilburn sur les protoconques, possèdent un signal phylogénétique mais ne sont pas exempts de cas de convergence également. Ajoutons à cela l'aspect continu des caractères morpho-anatomiques (cf. Puillandre, 2008), la forte variabilité intra-spécifique (voir Chapitre 1) et le très grand nombre d'espèces (souvent mal délimitées) à traiter chez les Conoidea et cela découragerait n'importe-quel systématicien.

Il faudra donc attendre la « révolution moléculaire » pour voir naître les premières phylogénies de Turridae. En 2007, Heralde III *et al.* publient une phylogénie basée sur un fragment d'ARN ribosomique 12S mitochondrial (12S) avec un échantillonnage de 23 espèces côtières de Turridae des Philippines (excepté pour le *Polystira* qui vient des Caraïbes) (Fig. 6a). Cette phylogénie sera améliorée trois ans plus tard (Heralde III *et al.*, 2010) par l'ajout des quelques espèces mais surtout d'un autre fragment d'ARN ribosomique mitochondrial, le 16S (Fig. 6b). Malgré leur échantillonnage réduit, ces phylogénies mettent

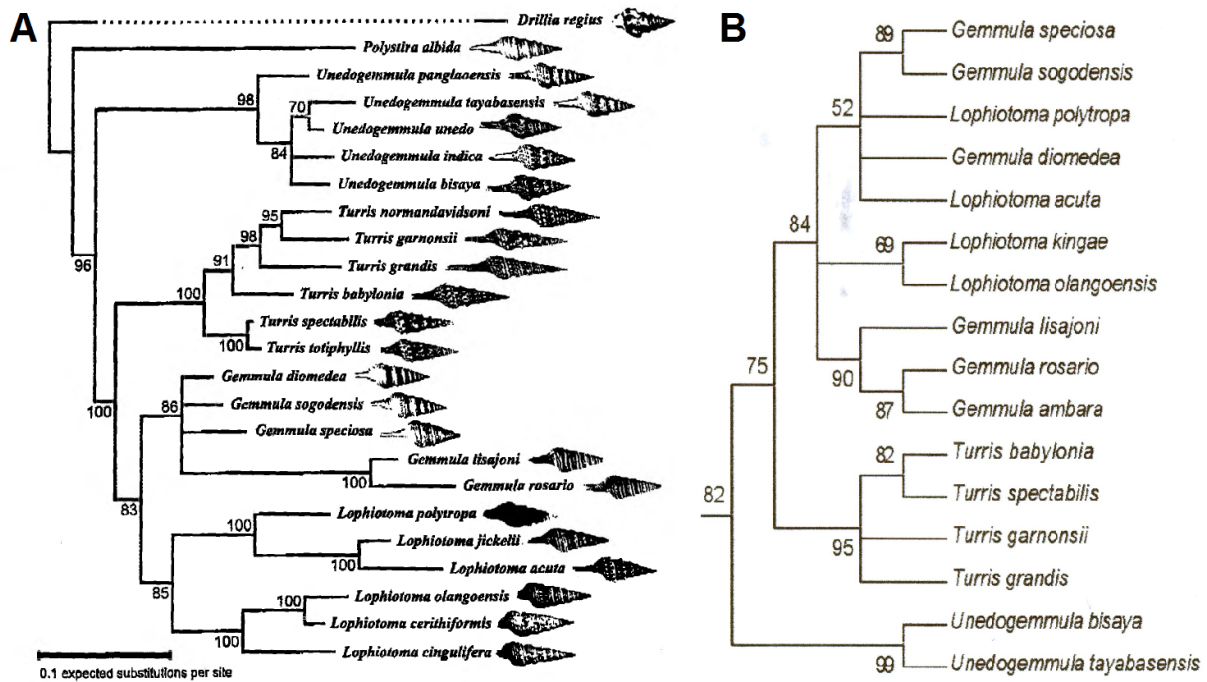


Figure 6 : Phylogénies des Turridae produites par Herald III *et al.* (2007) et Herald III *et al.* (2010). A. Arbre du fragment 12S obtenu en inférence bayésienne. B. Arbre du fragment 16S obtenu par analyse de minimum d'évolution sur MEGA 3.1.

déjà en évidence la paraphylie longtemps suspectée des genres *Gemmula* et *Lophiotoma*. Et, avec l'étude de Puillandre *et al.* (2012b), cette paraphylie va prendre une nouvelle ampleur : la diversité de *Lophiotoma* et *Gemmula* est sous-estimée, et la moitié des espèces de *Gemmula* seraient à décrire. Dans l'arbre établi à partir du fragment mitochondrial du gène cytochrome c oxydase sous-unité 1 (CO1; Fig.7), largement corroboré par l'arbre établi avec le fragment d'ARN ribosomique 28S, le genre *Gemmula* est reparti dans 7 clades (en restant le plus conservateur possible sur les limites des genres). De plus, parmi les groupes de niveau générique de Turridae considérés dans ce manuscrit (voir Chapitre 3), 16 sont retrouvés monophylétiques par l'arbre CO1 (Fig. 7) issu de l'étude de Puillandre *et al.* L'amplitude de ces résultats repose sur la mise en place d'un pipeline original de taxonomie intégrative (Puillandre *et al.*, 2012b : Fig. 1), misant sur un échantillonnage « à l'aveugle » et permettant la découverte de groupes « cryptiques ». C'est de ce pipeline dont je me suis grandement inspiré pour l'établissement de mes Unités Taxinomiques Opérationnelles Moléculaires (MOTU en anglais) et que je décrirai plus précisément dans le Chapitre 1. Car si la révolution moléculaire a permis de nombreuses publications de phylogénies de mollusques, elle a permis également des publications encore plus nombreuses de délimitations d'espèces (cf. Chapitre 1). Et, les méthodes de délimitation d'espèces nécessitant souvent la construction d'un arbre phylogénétique, on peut considérer ces études comme produisant également des petites phylogénies, comprenant quelques espèces, à l'échelle d'un genre par exemple. Enfin, des phylogénies peuvent être tout simplement établies à l'échelle d'un genre, pour prouver ou non



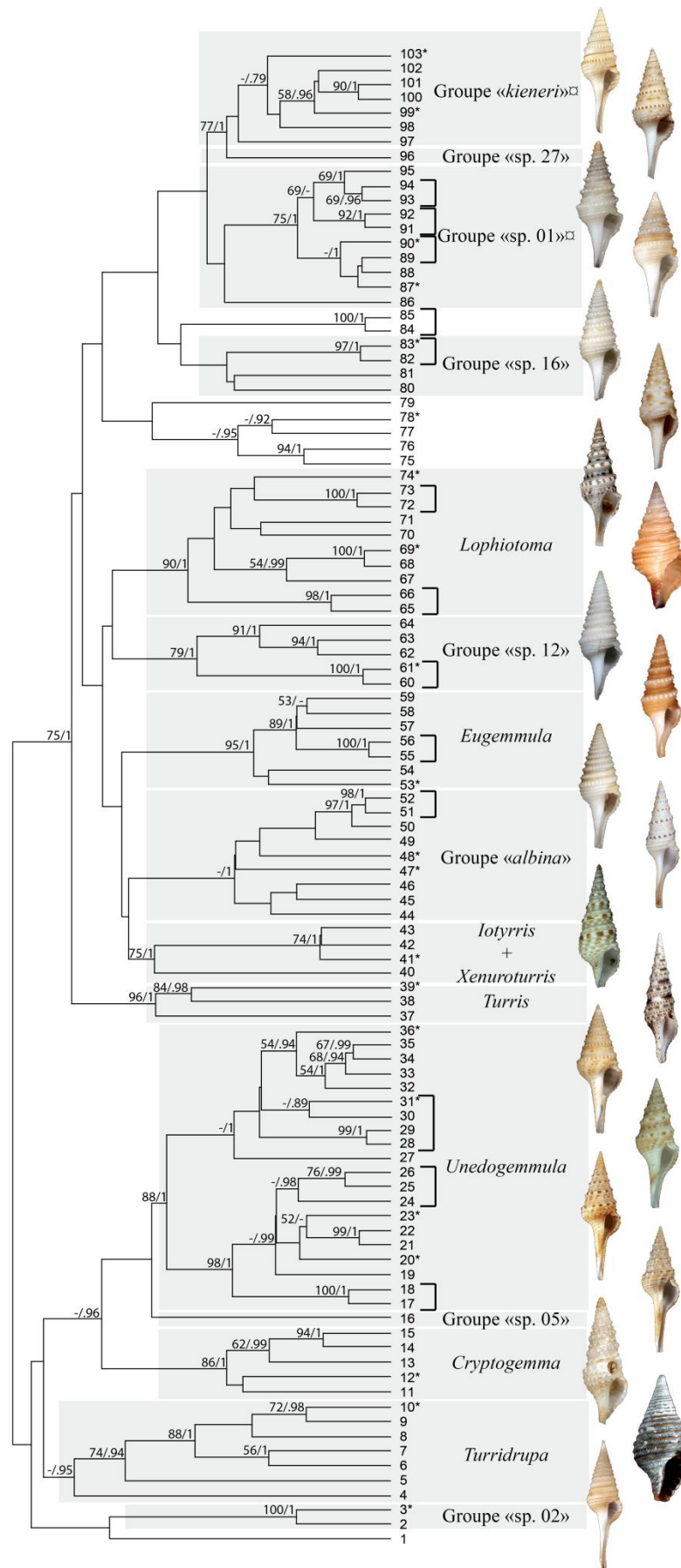


Figure 7 : Phylogénie des Turridae produite par Puillandre *et al.* (2012b), sur la base du fragment CO1. Les groupes nommés correspondent aux groupes trouvés dans le chapitre 3.

sa monophylie. Chez les Turridae nous citerons les phylogénies infra-familiales issues des travaux Olivera *et al.*, (2008) sur le genre *Gemmula*, Puillandre *et al.*, (2010) sur le genre *Gemmuloborsonia*, Fedosov *et al.*, (2011) sur le genre *Turris*, Todd & Rawlings (2014) sur le genre *Polystira*, Puillandre *et al.*, (2017 – voir Annexe 1) sur le genre *Lophiotoma*, Kantor *et al.*, 2008 et Abdelkrim *et al.*, (2018b – voir Annexe 3) sur les genres *Xenuroturris* et *Iotyrris*) et Kantor *et al.*, (2019) sur le genre *Lucerapex* (entre autres). A noter que l'étude de délimitation d'espèces d'Abdelkrim *et al.* (2018b – voir Annexe 3) est la seule publiée où des techniques de séquençage haut débit (HTS en anglais) ont été appliquées aux Turridae. J'ai pu, au cours de ma thèse, participer à certaines de ces études, en contribuant notamment à l'application de méthodes de délimitations d'espèces « traditionnelles » (Chapitre 1: Article 1 ; Annexe 1) et modernes (Annexe 3). D'autres applications de techniques HTS ont été utilisées dans cette étude, notamment pour établir une phylogénie robuste des Turridae, et seront présentées dans les chapitres 2 et 3.

### **3. Biologie**

Dans le chapitre 2, Les Turridae sont présentés comme un « taxon animal non-modèle », en définissant rapidement un taxon-non modèle comme un « taxon pour lequel peu ou pas de données génomiques et/ou transcriptomiques sont disponibles, et qui est n'est peu ou pas représenté dans les bases de données de loci orthologues ». Cette définition a été proposée dans un cadre strictement phylogénomique – en ayant parfaitement conscience qu'il ne s'agit probablement pas de la définition consensuelle de « taxon non-modèle », s'il en existe une. En réalité, le terme « taxon modèle » est dérivé du terme « organisme modèle », qui est plutôt utilisé pour parler d'une seule espèce (notons la confusion terminologique où le terme « organisme » désigne en fait une espèce) qui a été étudiée de manière approfondie à tous les niveaux de la biologie (voies métaboliques, régulateurs, développementales, comportementales, etc.), comme c'est le cas de la souris ou de la drosophile. On pourrait donc proposer une définition de « taxon modèle » plus large que celle utilisée dans le chapitre 2, par exemple en ajoutant qu'un taxon non-modèle est un taxon pour lequel il n'existe peu ou pas d'études sur les voies métaboliques, régulateurs, développementales, comportementales, etc. chez une ou plusieurs des espèces comprises dans ce taxon. Là encore, les Turridae correspondent à cette définition : en plus de n'avoir aucun génome de Turridae (ni même de Neogastropoda) complet assemblé et annoté, limitant la possibilité d'études d'expression de gènes ou d'évo-dévo moderne, il n'existe que très peu d'études sur la biologie des Turridae. Et bien que nous ayons à disposition d'innombrables descriptions de traits morpho-anatomiques, nous n'avons pas la moindre idée des processus à l'origine du développement et de la mise en place de ces traits, ni si ces traits sont homologues d'une espèce à une autre. Je vais toutefois essayer de faire une synthèse de ce que nous savons de la biologie des Turridae, en essayant d'identifier des caractéristiques qui permettraient de distinguer les Turridae d'autres familles de Conoidea.

#### **3.1 Morphologie**

Socle des hypothèses taxinomiques en malacologie pendant plusieurs siècles, l'étude de la morphologie reste indispensable pour identifier les coquilles vides ou les fossiles, c'est-à-dire l'immense majorité des coquilles retrouvées sur le terrain. Mais les Conoidea souffrent d'une absence de formalisation de cette morphologie, comme elle a pu être réalisée chez d'autres groupes de gastéropodes (par exemple Sanders *et al.*, 2017). L'étude poussée de la morphologie des Turridae ne fait pas partie des objectifs de cette thèse, les descriptions de la

coquille adulte (téléoconque) et de la coquille larvaire (protoconque) seront donc brèves et simplifiées.

### 3.1.1 Téléoconque

En synthétisant les définitions de Turridae, les Turridae sont caractérisés par la présence d'un sinus anal développé situé sur la périphérie du dernier tour de la coquille, laissant une bandelette périphérique sur les tours de la coquille (Fig. 8). Sur ces bandelettes on peut observer parfois la présence de gemmules, qui sont le produit de stries d'accroissement – ces dernières sont observables également sur des bandelettes sans gemmules. La coquille est fusiforme et peut présenter ou non un canal siphonal développé, et présente peu ou pas d'ornementation externe, ni de dents sur la columelle. Au niveau des patrons de coloration, la plupart des espèces des Turridae sont d'un teint uni blanc/gris clair, souvent avec une ligne orange/brune présente soit au niveau de la bandelette du sinus anal, soit au niveau de la ligne de suture, soit sur de multiples cordons. Cependant certaines espèces présentent un teint uni orange voire brun foncé (Planche C6). Il est également possible d'observer des patrons de coloration plus complexes, avec des coquilles tantôt tachetées, tantôt ponctuées de couleurs

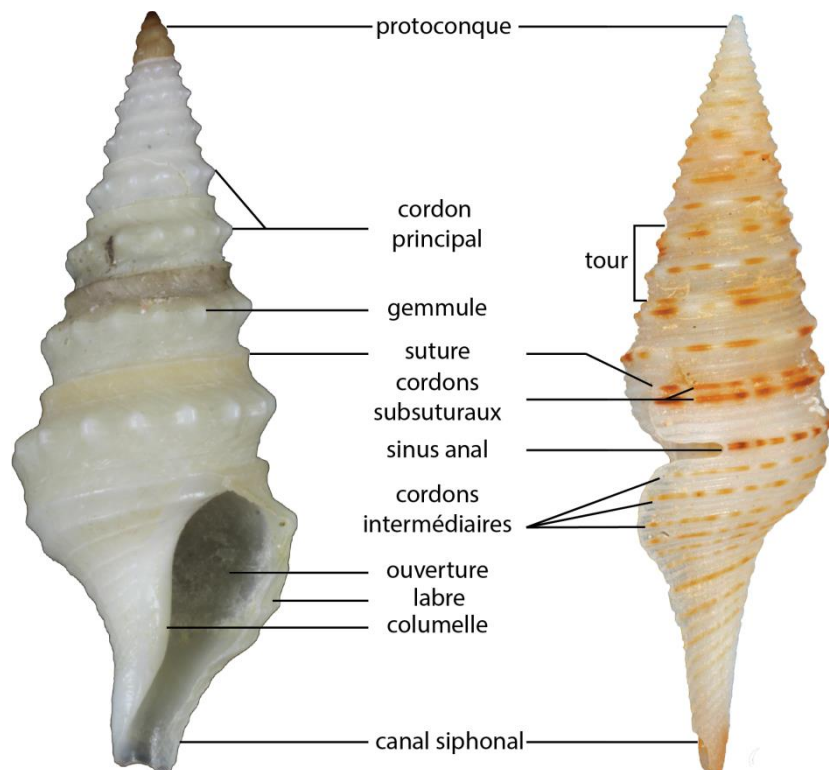


Figure 8 : Morphologie de la coquille. A gauche, spécimen de *Cryptogemma praesignis*. A droite, spécimen de *Gemmula hastula*.

oranges/brunes (Planche C7). Enfin, certaines espèces ont la particularité d'avoir un canal siphonal qui tend vers le rose (Planche C2H, C7H, C8K).

Aucun des traits décrits ci-dessus n'est exclusif à la famille Turridae. Le « sinus anal caractéristique », le plus souvent cité comme caractère diagnostique des Turridae, se retrouve aussi chez certaines espèces de Borsoniidae ou de Pseudomelatomidae. Les patrons de coloration sont classiques et se retrouvent chez tous les gastéropodes. C'est donc ailleurs que sur la coquille adulte qu'il faut chercher des caractéristiques propres aux Turridae.

### 3.1.2 Protoconque

Les traits de la protoconque peuvent servir à caractériser certaines familles de Conoidea : les protoconques avec une sculpture treillissée par exemple, ne se retrouvent que chez les Raphitomidae (Fedosov & Puillandre, 2012). Peu de taxonomistes se sont attardés sur une discussion générale sur la protoconque des Turridae ; la dernière réflexion à ce sujet en date est certainement celle de Kilburn (1983 :551-552) déjà citée plus haut. Kilburn fait état de trois types de protoconque (sans explicitement les nommer) qui suivent la séquence suivante : (i) « *Turris*-like », paucispirale (avec peu de tours) et lisse sauf sur le dernier tour où l'on peut observer des côtes axiales (ii) « *Lophiotoma*-like », avec plus de tours et des côtes axiales sur le dernier tour ou les deux derniers et (iii) « *Gemmula*-like », multispirale (« polygyrate ») et avec beaucoup de côtes axiales sur plus de 3 tours. A cela nous pourrions ajouter un quatrième type, « *Polystira*-like », paucispirale et complètement lisse, c'est-à-dire sans côtes axiales. En réalité, les mesures des tours de la protoconques et l'observation des côtes axiales dans un grand nombre d'espèces nous indique que les types de protoconques s'enchainent dans un continuum, allant de la protoconque paucispirale à des protoconques multispiralées pouvant atteindre plus de cinq tours (Fig. 9).



Figure 9 : Continuum des types de protoconque chez les Turridae. De gauche à droite, paucispirale (*Lucerapex*), multispirale à faible nombre de tours (*Lophiotoma*) et multispirale avec de nombreux tours (*Gemmula tesselata*). Le trait rouge marque la fin de la protoconque et donc de la vie larvaire.

En outre, Kilburn constate qu'il est incapable de savoir si ces différences sont d'origines phylogénétiques ou bien dues à des convergences liées à des adaptations écologiques ou des divergences géographiques. A ce jour, les mécanismes génétiques amenant un taxon à passer d'un type de protoconque à un autre sont toujours inconnus chez les mollusques. De plus, alors qu'il est admis que plus la protoconque est multispirale et plus la larve passe de temps à se nourrir dans la colonne d'eau (Powell, 1964-1967), et donc potentiellement plus elle est capable de disperser (étendant la répartition géographique de l'espèce), Kilburn constate que cette hypothèse ne se vérifie pas toujours. Face à ces constats, Kilburn décide d'utiliser une taxonomie opérationnelle (non-phylogénétique) des Turridae, en attendant « plus d'informations sur les systèmes d'organes ».

### 3.2 Anatomie

L'impossibilité d'établir des hypothèses d'homologies convaincantes – et donc de discuter d'innovations clefs et d'adaptation – sur la coquille des Turridae (et autres Conoidea) va conduire les systématiciens de la fin du 20<sup>ème</sup> à développer les études anatomiques, en particulier sur deux systèmes : la radula et le système digestif antérieur.

#### 3.2.1 Radula

La radula des gastéropodes est connue supposément depuis Aristote (Sollas, 1907), mais elle n'a été reconnue d'importance pour la classification qu'au milieu du 19<sup>ème</sup>, avec les travaux de Franz H. Troschel. Ce dernier est d'ailleurs le premier à illustrer une radula de Turridae dans son célèbre « Gebiss der Schnecken » (Troschel, 1866) (Fig. 10 ; cf. Planche R7A). Par la suite, des auteurs comme Thiele (1929), successeur de Troschel, Powell (1942, 1964-1967), Morrison (1966), Maes (1971) et McLean (1971) vont observer que les dents marginales des Turridae ont une structure en forme de furcula (« wishbone »), parfois également qualifiée de type « duplex » (Fig. 11; cf. Planche R5H).

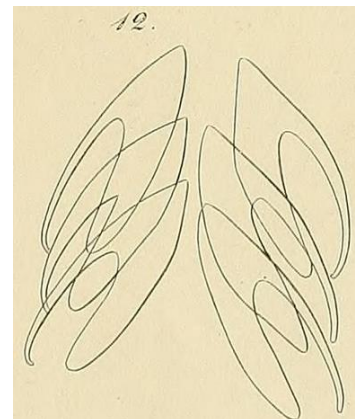


Figure 10 : Illustration de la radula de *T. babylonica* par Troschel (1866)

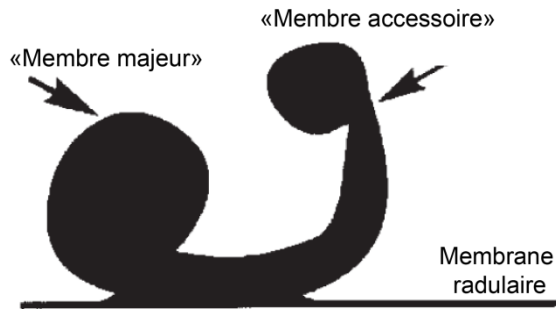


Figure 11 : Schéma d'une dent marginale de Turridae en coupe transversale. Modifié d'après Kantor & Taylor (2000)

Mais les observations au microscope optique sont limitées dans leur niveau de détail, et empêchent une représentation en trois dimensions des différents composants de la radula et de leur agencement entre eux. L'étude de la radula va prendre un nouvel essor dans les années 80-90, avec la démocratisation du microscope électronique à

balayage (MEB). Ce degré de précision supplémentaire va permettre à des auteurs comme Shimek & Kohn (1981) d'interpréter les radulas d'un point de vue fonctionnel, et dans la foulée de proposer des hypothèses d'apomorphie et de plésiomorphie sur les différents types de radulas. Les auteurs vont définir six grands types de radulas au sein des Conoidea, et les Turridae (même si l'étude ne comporte aucun Turridae *stricto sensu*) se voient attribuer le type « tranchant », en lien avec l'hypothèse que le développement des dents marginales et la réduction des dents centrales et latérales est une adaptation à la lacération des proies. Dans cette même étude, ce type de radula est considéré comme intermédiaire entre la radula « tranchante-râpante » primitive (à 5 dents) des Pseudomelatomidae et des Drillidae et la radula « hypodermique » des Conidae et Borsoniidae, « parfaitement adaptée » (d'après les auteurs) à l'injection de venin.

A partir de 1993, date de la publication de la première radula de Turridae prise au MEB (Taylor *et al.*, 1993), l'accumulation de photos haute définition de radulas de Turridae va s'accroître, notamment sous l'impulsion de malacologues russes, Alexandre Sysoev et Yuri Kantor. Ce dernier publie en 2006 (Kantor, 2006) une étude sur l'homologie de la dent centrale chez les Turridae, ayant pour objectif de proposer une hypothèse d'homologie entre les 5 dents (état plésiomorphe) des Drillidae et les 3 dents (état dérivé) des Turridae. En effet, entre les deux dents marginales, les Turridae possèdent soit une seule dent – plutôt importante (Planche R4A-D) ou plutôt réduite (Planche R3G-H) – soit aucune dent (Planche R3J-K). Kantor propose l'hypothèse que la dent centrale des Turridae est en réalité issue de la fusion de 3 dents (la centrale et les deux latérales). Cette hypothèse est appuyée par l'observation de radulas chez certaines espèces de *Turridrupa* (Planches R2A-B, E-F). Elle confirme ainsi la précédente hypothèse de Shimek & Kohn, comme l'étude de Kantor et Puillandre (2012) basée sur un arbre phylogénétique reconstruit à partir de données moléculaires : la réduction des dents centrales et latérales est un état dérivé. L'hypothèse est à nouveau corroborée par



Alexandre Fedosov (Fedosov *et al.*, 2011), qui montre que la radula du genre *Annulaturrhis*, groupe ayant divergé plus tôt dans la phylogénie des Turridae, possède une dent centrale importante alors que la radula du genre *Turrhis* est caractérisé entre autres par une réduction de la dent centrale.

Ces hypothèses sur la polarisation des états présence/réduction de la dent centrale seront testées dans le Chapitre 3. Il reste néanmoins que les études chez les Conoidea montrent que la radula des Turridae n'est pas caractéristique de la famille, et se retrouve chez les Clavatulidae et certains Pseudomelatomidae.

Encore une fois, donc, la radula fait partie de ces structures chez les Turridae qui ne permettent pas de caractériser formellement le groupe. Parallèlement à l'étude de la radula – toujours sous l'impulsion de l'école russe de malacologie – se sont développées des études sur les parties molles de l'anatomie.

### 3.2.2 Système digestif antérieur

Les Neogastropoda étant pour la plupart des prédateurs, il semble intuitif que l'étude du système digestif antérieur – zone utilisée pour la capture des proies – peut produire des informations pertinentes au niveau biologique, taxonomique et évolutif. Les Conoidea présentent une diversité particulièrement riche de ce système, et l'étude de celui-ci a rapidement conduit à des hypothèses sur son évolution (par exemple Sheridan *et al.*, 1973), mais seulement dans un cadre théorique néodarwinien. Ainsi, Medinskaya (2002a) évalue qu'environ une centaine de Turridae (*sensu lato*) ont déjà été étudiés, mais déplore malgré cela la confusion dans l'utilisation des données anatomiques à la fois pour la délimitation des familles et pour les relations entre celles-ci. De plus, sur la centaine de Turridae qu'évoque Medinskaya, seulement 6 schémas anatomiques publiés (Taylor, 1994 ; Medinskaya & Sysoev, 2001) correspondent à des Turridae *stricto sensu*, l'objet de notre étude.

Sur la base de deux spécimens de deux espèces de Turridae très proches (*Gemmula leucotropis* et *G. deshayesii*), Taylor (1994) établit que l'arrangement général du système

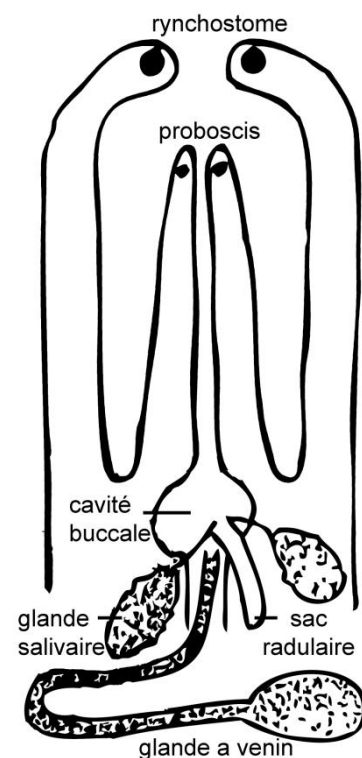


Figure 12 : Schéma du système digestif antérieur d'un Turridae. Modifié d'après Taylor (1994).



digestif antérieur des Turridae est similaire à celui des Drilliidae. La principale ressemblance étant la masse buccale située à la base du proboscis (Fig. 12), comme chez d'autres Neogastropoda, qui conduit Taylor à considérer cette configuration comme plésiomorphique. Posséder une masse buccale à la base du proboscis est de plus considéré comme une « contrainte fonctionnelle » par Taylor, car cette configuration met une distance entre la radula et l'extrémité du proboscis – lieu où la capture de la proie s'effectue. Deux principaux systèmes alternatifs permettant de contourner ce supposé handicap ont été observés : une masse buccale rapprochée de l'extrémité du proboscis (par exemple chez *Turricula*), ou bien le détachement d'une dent de la radula afin de la placer à l'extrémité du proboscis (par exemple chez les Conidae). Ce dernier *modus operandi* serait une synapomorphie des

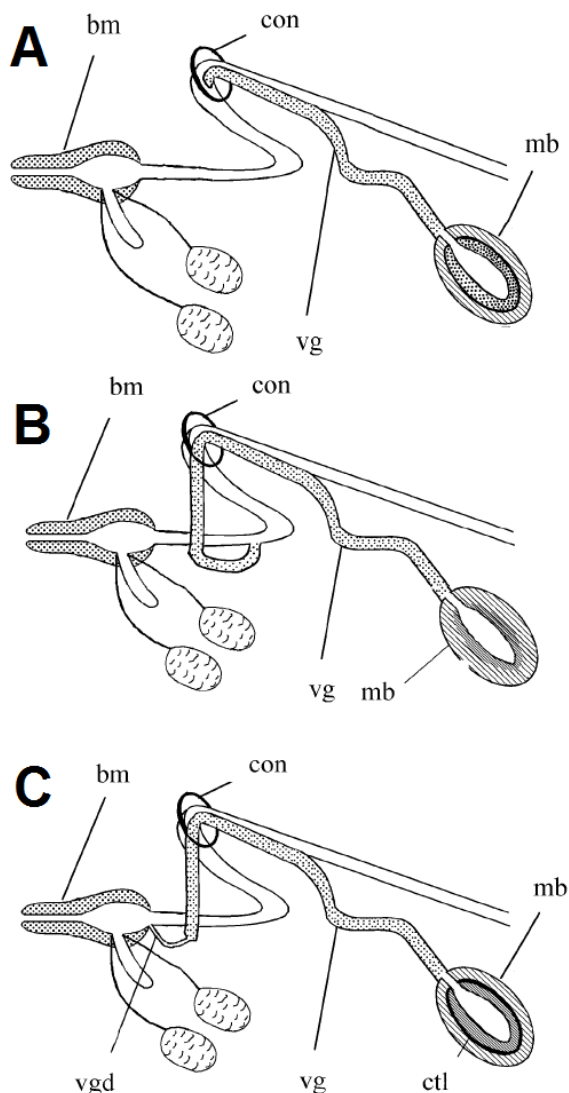


Figure 13 : Schémas de différents appareils venimeux, modifié d'après Medinskaya (2002). A. *Cochlespira*. B. *Lucerapex*. C. *Gemmula*. Bm – masse buccale, ctl, couche de tissus connecteurs, con – anneau nerveux, mb – bulbe musculéux, vg – glande à venin, vgd – canal de la glande à venin

Conoidea (Kantor & Puillandre, 2012), il y a donc fort à parier que les Turridae l'utilisent également, étant donné la position de leur masse buccale. En revanche, rien ne permet d'affirmer que tous les Turridae possèdent le même système digestif antérieur ; les Pseudomelatomidae (Taylor, 1994) par exemple possèdent plusieurs types de systèmes.

Les seules études sur la variation des types anatomiques intra-Turridae sont celles de Medinskaya sur le système digestif antérieur (Medinskaya, 2002a) et sur la glande à venin (Medinskaya, 2002b). Ces analyses incluent également quelques spécimens de Pseudomelatomidae et du groupe FHC (cf. Abdelkrim *et al.*, 2018a – voir Annexe 2). Quelques études anatomiques ont été publiées par la suite (Garcia-Lopez *et al.*, 2007 ; Santinabez *et al.*, 2013) mais sont principalement descriptives et ne s'inscrivent pas dans un cadre comparatif. L'analyse cladistique de Medinskaya réalisée à partir des caractères anatomiques ne permet de retrouver aucune des

trois familles ni des dix genres étudiés monophylétiques, contrairement aux résultats moléculaires. De fait il est impossible d'énoncer une seule synapomorphie des Turridae sur la base de cette analyse. Des différences entre espèces existent donc, mais les différents caractères proposés ne permettent pas de constituer des groupes considérés aujourd'hui comme monophylétiques. On notera tout de même que *Gemmuloborsonia jarrigei* possède une masse buccale située dans le proboscis et non à sa base. L'étude de l'appareil venimeux (Medinskaya, 2002b) offre peut-être plus de perspectives. Chez les Cochlespiridae, la glande à venin (qui en réalité à la forme d'un canal) se déverse dans l'œsophage au niveau de l'anneau nerveux (Fig. 13A). Chez la plupart des autres Conoidea – dont les Turridae – la glande à venin se connecte directement dans la masse buccale (Fig. 13C). Les Cochlespiridae étant groupe-frère de tous les autres Conoidea, il a été proposé que le déversement direct de la glande à venin dans la masse buccale est une adaptation fonctionnelle, augmentant l'efficacité de l'injection des toxines dans la proie (Abdelkrim *et al.*, 2018a – voir Annexe 2). Mais chez les genres *Gemmuloborsonia* et *Lucerapex*, la glande à venin se déverse dans l'œsophage à un niveau intermédiaire entre l'anneau nerveux et la masse buccale (Fig. 13B). Ces deux genres constituent un clade (Abdelkrim *et al.*, 2018a – voir Annexe 2) - si l'on y ajoute l'énigmatique '*Ptychosyrinx*' *chilensis* (cf. Chapitre 3) – qui aurait divergé des autres Turridae il y a environ 80 Ma. L'hypothèse que ce groupe possède un appareil venimeux intermédiaire entre celui des Cochlespiridae et des Turridae semble donc plausible. Une façon de corroborer cette hypothèse serait (i) d'observer si l'appareil venimeux de '*Ptychosyrinx*' *chilensis* possède la même structure (ii) observer si d'autres Turridae divergeant à la base de l'arbre (*Kuroshioturris*, *Gemmula hindsiana*, *G. closterion*) possèdent cette particularité ou non.

L'étude du système anatomique des Turridae, bien que proche de certains autres groupes de Conoidea, semble une voie prometteuse dans la recherche de synapomorphies et d'hypothèses adaptatives chez les Turridae. La variabilité du système digestif antérieur, même entre espèces proches, a été démontrée (Medinskaya & Sysoev, 2001 ; Medinskaya, 2002a), mais ces études manquaient d'un cadre phylogénétique pour interpréter correctement ces différences. L'anatomie comparée est une discipline en déclin, malgré son importance en systématique et en évolution (Boxshall & Self, 2011 ; De Carvalho *et al.*, 2014). Mais de nouvelles méthodes d'imagerie, non-invasives, permettant de scanner plusieurs spécimens à la fois en haute définition, pourraient permettre un renouveau de cette science et l'associer plus systématiquement aux études phylogénomiques.

### 3.3 Reproduction & Cycle de vie, Comportement & Alimentation

Cette partie est de loin la plus spéculative en ce qui concerne notre connaissance de la biologie des Turridae : aucun comportement reproducteur ou alimentaire en milieu naturel n'a jamais été publié. Toutes les considérations qui suivent découlent d'observations en aquarium ou d'inférences à partir de données morphologiques ou moléculaires.

#### 3.3.1 Reproduction

Comme pour la plupart des Caenogasteropoda, les individus sont sexués (mâles ou femelles) et la fertilisation est interne chez les Turridae. Un cas de dimorphisme sexuel sur la coquille a été documenté chez *Cryptogemma praesignis* (Kantor & Sysoev, 1991): les femelles matures possèdent un cordon supplémentaire qui permettrait le passage du pénis lors de la fécondation. Ce cas sera discuté dans le Chapitre 1. Mais en réalité ce phénomène pourrait s'appliquer à d'autres espèces de Turridae comme *Gemmula hombroni*, *G. graffei*, *G. diomedeae* (cf. Planche C4C), *G. sp. 20* (cf. Planche C8A), *Cryptogemma aethiopica*, ou bien même à d'autres Conoidea comme le Cochlespiridae *Aforia circinata* (Ronai & Shimek, 1984). Ce cordon supplémentaire apparaît de façon périodique sur la coquille (deux sont observables sur la coquille de *G. diomedeae*, Planche C4C), et ces périodes pourraient correspondre aux périodes de reproduction (s'il en existe). Dès lors, si le rythme de croissance d'un individu est connu – ce que l'on ne sait pas faire pour le moment – alors on pourrait éventuellement quantifier les périodes de reproduction sur plusieurs années et comparer les périodes de reproduction des différentes espèces citées plus haut.

#### 3.3.2 Cycle de vie

Lorsque la larve est planctotrophique, une partie de son développement est pélagique, c'est-à-dire qu'il s'effectue dans la colonne d'eau et constitue une partie du zooplancton, avant de se métamorphoser en juvénile, puis en adulte benthique. La période de vie pélagique de la larve planctotrophique reste un mystère pour quasiment toutes les espèces de Conoidea. Même si des périodes allant jusqu'à 50 jours de vie pélagique ont été observés chez l'espèce *Conus lividus* (Perron, 1981), il faut garder à l'esprit que ces observations ont été faites en aquarium et ne reproduisent probablement pas les conditions du milieu naturel. Il est donc possible que certaines larves puissent passer plusieurs mois à l'état planctotrophique, de quoi extrapoler des distributions d'adultes quasi-cosmopolites (cf. Chapitre 1). Lorsque la larve est dites « non-planctotrophique », le développement est soit partiellement planctotrophique, et

ne dure en comparaison que très peu de temps (24h chez *Conus pennaceus*), ne laissant – en théorie – pas le temps à la larve de disperser sur de grandes distances, soit à « développement direct », c'est-à-dire que la larve ne passe pas du tout de temps dans la colonne d'eau, mais ce cas n'a jamais été observé chez les Turridae. Sur les quelques études sur le développement des larves de cônes réalisées en aquarium (par exemple Leal *et al.*, 2017), seule l'étude de Perron (1981) rapporte une observation d'un cycle complet de Conoidea, de l'éclosion à l'âge adulte. Chez les Turridae la seule étude publiée est celle de Penchaszadeh (1982), effectuée sur 17 spécimens de l'espèce *Polystira albida* (mal-identifiée comme *P. barretti* ; cf. Todd & Rawlings, 2014), non-planctotrophique, placés en aquarium et ayant produits 48 capsules. La capsule, de forme lenticulaire, allant jusqu'à 10mm de diamètre, et presque complètement recouverte de grains de sables, inclue de 32 à 126 œufs. Un œuf fait environ 430 µm, puis se développe, et la larve véligère qui éclot possède une coquille d'environ 670 µm de diamètre. Des capsules similaires mais attachées à des spécimens vivants ont été observées entre autres chez *Polystira oxytropis* (Todd & Rawlings, 2014). Les différents auteurs s'accordent à dire que ce type de capsule est assez répandu chez les Conoidea.

### 3.3.3 Comportement prédateur

Encore une fois, il existe de nombreuses études du comportement de prédation chez les cônes, mais quasiment rien chez les Turridae ou chez d'autres Conoidea. Une des dernières découvertes majeures chez les cônes et celle de l'utilisation de venins différents, venant de différentes parties de la glande à venin, en fonction de leur utilisation en attaque (prédation) ou en défense (Dutertre *et al.*, 2014). Des observations de différentes stratégies de chasse complexes chez les Conidae (Olivera *et al.*, 2015) ou même chez les Terebridae (Miller, 1970) laissent supposer une diversité de comportements insoupçonnée chez les Conoidea. Chez les Turridae la seule observation directe de prédation, en aquarium, est publiée par Olivera *et al.* (2017) sur un spécimen de *Turridrupa* attaquant un ver polychète (Fig. 14). Pendant le processus



Figure 14 : Processus d'envenimation d'un ver polychète par un spécimen de *Turridrupa*, d'après Olivera *et al.*, (2014). ©Apple Academic Press

d'envenimation, les auteurs observent un changement abrupt dans la couleur de l'hémolymphe du ver, proposant la possibilité que le venin affecte avant tout le système circulatoire plutôt que le système nerveux comme chez la majorité des cônes. Olivera *et al.* (2014) font également référence à une observation d'un spécimen de *Turris* pour lequel le processus d'envenimation s'est déroulé alors que l'individu était enfoui sous le substrat : la présence d'un canal siphonal (voir Fig. 8) permettrait apparemment de manœuvrer efficacement le proboscis vers la proie (polychète) tout est restant caché sous le substrat.

### 3.3.4 Alimentation

Les résultats présentés ci-après sont issus du mémoire de stage de Claire Vincent (2019), portant sur une étude de métabarcoding des contenus stomacaux de spécimens issus de différents genres de Turridae. Même si quelques rares articles existent sur l'identification de proies (par exemple Taylor, 1980), il s'agit de la seule étude comparative de régimes alimentaires chez les Turridae. Cette approche a permis de séquencer le marqueur 16S de restes de polychètes issus de contenus stomacaux d'espèces de *Iotyrris/Xenuroturris*, *Lophiotoma*, *Turridrupa* et *Gemmula*, ainsi que des représentants d'autres familles de Conoidea. Les résultats, bien que limités par le nombre d'échantillons, semblent indiquer que les espèces d'un même genre partagent un régime alimentaire commun : les *Iotyrris/Xenuroturris* sont les seuls Turridae à se nourrir des polychètes Maldanidae et de Phyllodocidae par exemple, tandis que les spécimens de *Lophiotoma* semblent être les seuls à se nourrir de certains sous-groupes de polychètes Nereididae ou d'oligochètes. Le fait que des espèces de même genre, très proches phylogénétiquement (voir même espèces sœurs) et présentes en sympatrie consomment le même type de proie pose des questions sur les mécanismes de spéciations chez les Turridae. En effet, chez les Conoidea, le modèle de diversification régulièrement avancé est l'hypothèse de spéciation par différenciation de niche écologique : l'acquisition de nouvelles toxines au sein d'une population permettrait la prédation de proies nouvelles, amorçant un processus de spécialisation de niche écologique des populations et aboutissant à l'apparition de nouvelles espèces (Duda, 2008). Les résultats apportés par Claire Vincent semblent à première vue contredire cette hypothèse, remettant en question ce modèle. Evidemment, il est possible que les mécanismes à l'origine de la spéciation soient liés à des régimes alimentaires différents et que les espèces sœurs actuelles aient acquis un régime similaire secondairement, mais dans ce cas il n'existe – à ma connaissance – aucun moyen de tester cette hypothèse. L'auteure propose dans sa conclusion

que le conservatisme de niche (PNC pour *Phylogenetic Niche Conservatism* en anglais) – la tendance d’espèces proches à garder des traits morphologiques ou écologiques similaires – déjà avancée par Darwin (1859), pourrait expliquer pourquoi aucune différence n’est observée au sein des genres. Cependant le moteur de spéciation avancé par les théoriciens du PNC est principalement celui de la spéciation allopatrique, c’est-à-dire par la mise en place de barrières (principalement géographiques) à la reproduction. Or cette hypothèse a déjà été rejetée chez certains Turridae, avec l’argument que la plupart des espèces sont planctotrophiques, à large répartition, et donc potentiellement peu affectées par la mise en place de barrières géographiques (voir également Chapitre 1).

Même si le rejet de l’hypothèse écologique de diversification des Turridae par diversification de régime alimentaire se confirme dans les études à venir, cela ne discrédite pas pour autant un modèle de diversification écologique chez les Turridae. Comme présenté plus haut, les diversités de comportement (sexuels, larvaires, prédation) sont encore méconnues et peuvent être à l’origine de diversifications écologiques. Des évènements de polyploïdisation peuvent également être à l’origine de spéciation sympatriques. Il y a donc fort à parier que les mécanismes de spéciation chez les Conoidea soient multifactoriels. Enfin, l’hypothèse première de diversification écologique des Turridae (toxines/proies) est encore loin d’avoir été complètement rejetée : l’étude doit se poursuivre, notamment avec un échantillonnage accru des contenus stomacaux mais surtout avec la mise en relation des proies avec les cocktails venimeux injectés lors de la prédation, composés de toxines.

### 3.4 Toxines

La toxicité – parfois mortelle pour l’Homme – du venin des cônes et Conoidea est connue depuis plus de 300 ans. En 1705, Rumphius rapporte l’histoire d’une femme esclave sur les îles Banda (Indonésie) autour de l’an 1670 qui, ayant été piquée à la main par un cône, senti une douleur parcourant tout son corps et mourut rapidement sur le coup (Kohn, 2016). Kohn propose l’hypothèse qu’il s’agit de l’espèce *Conus geographus*, qui possède (d’après nos connaissances actuelles) le venin le plus mortel du genre *Conus*, et par extension des Conoidea. Bien que l’anatomie et le mode d’injection du venin des cônes soient connus depuis longtemps, les premières études sur les effets du venin ont commencé dans les années 60 (Kohn *et al.*, 1960) et les premières isolations et caractérisations des toxines présentes dans le venin qu’à partir de la fin des années 70 (Cruz *et al.*, 1978). Pour les autres Conoidea, et en

particulier les Turridae, il faut attendre les années 2000 pour voir les premières études de toxinologie.

### 3.4.1 Les premières « turrtoxines »

La majorité des études sur les toxines de Turridae ont été réalisées par (ou en collaboration avec) Baldomero Olivera (et son équipe), également pionnier dans l'étude des toxines chez les cônes (Tbl. 1). Comme pour les cônes, la motivation principale de ces études est la recherche de composés bioactifs, potentiellement d'intérêt pharmacologique, comme c'est le cas du ziconotide (Prialt®, analgésique 1000 fois plus puissant que la morphine), version synthétique du peptide  $\omega$ -conotoxin MVIIA trouvé dans le venin de *Conus magus*. Profitons du nom complet de ce dernier peptide pour présenter rapidement le système de classification des toxines (Mary, 2018):

- (i) La lettre grecque identifie l'une des familles de cible moléculaire (enzyme, canal ionique, etc.).
- (ii) La lettre latine identifie l'une des superfamilles de gènes (28 sont référencées sur la base de données [www.conoserevr.org](http://www.conoserevr.org)), sur la base de la similarité de leur séquence signal.
- (iii) Les chiffres romains identifient l'un des patrons (« framework ») de cystéines conservés dans la protéine mature.

Ce système a d'abord été basé sur les toxines trouvées chez les cônes, baptisées « conotoxines » (Cruz *et al.*, 1978), puis étendu aux « turrtoxines » pour les Turridae (López-Vera *et al.*, 2004) et aux « teretoxines » pour les Terebridae (Puillandre & Holford, 2010; parfois aussi trouvées sous le nom « augertoxins » dans la littérature). Cependant cette distinction a été remise en question et il a été proposé d'utiliser le terme « conotoxine » pour tous les Conoidea (Heralde III *et al.*, 2008), mais cette idée ne semble pas avoir été suivie (d'après la littérature récente).

La stratégie « traditionnelle » de recherche de composé bioactifs (d'après Puillandre & Holford, 2010) consiste à :

- (i) Sélectionner une ou quelques espèces facilement échantillonnables.
- (ii) Effectuer un fractionnement de leur venin à l'aide de la méthode de chromatographie en phase liquide à haute performance (HPLC en anglais).
- (iii) Identifier les composants bioactifs d'intérêt.

- (iv) Séquencer les peptides correspondants.
- (v) Synthétiser chimiquement les peptides et réaliser des études fonctionnelles.

Toutefois, la plupart des études ne comprennent qu'une partie des étapes ; chez les Turridae, très peu d'études fonctionnelles ou de recherche de cible moléculaire ont été réalisées. Dans certaines études, l'étape (ii) est directement remplacée par une recherche des peptides dans l'ADNc (Watkins *et al.*, 2006). Enfin, ce n'est que récemment que des approches transcriptomiques – donc à plus grande échelle – ont été développées pour identifier des toxines (Gonzales & Saloma, 2014). Les principaux résultats des études de toxinologie chez les Turridae ont été synthétisés dans le tableau ci-après (Tableau 1) :

Tableau 1 : Tableau récapitulatif des études toxinologiques chez les Turridae

Publication	Espèce(s) de Turridae	Superfamille(s) de toxines	Résultats principaux
López-Vera <i>et al.</i> (2004)	<i>C. periscelida</i> <i>P. albida</i>	<i>n.a.</i>	Peptides PaIAa et GpIAa, riches en méthionine. Structure des peptides en hélice $\alpha$ .
Watkins <i>et al.</i> (2006)	<i>I. olangoensis</i>	“I-like”, “O-like”, P	Peptides plus grands (80-110 acides aminés) que chez les cônes (10-30 a.a.)
Heralde III <i>et al.</i> (2008)	<i>G. speciosa</i>	« Pg »	Superfamille « Pg ». Modifications post-transcriptionnelles importantes.
Aguilar <i>et al.</i> (2009)	<i>P. albida</i>	« P-like »	Turrtoxine « pal9a » – appartient aux « peptides spasmodiques ».
Arumugam <i>et al.</i> (2013)	<i>U. indica</i>	<i>n.a.</i>	Activité hémolytique, caséinolytique et gélatinolytique du venin.
Gonzales & Saloma (2014)	<i>G. speciosa</i> <i>U. bisaya</i>	B, D, I, L, M, O, P, Y	Nouvelles superfamilles. Peptides jusqu'à 190 a.a. Turrtoxines pauvres en cystéines.
Omega <i>et al.</i> (2017)	<i>U. bisaya</i>	M	Turrtoxine « ubi3a » – activité excitante sur le ganglion spinal chez la souris.

L'identification des toxines en suivant une terminologie faisant référence aux familles de Conoidea part d'une bonne intention, celle de garder un cadre phylogénétique pour la désignation des toxines. Cependant, dans la pratique, il y a des limites à cette classification. Tout d'abord, les équipes étudiant les turrtoxines utilisent toujours le terme Turridae pour se référer à tous les Conoidea qui ne sont ni des cônes ni des térébres. Pour rester cohérent il



faudrait créer une catégorie de toxine pour chacune des 17 familles de Conoidea (par exemple « bouchetispiritoxines » pour les Bouchetispiridae). Ensuite, comme il est présenté dans le Tableau 1, les auteurs font souvent référence à la classification des conotoxines (par exemple « P-like ») au lieu de désigner une nouvelle superfamille, propre aux turrtoxines, lors de la découverte de nouvelles toxines de Turridae ayant des propriétés similaires avec les toxines de Conidae. Une confusion s'installe : les « turrtoxines-like » sont-elles proches des conotoxines par convergence ou s'agit-il réellement d'homologues ? S'il s'agit réellement d'homologues, y-a-t-il un sens à continuer de distinguer conotoxines et turrtoxines ? De façon plus générale, la classification des toxines de Conoidea en superfamilles se base sur la séquence signal, partie de la toxine la plus conservée phylogénétiquement. Mais, même conservée, cette séquence exprime des niveaux de divergence très élevés, même au sein des superfamilles de toxines, et même au sein du genre *Conus* (Puillandre *et al.*, 2012c; Gorson, 2018). De plus, la taille limitée de ces séquences signal (10-20 acides aminés) fournit un nombre de caractères limités pour des analyses phénétiques ou phylogénétiques. Face à ces obstacles, confrontés aux dizaines de millions de toxines potentielles estimées par Puillandre *et al.* (2012c), est-il réaliste d'envisager un jour d'avoir une phylogénie résolue et robuste des toxines de Conoidea ? Une solution consiste peut-être à créer une classification plus arbitraire, fondée sur des caractères diagnostiques, à la manière d'une clef d'identification, pourquoi pas en réseaux. Les analyses phylogénétiques pourraient être restreintes à des groupes au sein de superfamilles pour continuer d'étudier, par exemple, les taux d'évolutions des différents groupes de toxines.

Si la promesse d'une classification stable des toxines semble encore éloignée, il reste un obstacle que je n'ai pas encore évoqué, mais qui est peut-être en passe d'être surmonté, comme détaillé dans la partie suivante.

### 3.4.2 La recherche de nouvelles toxines

Les résultats de cette partie sont tous issus du mémoire de stage de Mary (2018) effectué pendant ma thèse et avec laquelle j'ai pu échanger sur l'assemblage et l'analyse de transcriptomes. Ses travaux ont résulté en un pipeline bio-informatique original qui est actuellement finalisé en vue d'une publication.

Parmi les obsessions des scientifiques, il y a la découverte d'entités nouvelles, et si possible que l'on n'attendait pas. Mais comment découvrir sans savoir ce que l'on cherche ? Dans le cas de la recherche de toxines, comment trouver des toxines totalement nouvelles, qui

présentent par exemple des séquences signal inattendues ou bien des patrons de cystéine peu communs ? Jusqu'à présent l'approche traditionnelle consistait à utiliser des données (toxines) de référence dans une base de données, puis de chercher des peptides similaires lors de l'étude d'un nouveau venin. Cependant les approches de recherche par similarité sont limitées et en particulier pour les toxines puisque, comme nous l'avons vu plus haut, les séquences de toxines peuvent être très divergentes, et donc potentiellement non-déTECTABLES. Lou Mary a proposé une alternative à cette approche traditionnelle : une approche structurale, basée sur les propriétés-clefs conservées des peptides. Ainsi, au lieu chercher un peptide similaire à une autre, on va chercher un peptide qui possède les propriétés régulières qui nous intéressent, en accordant un certain degré de liberté. Par exemple, la présence d'au moins quatre cystéines dans la séquence signal est un bon indicateur pour chercher des toxines. Même si cela ne suffit pas, on aura accès à une gamme de toxines potentielles beaucoup plus importante, que l'on peut à la suite affiner avec d'autres critères tels que le nombre de cystéines paires.

Je ne présente ici en guide de conclusion qu'une partie des résultats – ceux pertinents pour cette thèse – sur les toxines trouvées à partir des 14 transcriptomes de glandes (à venin ou salivaires) de Turridae (Fig. 15). Le pipeline de Lou Mary a permis de retrouver 35103 toxines potentielles contre 2376 pour le pipeline traditionnel, soit près de 15 fois plus de toxines. Et il ne s'agit ici que de résultats fondés sur une analyse utilisant des valeurs de paramètres stringents. Avec des paramètres relâchés, les nombres de toxines potentielles pourrait bien être des dizaines de fois plus élevé en préférant le pipeline structurel plutôt que le traditionnel. Bien évidemment, ces résultats sont à relativiser, en gardant en tête que seule une véritable analyse fonctionnelle permet de décider si oui ou non tel peptide est bien une toxine. Cependant, ce

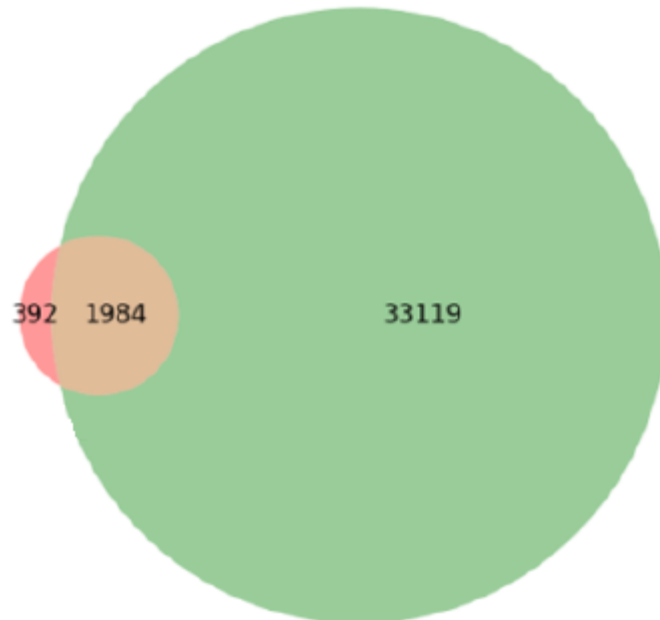


Figure 15 : Résultat du pipeline sur les transcriptomes de glandes à venin de Turridae. En rouge, pipeline traditionnel, en vert le nouveau pipeline. D'après Mary (2018)

pipeline ouvre une voie prometteuse dans la recherche des toxines, et plus largement dans la recherche de composants bioactifs. Les nouveaux défis dans la recherche de toxines se situent maintenant dans l'automatisation des tests fonctionnels de toxines potentielles par les

biochimistes. Pour le phylogénéticien, l'étude d'un panel de toxines potentielles plus vaste permettrait d'évaluer plus précisément le nombre d'évènements de duplication et de rendre compte des dynamiques évolutives aboutissant à la création d'entités nouvelles.

Pour des questions de temps, il n'a pas été possible dans le cadre de cette thèse d'étendre l'étude de la diversité des toxines produites par les Turridae à plus de spécimens que celles présentées ci-dessus, ce qui m'aurait notamment permis de tester l'impact potentiel de la diversification des turrtoxines sur la diversification des Turridae (voir Chapitre 3). Cependant, dans le cadre de l'analyse des transcriptomes de Turridae présentée dans le chapitre 2, j'ai eu l'occasion au cours de la thèse de participer à l'assemblage de transcriptomes d'autres Conoidea, notamment dans le cadre du stage de Master 2 de Lou Mary et dans le cadre du séjour post-doctoral de Giulia Fassio (voir partie 2 du Chapitre 2). En plus du pipeline d'analyse des transcriptomes de Conoidea déjà mentionné, ma participation à l'analyse de ces transcriptomes a été valorisée par la publication d'un article (Fassio *et al.*, 2019) dont je suis co-auteur.

### 3.5 Géographie, Bathymétrie & Habitat

Fondamentale en systématique, l'attribution d'aires géographiques aux espèces sert de point de départ à nombre de disciplines, que ce soit en taxonomie pour relier des spécimens-types à des unités taxinomiques opérationnelles, en écologie pour l'étude des stations et des successions écologiques, en phylogénie pour la reconstruction d'aires géographiques ancestrales, ou en biogéographie pour la reconstruction d'aires biogéographiques ancestrales (Cecca & Zaragüeta, 2015). La distribution d'espèces constituant le point de départ aux études micro- et macro-évolutives, sa caractérisation ne doit pas être négligée. Non moins importants mais bien moins étudiés, les distributions bathymétriques et la caractérisation d'habitats devraient également être intégrées dans les facteurs explicatifs des différents processus de diversification.

#### 3.5.1 Géographie

Exercice difficile que de caractériser l'aire de répartition d'une espèce, encore plus peut-être chez les Turridae. Difficile conceptuellement, parce-que nous sommes toujours en droit de nous demander, par exemple, si l'aire de répartition d'une espèce concerne seulement les adultes ou bien des représentants vivants à tous les stades ontogénétiques. L'alternative impliquerait d'identifier des Turridae dans les populations planctoniques de par le monde via

des approches, par exemple, de métabarcoding. Un article de Su *et al.* (2018), dont les Turridae ne constituent pas l'objet premier de l'étude, fait état d'un spécimen d'*Unedogemmula undeo* retrouvé vraisemblablement à l'état larvaire planctonique dans le contenu stomacal d'un maquereau (*Scomberomorus commerson*) en Décembre 2013 dans la baie de Zhanjiang en mer de Chine méridionale. La quasi-absence de données sur les larves de Turridae et l'ampleur de la tâche justifient de ne pas s'attarder sur la question de la prise en compte des larves dans l'attribution d'une aire géographique à une espèce. Pourtant, si l'on observe les schémas d'aires de distribution proposés par des taxinomistes, par exemple sur une espèce exclusive aux milieux côtiers mais présentes dans toutes les îles de l'Indo-Pacifique, le taxinomiste ne s'embarrasse pas de dessiner l'aire de distribution de façon parcellaire (par exemple en surlignant le contour des îles) mais va simplement considérer la région Indo-Pacifique comme l'aire de distribution de cette espèce (Powell, 1964-1967). Ainsi, la majorité de la surface géographique représentée dans l'aire de distribution ne contient pas de représentants de l'espèces (même en considérant les larves ; ces dernières suivent probablement des courants bien précis). Cette considération un peu naïve a son importance : dans le cas de reconstruction d'aires ancestrales ou d'études biogéographiques, ce sont les aires de distribution qui sont considérées et non les aires effectives de vie des spécimens. Et des espèces considérées comme sympatriques pourraient en réalité ne jamais se retrouver aux mêmes endroits, car distribuées à différentes profondeurs ou dans différents habitats. Or la sympatrie est souvent la condition initiale requise pour discuter de diversification écologique (Coyne & Orr, 2004; voir aussi partie 3.3). Ces considérations ont conduit certains auteurs à proposer d'abandonner le concept de sympatrie (Fitzpatrick *et al.*, 2008).

Les Turridae sont présents dans toute la zone intertropicale, mis à part sur la côte Ouest-Africaine. Le MNHN dispose d'une collection inégalée, rassemblant des dizaines de milliers de spécimens – pour la plupart des coquilles vides – collectés lors de missions d'exploration distribuées dans le monde entier. Cependant, la non-fiabilité de la morphologie pour la délimitation d'espèce (voir parties 2 et 3.1) nous oblige à n'utiliser que les spécimens séquencés pour caractériser les aires de chaque espèce. Ces spécimens, bien que nombreux, ne sont pas suffisants pour une telle entreprise. Il est fort probable que la plupart des espèces sœurs planctotrophiques aient une distribution sympatrique, que ce soit en Atlantique-Ouest ou en Pacifique-Est pour les espèces du genre *Polystira*, ou bien dans l'Indo-Pacifique pour la quasi-totalité des autres genres de Turridae (voir aussi Chapitre 1). Cependant, beaucoup de ces espèces ne sont représentées que par un ou quelques spécimens, souvent échantillonnés au

même endroit, résultant en une distribution allopatrique erronée avec leurs espèces sœurs. L'accumulation de nouveaux spécimens campagne après campagne ne cesse de redéfinir les aires de distribution d'espèces que l'on croyait confinées à une zone limitée. Il est ainsi possible de passer d'une distribution confinée à la Nouvelle-Calédonie à une zone comprenant la Nouvelle-Calédonie + le Nord de Madagascar en augmentant l'effort d'échantillonnage. De plus, il faut compter trois autres biais : (i) les campagnes ne sont pas réparties uniformément dans la zone intertropicale, (ii) les spécimens d'espèces plus petites (plus difficiles à collecter et à séquencer) sont souvent sous-représentés (iii) les spécimens d'espèces de milieux profonds sont souvent moins bien conservés pour des analyses moléculaires. Face à ces obstacles, il est encore trop tôt pour entreprendre des analyses (bio)géographiques sur les Turridae. Et dans tous les cas, comme cela a été évoqué précédemment, d'autres paramètres extrinsèques peuvent également être un moteur de diversification chez les Turridae.

### 3.5.2 Bathymétrie

Tout comme les données géographiques, les données bathymétriques ont leur importance en systématique, mais ont surtout été utilisées pour répondre à des questions d'ordre écologiques, comme par exemple sur la dispersion verticale des larves d'espèces abyssales, ou sur les liens entre la profondeur des espèces et la planctotrophie (Bouchet & Waren, 1979). Chez les Conoidea, il a été proposé que les espèces « toxoglosses » (possédant une radula hypodermique) sont plus représentées dans les grandes profondeurs que les espèces non-toxoglosses (Sysoev, 1991).

Même s'il existe un continuum dans la distribution des espèces en profondeur, il peut exister des différences bien marquées entre espèces. Certaines espèces sont strictement littorale (< 25 m) et ne se retrouvent donc jamais lors de dragages dans des milieux profonds. Contrairement à la géographie donc, il existe un fractionnement plus important au niveau des distributions bathymétriques chez les Turridae. De plus, l'ajout de nouveaux échantillons ne modifie pas en profondeur les moyennes bathymétriques propres à chaque espèce. Les campagnes Tropical Deep-Sea Benthos, échantillonnant régulièrement entre 100 et 1000m de profondeur, ont mis en évidence une diversité insoupçonnée de Turridae (et des Conoidea) dans cette tranche de profondeur (Bouchet *et al.*, 2009), à tel point qu'il y a plus de MOTUs de Turridae situés dans des profondeurs supérieures à 100m que dans le milieu côtier. Enfin, deux paires d'espèces sœurs très proches génétiquement distribuées en zonation bathymétrique ont été identifiés : celui de *Lophiotoma abbreviata* et *L. brevicaudata*

(Puillandre *et al.*, 2017 – voir Annexe 1) et celui de *Cryptogemma praesignis* et *C. phymatias* (voir Chapitre 1).

### 3.5.3 Habitat

Certes, des spécimens ont pu être photographiés vivant sur leur habitat naturel (Fig. 16), mais il ne s'agit souvent que d'une seule zone géographique explorée par un plongeur amateur de photographies naturalistes, qui n'a pas vocation à caractériser tous les habitats d'une espèce en particulier, à l'échelle de son aire de répartition. Plus important :

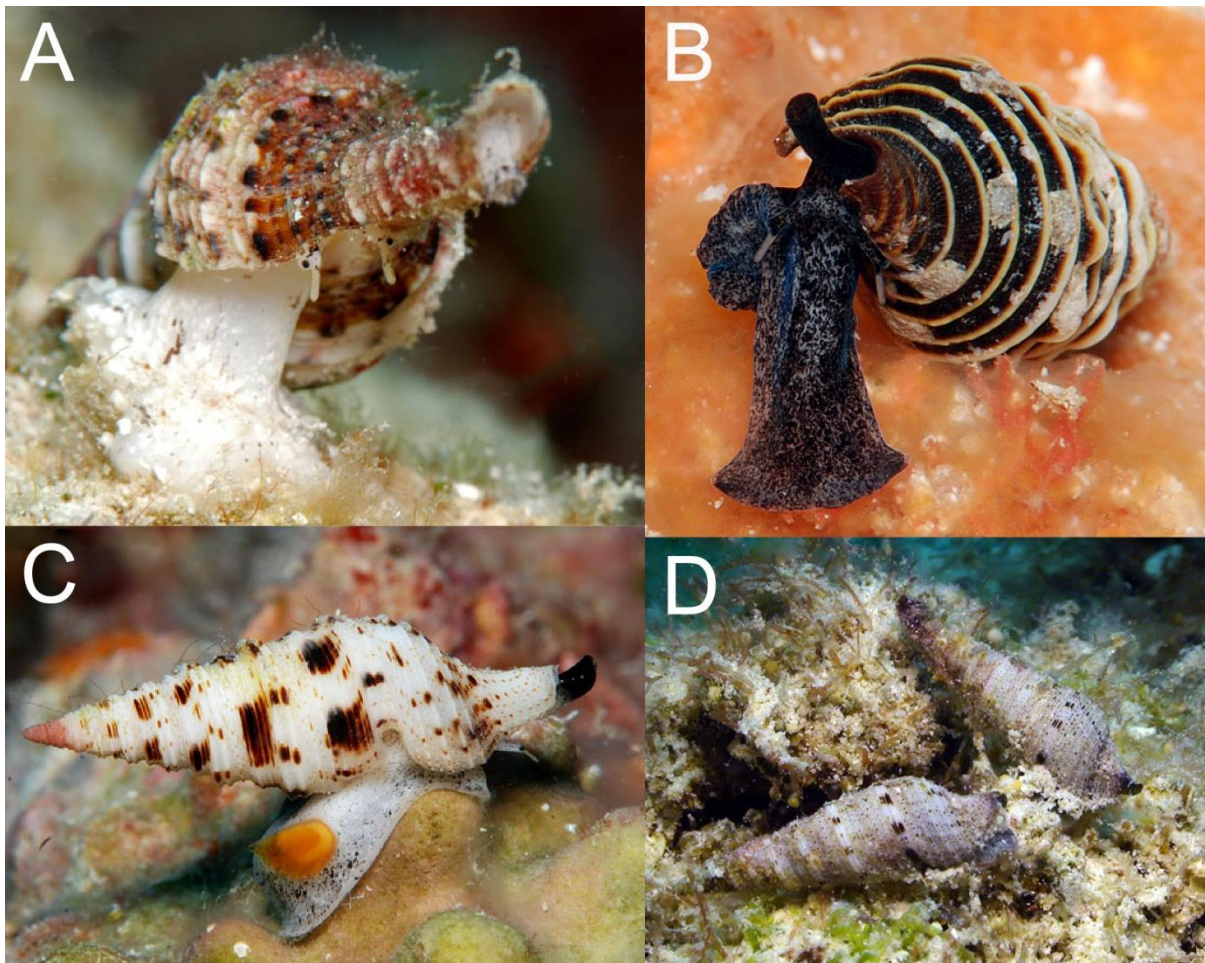


Figure 16 : Photographies de spécimens vivants de Turridae. A. *Turris* B. *Turridrupa* C. *Lophiotoma* D. *lotyrris*.  
©In-Depth Images Kwajalein.

l'identification des espèces – déjà difficile à partir d'une étude morfo-anatomique poussée – peut s'avérer impossible à partir d'une simple photographie. En outre, il est supposé que la plupart des espèces de Turridae (comme de Conoidea) s'enfouissent dans le sédiment. Enfin, comme énoncé plus haut, beaucoup d'espèces de Turridae vivent à des profondeurs inférieures à 100 mètres, difficiles d'accès (voire inaccessibles) à des plongeurs, et difficilement photographiables par des robots sous-marins comme un ROV car trop petites et/ou enfouies. Dès lors, la caractérisation de l'habitat d'une espèce de Turridae ne peut

s'appuyer que sur les données de dragage/chalutage et la combinaison éventuelle de celles-ci avec des photographies de robots sous-marins. Cela fait quelques années seulement que les retours de dragues/chaluts des campagnes TDSB sont systématiquement pris en photos, permettant de déterminer la nature des fonds. Mais la plupart des stations de campagnes TDSB sont soit accompagnées au mieux de quelques notes sur le type de milieu, au pire d'aucune information si ce n'est la profondeur et les coordonnées GPS. Les informations sur les différents milieux sur lesquels ont été échantillonnés les Turridae sont donc encore moins complètes que pour les données géographiques. Toutefois, il est possible de spéculer sur la spécialisation possible de quelques espèces pour lesquels les spécimens n'ont été retrouvés que sur du sable foncé ou de la vase comme les *Gemmula cf. hombroni* ou *Lophiotoma polyropa*, ou encore des espèces de *Turridrupa* trouvées systématiquement sur des fonds plutôt rocheux. Mais ces spéculations ne concernent que des espèces pour lesquelles peu de spécimens ont été échantillonnés. Dans le cas contraire, comme chez les espèces de *Cryptogemma* par exemple, les espèces peuvent être qualifiées de généralistes. Un cas extrême de spécimens de *Gemmula tessellata* retrouvés dans des environnements drastiquement différents est présenté dans le Chapitre 1.

Même s'ils ne possèdent pas le même attrait que les cônes ou les porcelaines, les Turridae sont abondamment pêchés dans certaines parties du monde pour la vente de coquilles ou bien en vue d'études toxinologiques. Une étude sur la double barrière de corail dans la mer de Visayan, aux Philippines a été réalisée entre 2008 et 2010 afin d'évaluer l'impact des « turrid fisheries » (Seronay *et al.*, 2014). La densité des Turridae *s.l* est évaluée à 217 individus/hectare, et parmi les 6 familles échantillonnées, les Turridae *s.s.* y sont les plus représentés, tant en diversité (46%) qu'en abondance (54%). L'étude estime un déclin dans les taux de capture entre 2008 et 2010 et propose un arrêt périodique des opérations de pêches. Actuellement, aucun Conoidea exceptés quelques cônes (Peters *et al.*, 2013) n'est sur la liste rouge mondiale des espèces menacées. Une autre étude similaire (Serronay *et al.*, 2016) rapporte que l'abondance de spécimens représentant de l'espèce *Turridrupa bijubata* varie selon les cycles de mousson. Toutefois ces études ne se basent que sur des reconnaissances morphologiques d'espèces, alors que l'on sait que plusieurs des espèces échantillonnées présentent une diversité cryptique (par exemple *Lophiotoma acuta* ; Puillandre *et al.*, 2017 - voir Annexe 1).

#### **4. Conclusion**

Après plusieurs siècles de confusion sur la classification des Turridae *sensu lato* puis, ces 20 dernières années, des avancées régulières sur l'étude des Turridae *sensu stricto* la classification par genres des Turridae semble petit à petit se stabiliser. Cependant, la taxonomie spécifique et générique, ainsi que la phylogénie du groupe, restent instables, comme le montre l'étude de Puillandre *et al.* (2012b). Cette instabilité peut avoir des conséquences sur notre connaissance de la biologie (moléculaire, physiologique) et l'écologie des Turridae, encore largement méconnue. Le point de départ de cette thèse va donc consister à appliquer une stratégie de taxonomie intégrative, déjà utilisée avec succès (Annexes 1 et 3) dans ce groupe. Cette stratégie va permettre de s'assurer de l'établissement d'hypothèses d'espèces robustes chez les Turridae, comme présenté dans le chapitre suivant.

---

#### **5. Planches**

Liste des abréviations utilisées dans les planches :

MNHN: Muséum national d'Histoire naturelle

SMF: Forschungsinstitut und Natur-Museum Senckenberg, Senckenberg-Anlage 25, 6000 Frankfurt-am-Main 1, Germany.

NMSA: KwaZulu-Natal Museum



## PLANCHE C1

### *Lucerapex*

- A. *Lucerapex* sp., MNHN-IM-2013-19989, PAPUA NIUGINI (CP4012)
- B. *Lucerapex laevicarinatus* MNHN-IM-2013-52063, EXBODI (CP3934)
- C. *Lucerapex cracens*, MNHN-IM-2013-19988, PAPUA NIUGINI (CP4012)
- D. *Lucerapex cracens*, MNHN-IM-2013-52064, BIOPAPUA (CP3738)

### *Gemmuloborsonia*

- E. *Gemmuloborsonia colorata*, MNHN-IM-2013-62922, BIOMAGLO (DW4828)
- F. *Gemmuloborsonia clandestina*, MNHN-IM-2007-41937, AURORA 2007 (CP2658)
- G. *Gemmuloborsonia clandestina*, MNHN-IM-2007-41954, SALOMONBOA 3 (CP2857)
- H. *Gemmuloborsonia clandestina*, MNHN-IM-2013-68044, KANACONO (DW4661)

### Groupe “*chilensis*”

- I. “*Ptychosryinx*” *chilensis*, MNHN-IM-2013-43303, Hors Campagne (Chili)

### *Gemmula*

- J. *Gemmula hindsiana*, holotype NHMUK 1879.2.26.34, Golfe de Magdalena, California

### *Kuroshioturris*

- K. *Kuroshioturris nipponicus*, MNHN-IM-2009-19091, AURORA 2007 (CP2684)

### Groupe « *closterion* »

- L. *Gemmula* sp., MNHN-IM-2013-49903, Hors Campagne (Japon)
- M. *Gemmula closterion*, MNHN-IM-2007-40789, NORFOLK 2 (DW2156)



## PLANCHE R1

### *Gemmuloborsonia*

- A. *Gemmuloborsonia* aff. *jariggei*, MUSORSTORM 8 (DW1061)
- B. *Gemmuloborsonia clandestina*, MNHN-IM-2007-41943, AURORA 2007 (CP2658)
- C. *Gemmuloborsonia laeva*, MNHN-IM-2009-29322, TERRASSES (CP3089)
- D. *Gemmuloborsonia neocaledonica*, MNHN-IM-2009-17163, BIOPAPUA (DW3748)
- E. *Gemmuloborsonia incognita*, MNHN-IM-2009-14919, ATIMO VATAE (CP3585)

### *Lucerapex*

- F. *Lucerapex cracens*, MNHN-IM-2009-13549, CONCALIS (CP3027)
- G. *Lucerapex casearia*, BATHUS 4 (CP948)
- H. *Lucerapex* sp., BIOCAL (DW44)
- I. *Lucerapex laevicarinatus*, MNHN-IM-2013-52063, , EXBODI (CP3934)

### Groupe “*chilensis*”

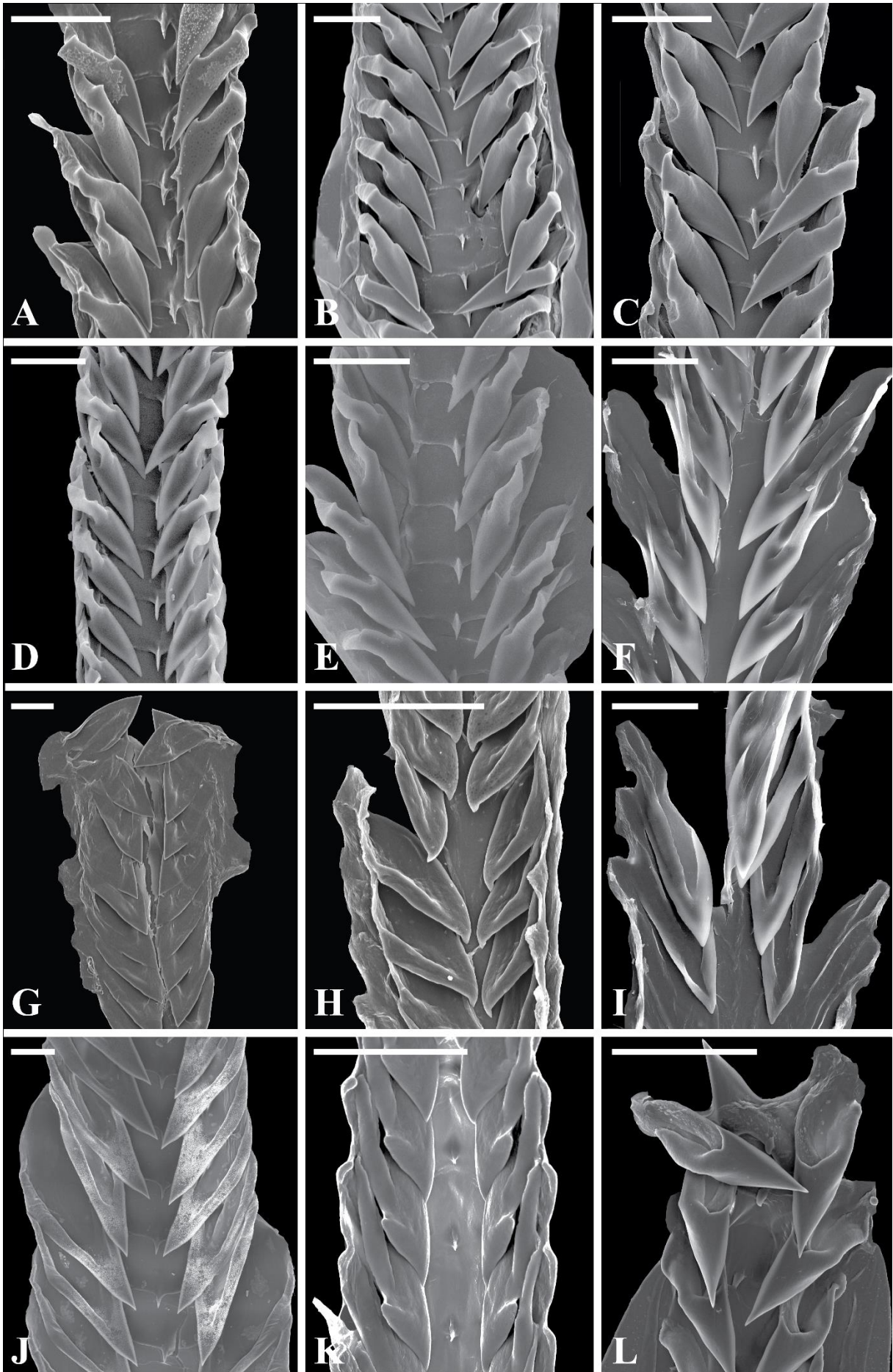
- J. “*Ptychosryinx*” *chilensis*, MNHN-IM-2013-43303, Hors Campagne (Chili)

### Groupe “*closterion*”

- K. *Gemmula closterion*, BATHUS 2 (DW721)

### *Kuroshioturris*

- L. *Kuroshioturris nipponicus*, MNHN-IM-2009-18984, AURORA 2007 (CP2680)





## PLANCHE C2

### *Turridrupa*

- A. *Turridrupa* aff. *magnifica*, MNHN-IM-2013-68165, KANACONO (DW4706)
- B. *Turridrupa consobrina*, MNHN-IM-2007-40841, PANGLAO 2004 (L46)
- C. *Turridrupa* sp., MNHN-IM-2013-48128, KANADEEP (DW4981)
- D. *Turridrupa albofasciata*, MNHN-IM-2013-14953, PAPUA NIUGINI (PS23)
- E. *Turridrupa acutigemmata*, MNHN-IM-2013-68187, KANACONO (DW4708)
- F. *Turridrupa cincta*, MNHN-IM-2009-14766, ATIMO VATAE (BS04)
- G. *Turridrupa* aff. *elongata*, MNHN-IM-2007-41300, SANTO 2006 (ZS19)
- H. *Turridrupa erythraea*, MNHN-IM-2013-57409, Hors Campagne (Arabie Saoudite)

### *Polystira*

- I. *Polystira albida*, MNHN-IM-2009-24360, KARUBENTHOS 2012 (GD37)
- J. *Polystira coltrorum*, MNHN-IM-2013-56055, GUYANE 2014 (CP4354)
- K. *Polystira* sp. 1, MNHN-IM-2013-60384, KARUBENTHOS 2 (DW4545)
- L. *Polystira gruneri*, MNHN-IM-2013-72822, MADIBENTHOS (AD298)
- M. *Polystira* sp. 3, MNHN-IM-2013-61170, KARUBENTHOS 2 (DW4613)
- N. *Polystira florencae*, MNHN-IM-2013-56571, GUYANE 2014 (CP4384)
- O. *Polystira staretti*, MNHN-IM-2013-61049, KARUBENTHOS 2 (DW4597)



## PLANCHE R2

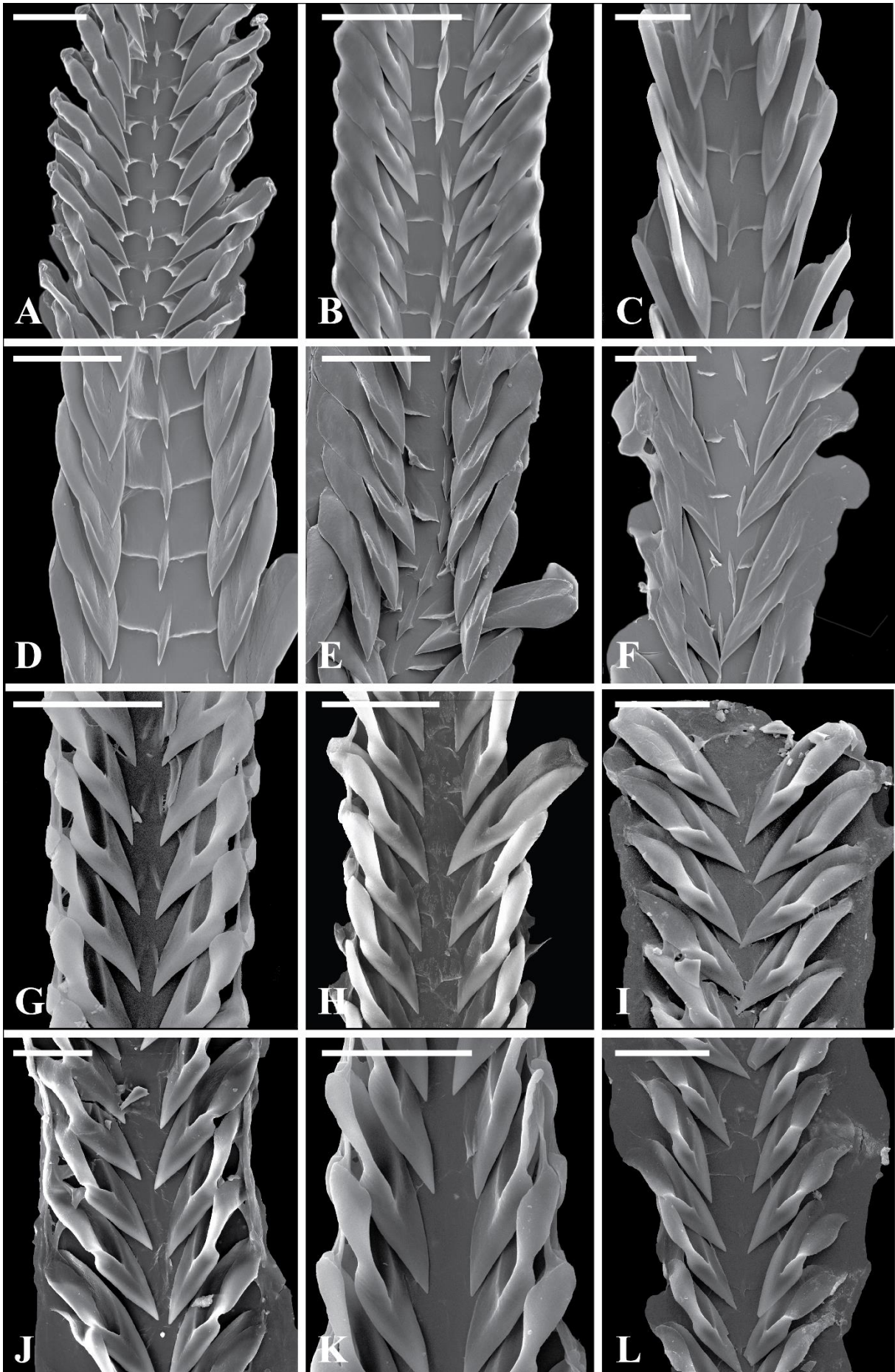
### *Turridrupa*

- A. *Turridrupa fedosovi*, holotype, LAGON (DW1148)
- B. *Turridrupa maestrati*, BATHUS 4 (DW942)
- C. *Turridrupa elongata*, MNHN-IM-2013-3038, PAPUA NIUGINI (PR48)
- D. *Turridrupa albofasciata*, LIFOU 2000 (1453)
- E. *Turridrupa cincta*, ATIMO VATAE (TA01)
- F. *Turridrupa erythraea*, SMF 346016 Hors Campagne (Egypte)

### *Polystira*

- G. *Polystira florencae*, MNHN-IM-2013-60119, KARUBENTHOS 2 (CP4517)
- H. *Polystira* sp. 3, MNHN-IM-2013-61158, KARUBENTHOS 2 (DW4612)
- I. *Polystira* sp. 4, MNHN-IM-2013-60870, KARUBENTHOS 2 (DW4583)
- J. *Polystira tellea*, MNHN-IM-2013-64847, PROTEUS-GUYANE (DW5071)
- K. *Polystira* sp. 1, MNHN-IM-2013-60384, KARUBENTHOS 2 (DW4545)
- L. *Polystira* sp. 8, MNHN-IM-2013-61038, KARUBENTHOS 2 (DW4594)



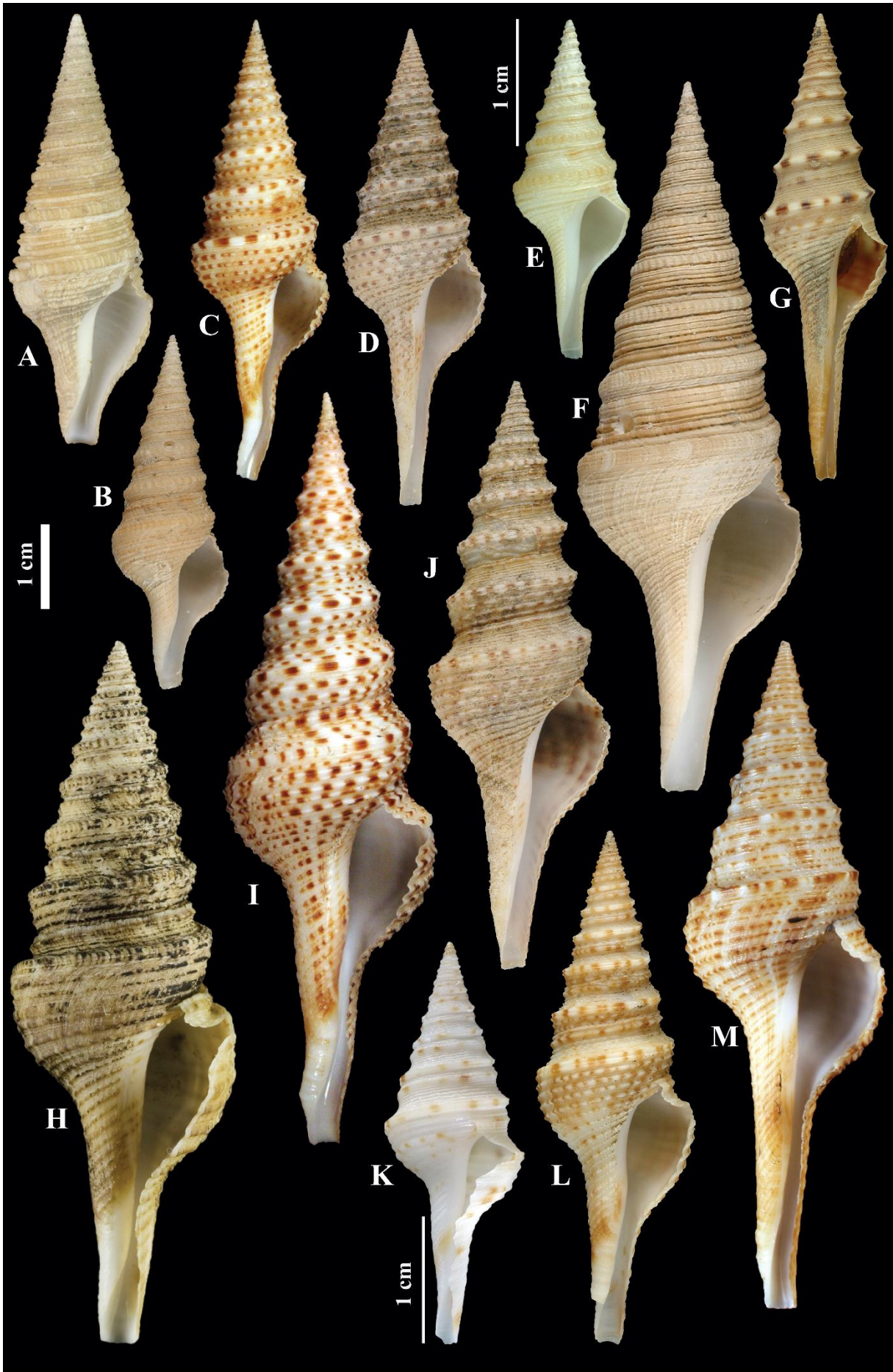




## PLANCHE C3

### *Unedogemmula*

- A.** *Unedogemmula sibukoensis*, MNHN-IM-2013-45247, MADEEP (CP4264)
- B.** *Unedogemmula* sp. 1, MNHN-IM-2009-16974, BIOPAPUA (CP3707)
- C.** *Unedogemmula bisaya*, MNHN-IM-2007-40867, PANGLAO 2005 (DW2375)
- D.** *Unedogemmula* cf. *indica* 5, MNHN-IM-2009-14929, ATIMO VATAE (CP3584)
- E.** *Unedogemmula* cf. *indica* 4 (juvenile), MNHN-IM-2007-17643, BOA1 (CP2430)
- F.** *Unedogemmula* cf. *indica* 4, MNHN-IM-2009-16946, BIOPAPUA (CP3707)
- G.** *Unedogemmula* cf. *indica* 6, MNHN-IM-2009-29708, Hors Campagne (ND7, Viêt Nam)
- H.** *Unedogemmula* cf. *panglaoensis* 3, MNHN-IM- 2013-9875, PAPUA NIUGINI (CP4028)
- I.** *Unedogemmula* aff. *friedrichbonhoefferi*, MNHN-IM-2007-41286, SANTO 2006 (EP27)
- J.** *Unedogemmula madagascarensis*, MNHN-IM-2009-15015, ATIMO VATAE (CP3583)
- K.** *Unedogemmula unedo* (juvenile), MNHN-IM-2007-40757, EBISCO (DW2512)
- L.** *Unedogemmula unedo*, MNHN-IM-2007-40625, AURORA 2007 (CP2666)
- M.** *Unedogemmula tayabasensis*, MNHN-IM-2007-40675, AURORA 2007 (CP2719)



## PLANCHE R3

### *Unedogemmula*

**A., B.** *Unedogemmula sibukoensis*, MNHN-IM-2013-45350, MADEEP (CP4265)

**C.** *Unedogemmula* sp. 1, MNHN-IM-2009-17225, BIOPAPUA (CP3708)

**D.** *Unedogemmula unedo*, BATHUS 4 (CP906)

**E.** *Unedogemmula* aff. *friedrichbonhoefferi*, VAUBAN 1978-1979 (9)

**F.** *Unedogemmula* aff. *indica*, MUSORSTORM 8

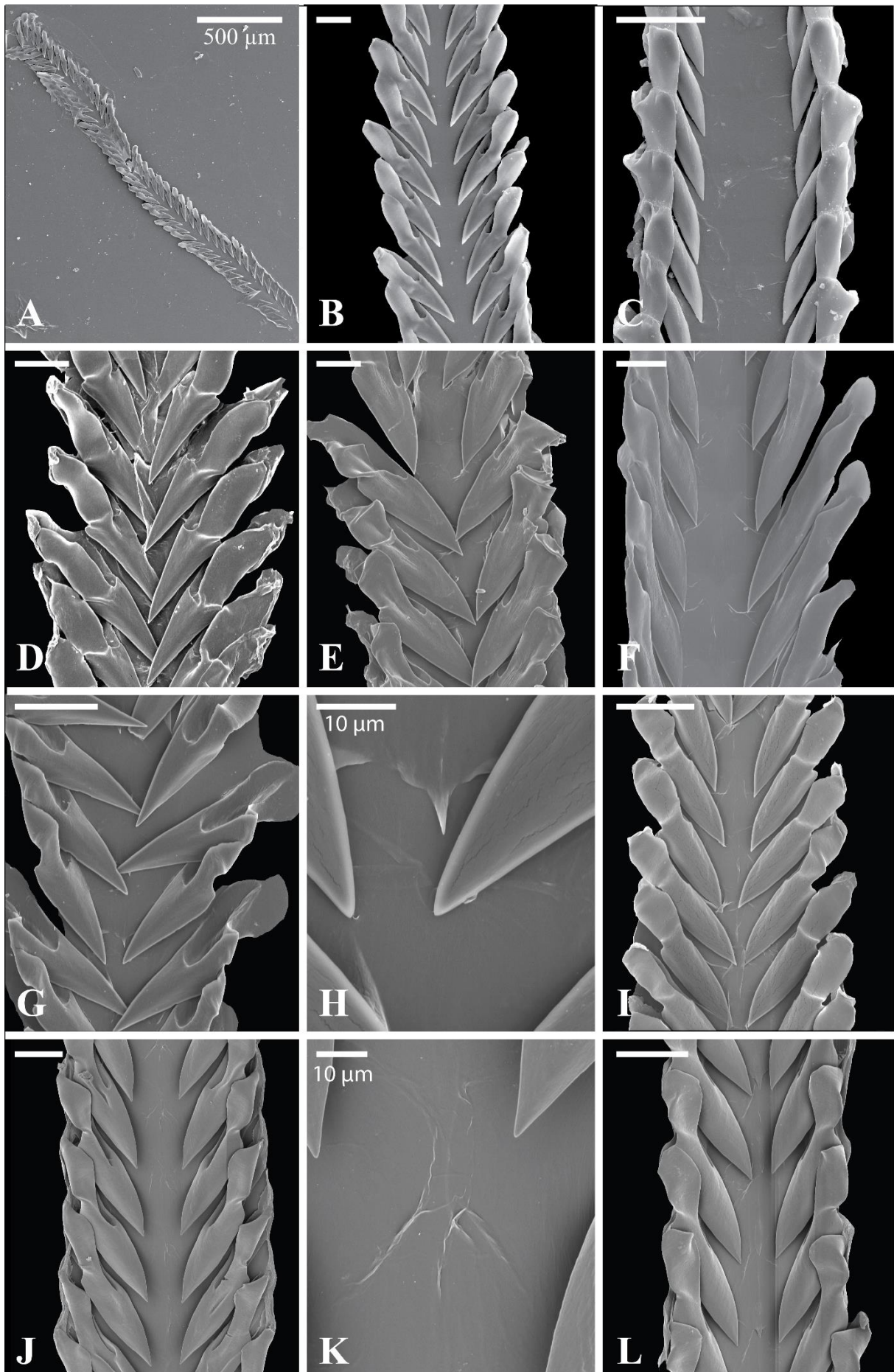
**G., H.** *Unedogemmula* aff. *indica*, BATHUS 2 ou MUSORSTOM 8

**I.** *Unedogemmula* aff. *indica*, MONTROUZIER (1237)

**J., K.** *Unedogemmula* aff. *bisaya*, Hors Camapgne (Philippines)

**L.** *Unedogemmula* aff. *bisaya*, Hors Camapgne (Philippines)





## PLANCHE C4

### Groupe “*kieneri*”

- A. *Gemmula speciosa*, MNHN-IM-2007-40884, PANGLAO 2005 (CP2377)
- B. *Gemmula* cf. *sogodensis* 1, MNHN-IM-2007-41011, SANTO 2006 (EP12)
- C. *Gemmula diomedea*, MNHN-IM-2007-40852, PANGLAO 2005 (CP2331)
- D. *Gemmula kieneri*, MNHN-IM-2013-68004, KANACONO (CP4656)
- E. *Gemmula* cf. *sogodensis* 2, , MNHN-IM-2007-40572, AURORA 2007 (CP2653)
- F. *Gemmula* cf. *cosmoi* 2, MNHN-IM-2009-29030, EXBODI (CP3833)
- G. *Gemmula martini*, MNHN-IM-2013-61824, ZHONGSHA 2015 (DW4143)
- H. *Gemmula* cf. *cosmoi* 1, MNHN-IM-2009-14830, ATIMO VATAE (CP3583)

### Groupe “sp. 05”

- I. *Gemmula* sp. 05, MNHN-IM-2013-45289, MADEEP (CP4264)
- J. *Gemmula* sp. 28, MNHN-IM-2009-29347, BIOPAPUA (CP3728)
- K. *Gemmula* sp. 39, MNHN-IM-2013-69308, KANACONO (DW4711)

### Groupe “*deshayesii*”

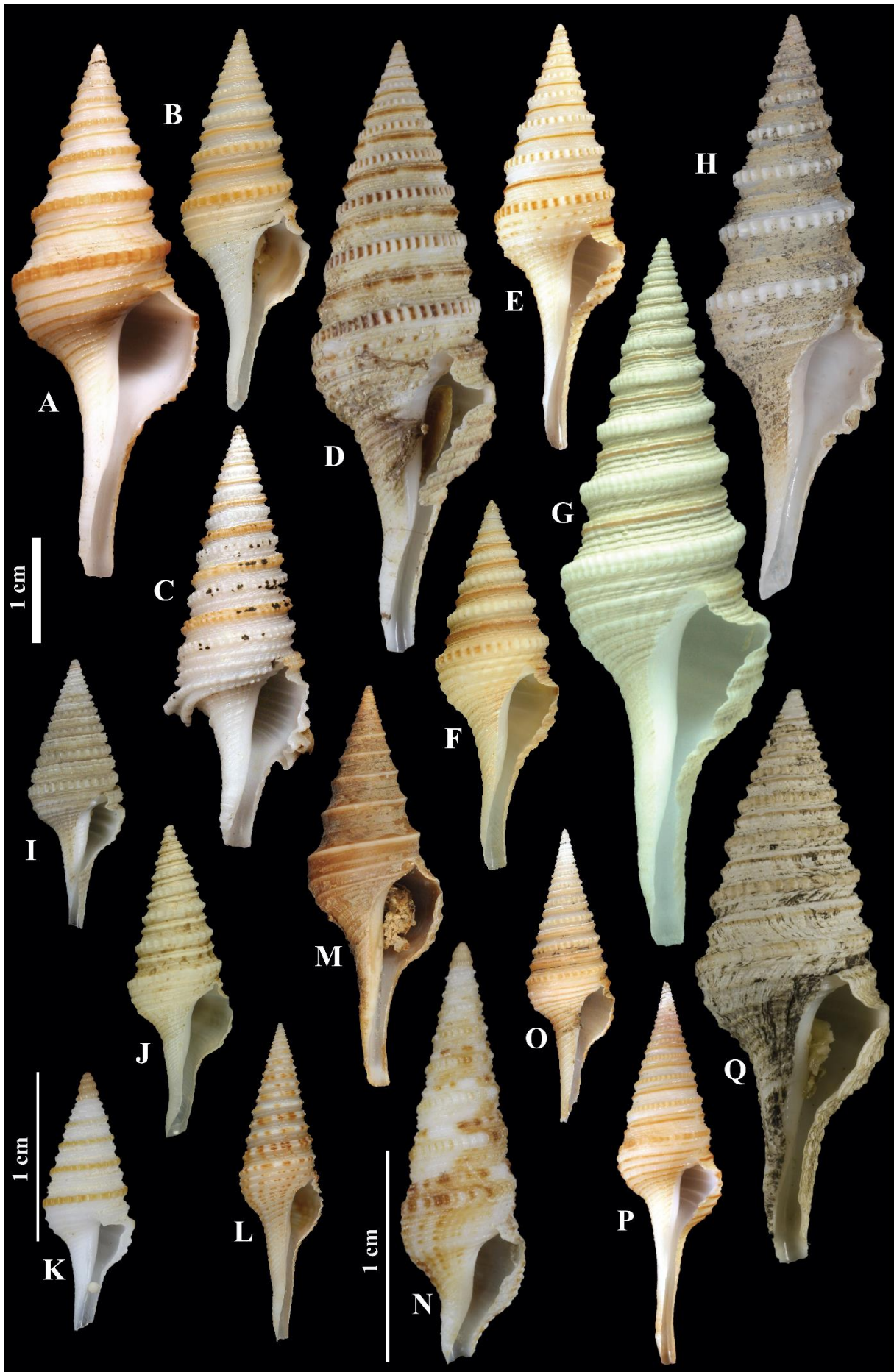
- L. *Gemmula hastula*, MNHN-IM- 2007-41157, SANTO 2006 (AT44)
- M. *Gemmula leucotropis*, MNHN-IM-2013-49934, Hors Campagne (Japon)

### Groupe “*gemmaoides*”

- N. *Gemmula gemmaoides*, MNHN-IM-2013-73348, KANACONO (DW4764)

### Groupe “sp. 02 ”

- O. *Gemmula* sp. 02, MNHN-IM-2007-40907, PANGLAO 2005 (CP2377)
- P. *Gemmula* sp. 03, MNHN-IM-2007-40715, AURORA 2007 (CP2760)
- Q. *Gemmula* sp. 29, MNHN-IM-2013-9780, PAPUA NIUGINI (CP4022)



## PLANCHE R4

### Groupe “*kieneri*”

- A. *Gemmula speciosa*, Hors Campagne (Philippines)
- B. *Gemmula* cf. *kieneri*, BATHUS 1 (DW886)
- C. *Gemmula diomedea*, MNHN-IM-2013-46051, MADEEP (CP4331)
- D. *Gemmula* cf. *sogodensis*, Hors Campagne (Philippines)

### Groupe “sp. 02”

- E. *Gemmula* sp. 29, MNHN-IM-2013-19161, PAPUA NIUGINI (CP4065)
- F. *Gemmula* sp. 03, MNHN-IM-2007-40715, AURORA 2007 (CP2760)

### Groupe “*gemmaoides*”

- G. *Gemmula gemmaoides*, MNHN-IM- 2013-58747, KAVIENG (DW4473)

### Groupe “*deshayesii*”

- H. *Gemmula hastula*, MNHN-IM-2007-41171, SANTO 2006 (AT56)
- I. *Gemmula leucotropis*, MNHN-IM-2013-49934, Hors Campagne (Japon)

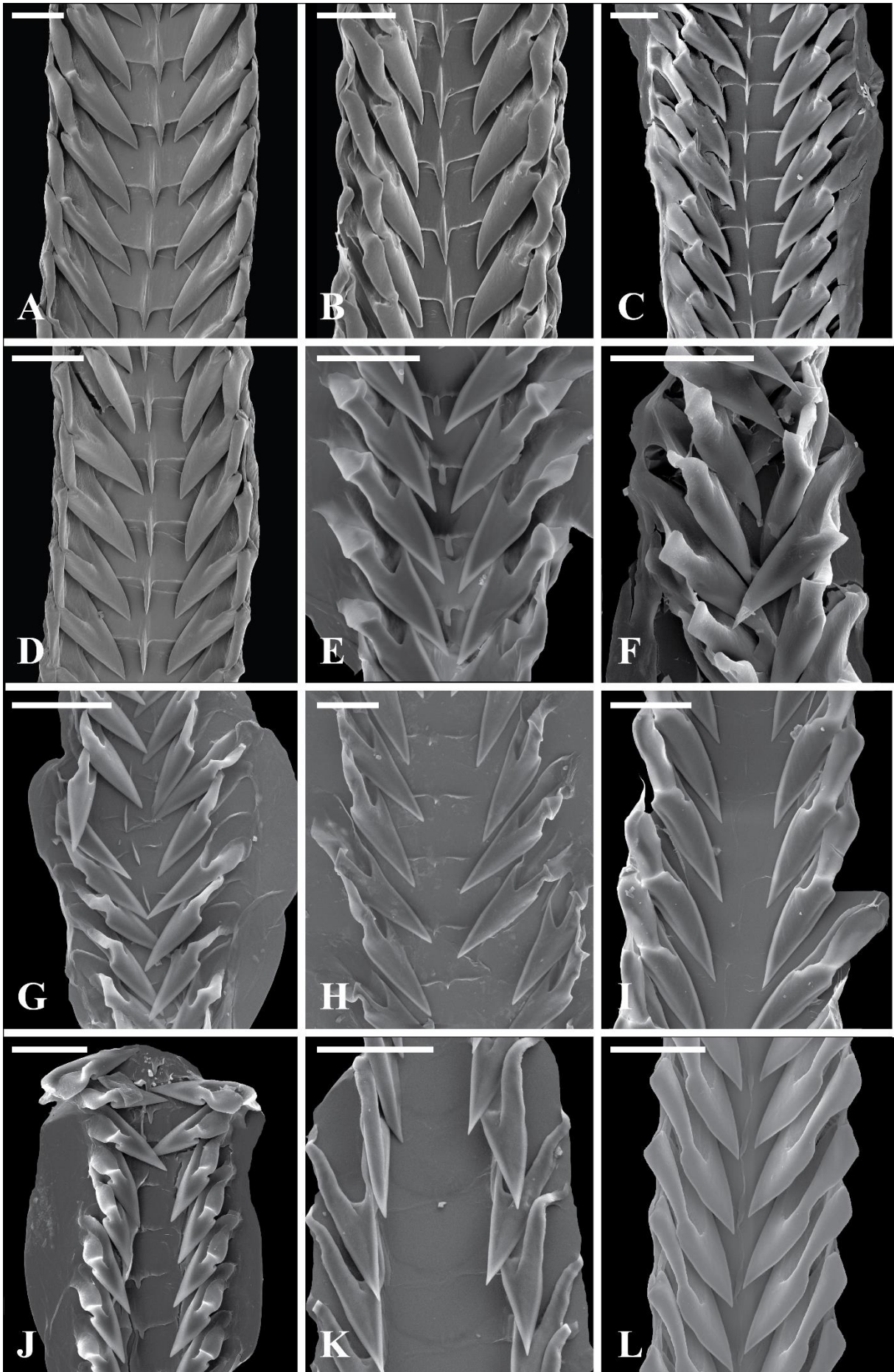
### Groupe “sp. 05”

- J. *Gemmula* sp. 28, MNHN-IM-2013-46190, MADEEP (CP4333)
- K. *Gemmula* sp. 05, MNHN-IM-2013-15366, PAPUA NIUGINI (CP3948)

### ?Groupe “*lululimi*”

- L. *Gemmula lululimi*, MNHN-IM-2009-17042, BIOPAPUA (DW3770)







## PLANCHE C5

### Groupe “*rarimaculata*”

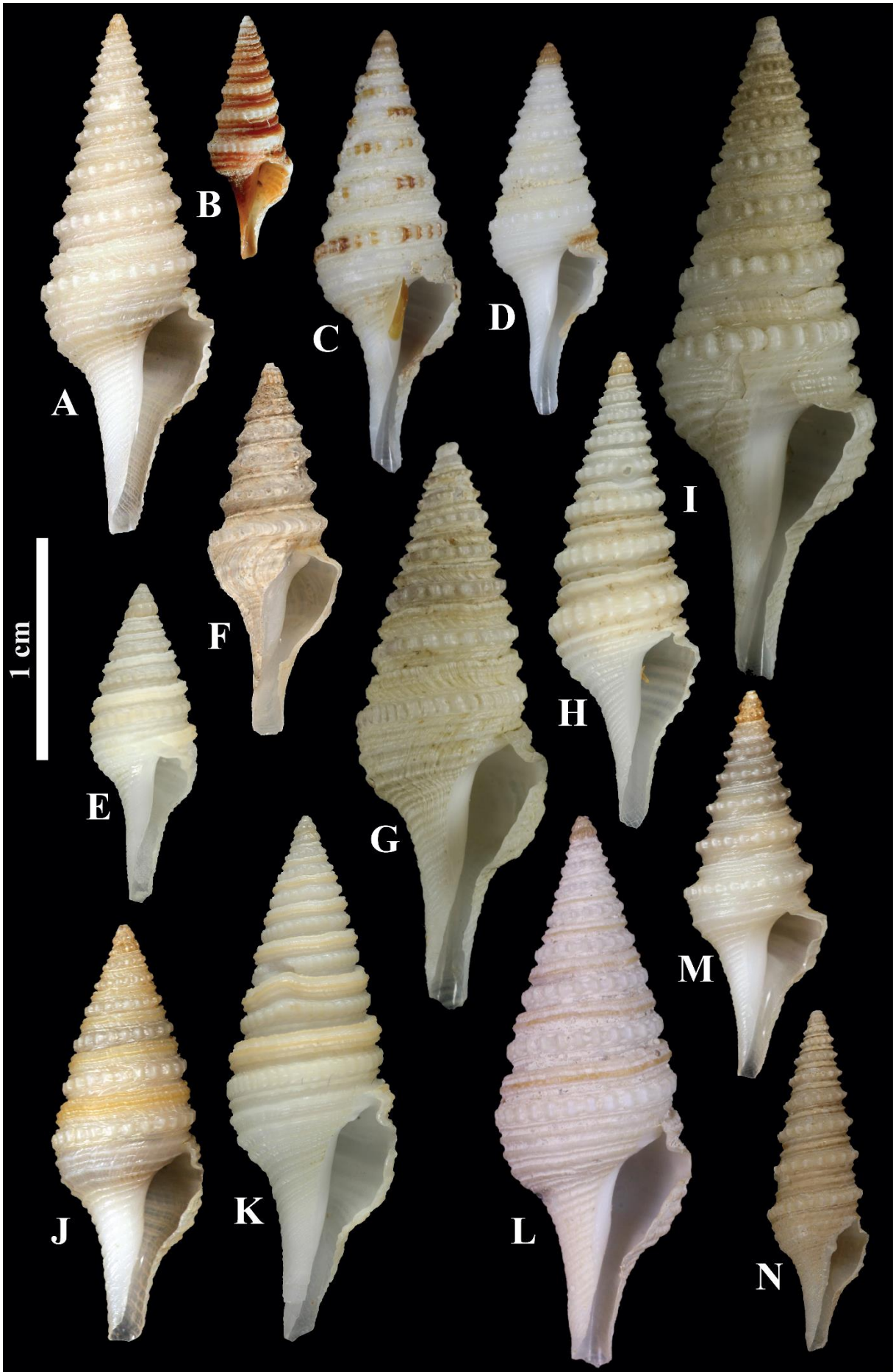
- A. *Gemmula* sp. 19, MNHN-IM-2007-41240, SANTO 2006 (AT73)
- B. *Gemmula* cf. *monilifera* 4, MNHN-IM-2007-40818, PANGLAO 2004 (T9)
- C. *Gemmula* cf. *rarimaculata* 2, MNHN-IM-2013-68775, KANACONO (DW4701)
- D. *Gemmula* cf. *rarimaculata* 1, MNHN-IM-2013-68154, KANACONO (DW4706)

### Groupe “sp. 01”

- E. *Gemmula* sp. 01, MNHN-IM-2007-40955, SALOMONBOA 3 (DW2775)
- F. *Gemmula* sp. 30, MNHN-IM-2009-18350, MIRIKY (CP3252)
- G. *Gemmula* sp. 21, MNHN-IM-2013-19021, PAPUA NIUGINI (CP4057)
- H. *Gemmula* sp. 24, MNHN-IM-2007-40927, SALOMON 2 (CP2184)
- I. *Gemmula* sp. 22, MNHN-IM-2013-19734, PAPUA NIUGINI (CP3981)
- J. *Gemmula* sp. 26, MNHN-IM-2007-41048, SANTO 2006 (AT19)
- K. *Gemmula* sp. 25, MNHN-IM-2007-40700, AURORA 2007 (CP2737)
- L. *Gemmula* sp. 40, MNHN-IM-2013-45818, MADEEP (DW4315)
- M. *Gemmula* sp. 23, MNHN-IM-2007-41019, SANTO 2006 (AT10)

### Groupe “sp. 27”

- N. *Gemmula* sp. 27, MNHN-IM-2009-17202, BIOPAPUA (CP3708)



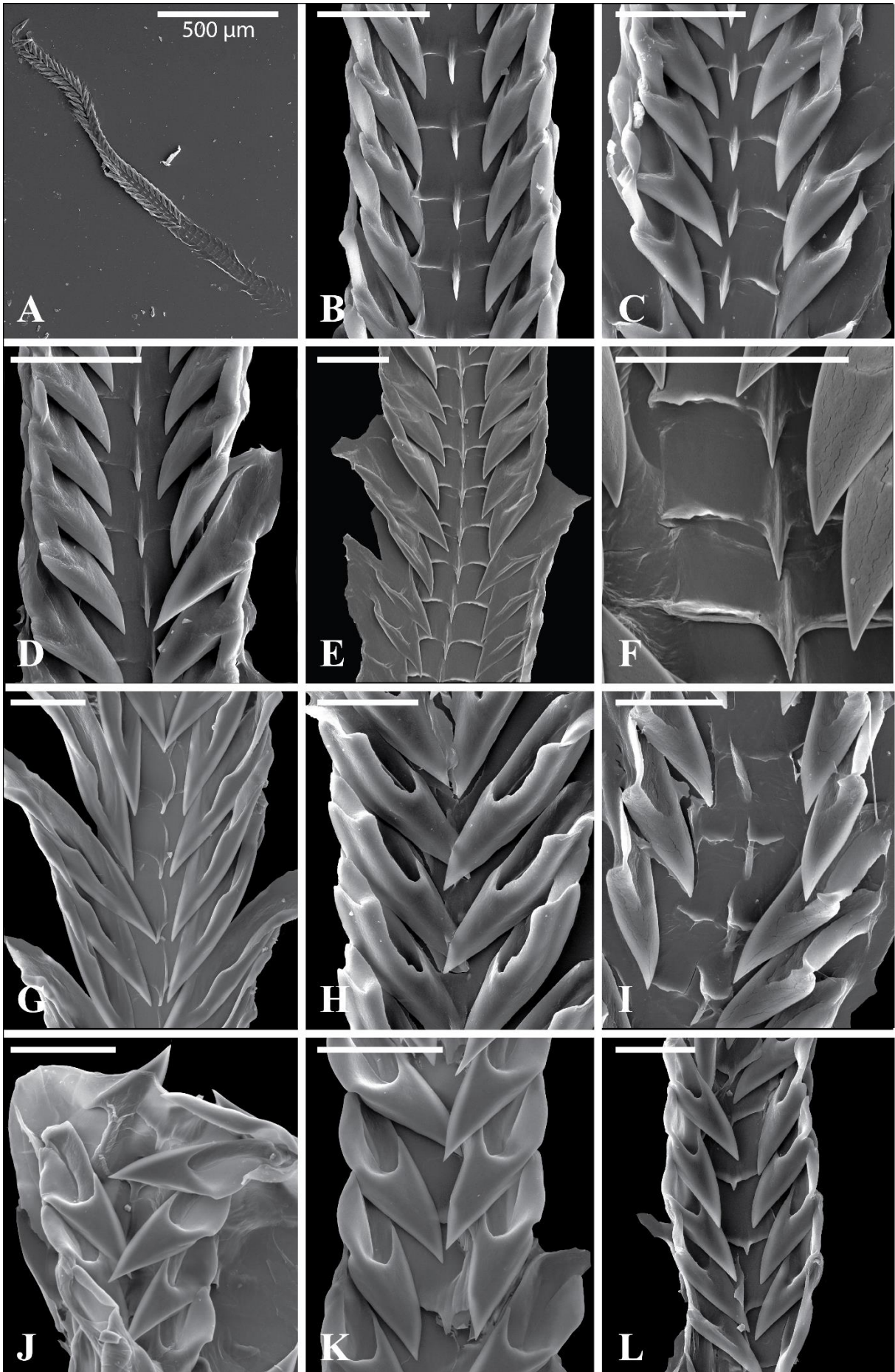
## PLANCHE R5

### Groupe “*rarimaculata*”

- A., B.** *Gemmula* sp. 19, MNHN-IM- 2013-61846, ZHONGSHA 2015 (DW4143)  
**C.** *Gemmula* cf. *rarimaculata* 1, MNHN-IM-2013-68121, KANACONO (DW4699)  
**D.** *Gemmula* cf. *rarimaculata* 2, MNHN-IM-2013-68015, KANACONO (CP4657)  
**E., F.** *Gemmula* cf. *rarimaculata*, MUSORSTOM 5 (263)

### Groupe “sp. 01”

- G.** *Gemmula* sp. 01, MNHN-IM-2007-40955, MADEEP (CP4335)  
**H.** *Gemmula* sp. 23, MNHN-IM- 2013-9831, PAPUA NIUGINI (CP4025)  
**I.** *Gemmula* sp. 25, MNHN-IM-2007-40693, AURORA 2007 (CP2730)  
**J., K.** *Gemmula* sp. 21, MNHN-IM-2013-15933, PAPUA NIUGINI (CP3961)  
**L.** *Gemmula* sp. 40, MNHN-IM-2013-45818, MADEEP (DW4315)



## PLANCHE C6

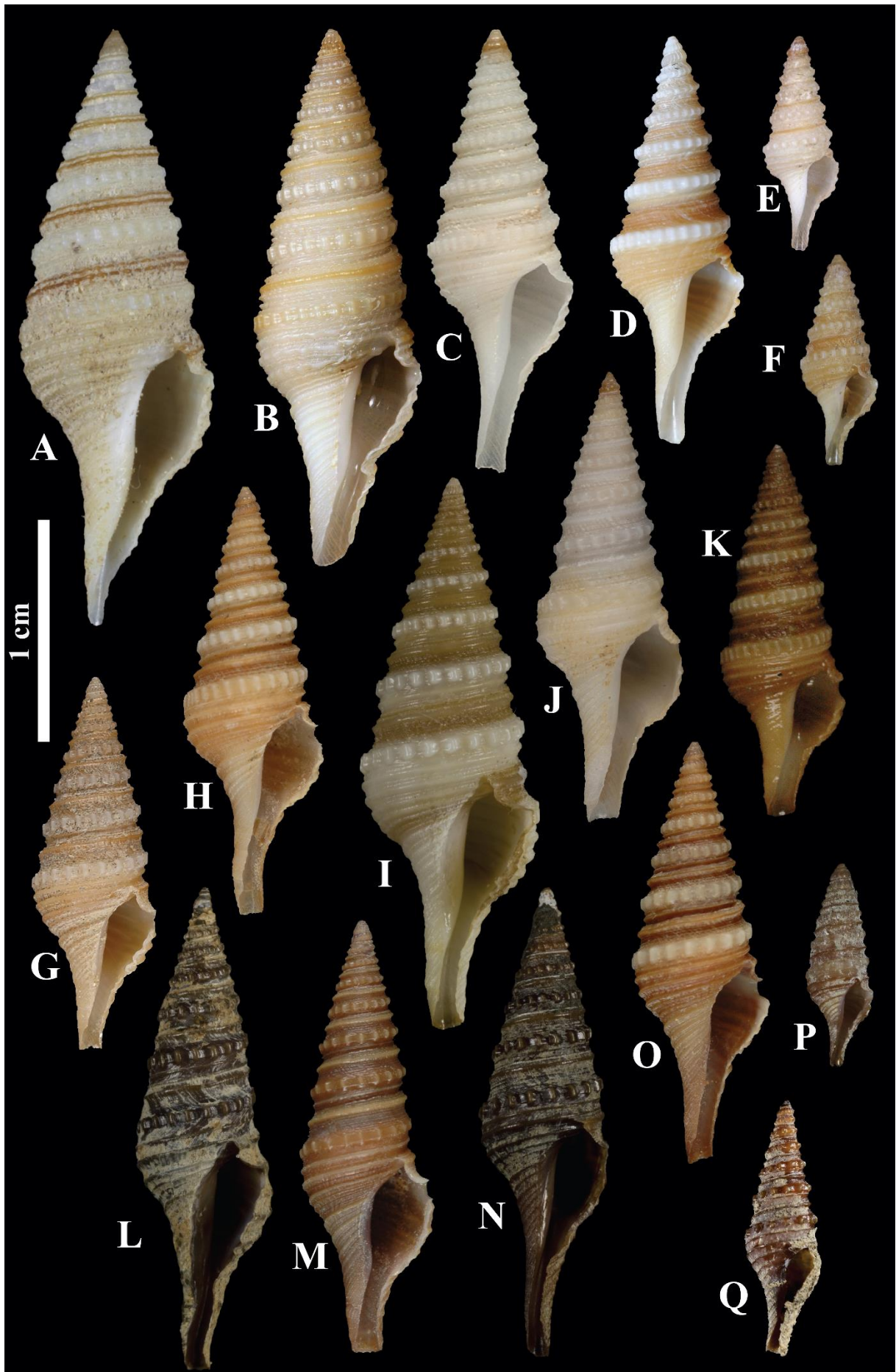
### Groupe “sp. 16”

- A. *Gemmula* sp. 17, MNHN-IM-2013-62909, BIOMAGLO (DW4842)
- B. *Gemmula* sp. 18, MNHN-IM-2007-41015, SANTO 2006 (AT10)
- C. *Gemmula* sp. 16, MNHN-IM-2009-16883, MIRIKY (DW3217)

### *Eugemmula*

- D. *Eugemmula* cf. *hombroni* 2, MNHN-IM- 2007-40870, PANGLAO 2005 (CP2377)
- E. *Gemmula* sp. 44, MNHN-IM-2013-46280, MADEEP (CP4336)
- F. *Gemmula* sp. 43, MNHN-IM-2013-46250, MADEEP (CP4335)
- G. *Gemmula* sp. 31, MNHN-IM-2009-18343, MIRIKY (CP3203)
- H. *Eugemmula* cf. *monilifera* 1, MNHN-IM-2007-41433, SANTO 2006 (AT128)
- I. *Eugemmula* cf. *monilifera* 3, MNHN-IM-2013-15500, PAPUA NIUGINI (PR107)
- J. *Eugemmula* cf. *monilifera* 5, MNHN-IM-2009-14878, ATIMO VATAE (TP11)
- K. *Gemmula* sp. 32, MNHN-IM-2009-26527, INHACA 2011 (MR11)
- L. *Eugemmula* cf. *hombroni* 4, MNHN-IM-2013-12875, PAPUA NIUGINI (PD21)
- M. *Eugemmula* cf. *hombroni* 3, MNHN-IM-2007-41086, SANTO 2006 (ED05)
- N. *Gemmula* sp. 45, MNHN-IM-2013-12723, PAPUA NIUGINI (PD20)
- O. *Eugemmula* cf. *monilifera* 2, MNHN-IM-2007-41432, SANTO 2006 (AT128)
- P. *Gemmula* sp. 42, MNHN-IM-2013-11188, PAPUA NIUGINI (PD01)
- Q. *Gemmula* sp. 33, MNHN-IM-2013-5651, PAPUA NIUGINI (PD67)





## PLANCHE R6

### Groupe “sp. 16”

**A.** *Gemmula* sp. 17, MNHN-IM-2013-62909, BIOMAGLO (DW4842)

**B., C.** *Gemmula* sp. 18, MNHN-IM- 2007-41379, SANTO 2006 (AT101)

### *Eugemmula*

**D.** *Eugemmula* cf. *monilifera*, MUSORSTORM 8 (DW988)

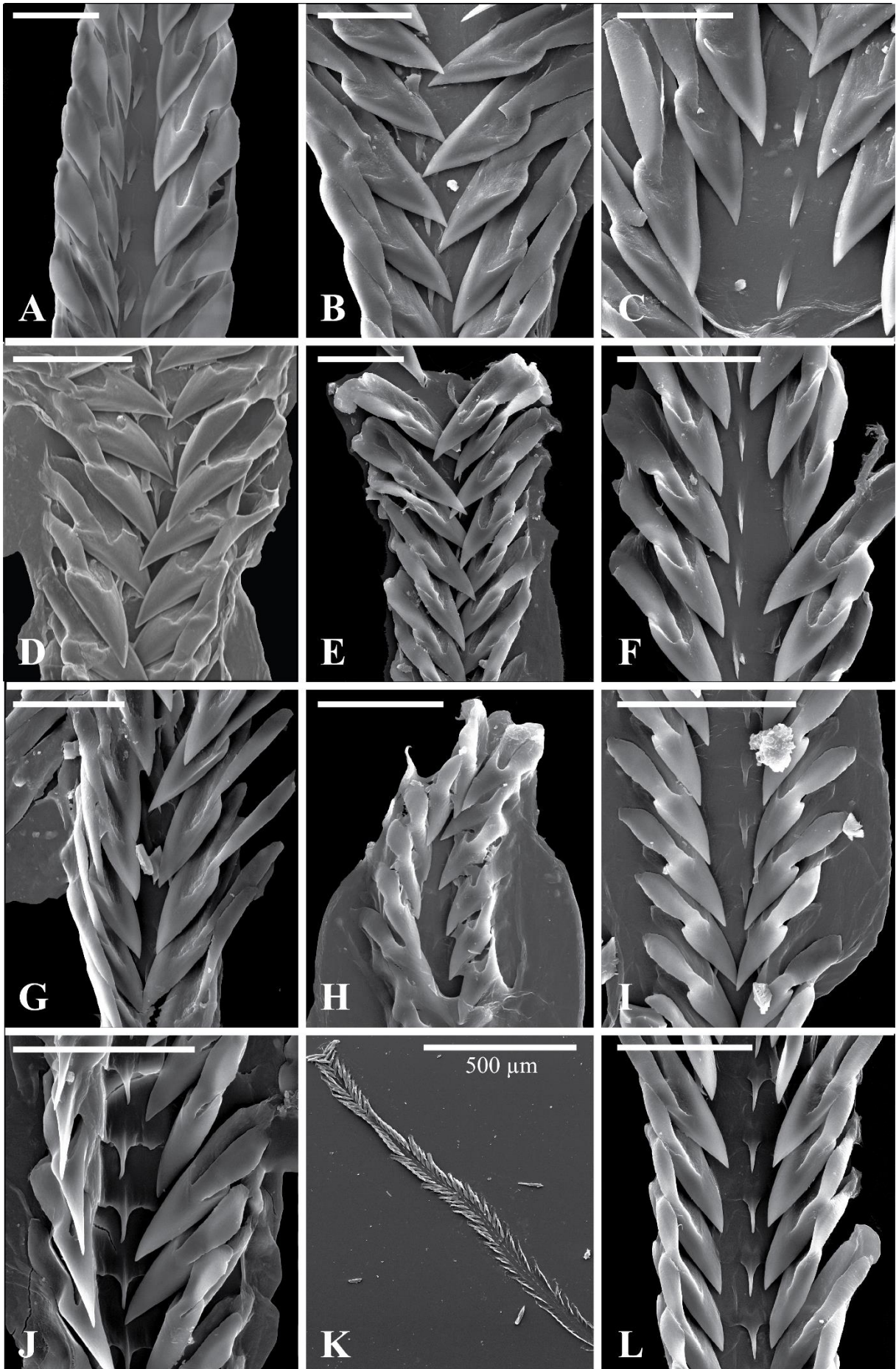
**E., F.** *Eugemmula* cf. *monilifera* 2, MNHN-IM-2013-58542, KAVIENG (CP4456)

**G.** *Gemmula* sp. 31, MNHN-IM-2009-18337, MIRIKY (CP3203)

**H., I.** *Gemmula* sp. 44, MNHN-IM-2013-46230, MADEEP (CP4335)

**J.** *Eugemmula* cf. *hombroni* 2, MNHN-IM- 2013-58707, KAVIENG (DW4470)

**K., L.** *Eugemmula* cf. *hombroni* 4, MNHN-IM-2013-12719, PAPUA NIUGINI (PD20)





## PLANCHE C7

### *Turris*

- A.** *Turris babylonia*, MNHN-IM-2013-55794, KAVIENG (KR184)
- B.** *Turris crispa*, MNHN-IM-2007-36648, MIRIKY (CP3288)
- C.** *Turris normandavidsoni*, MNHN-IM-2013-17444, PAPUA NIUGINI (PR165)
- D.** *Turris ambages*, Hors Campagne (Afrique du Sud)
- E.** *Turris chaldaea*, MNHN-IM-2013-53855, KAVIENG (KR100)
- F.** *Turris condei*, MNHN-IM-2013-12546, PAPUA NIUGINI (PR24)

### *Annulaturris*

- G.** *Annulaturris kathiewayae*, MNHN-IM-2013-69307, KANACONO (DW4711)
- H.** *Annulaturris* cf. *nadaensis* 3, MNHN-IM-2009-13365, PANGLAO 2004 (L77)
- I.** *Annulaturris* cf. *nadaensis* 2, MNHN-IM-2013-15527, PAPUA NIUGINI (PR110)
- J.** *Annulaturris pagasa*, MNHN-IM-2009-13354, CONCALIS (DW3017)
- K.** *Annulaturris* cf. *nadaensis* 1 MNHN-IM-2013-17604, PAPUA NIUGINI (PR173)



## PLANCHE R7

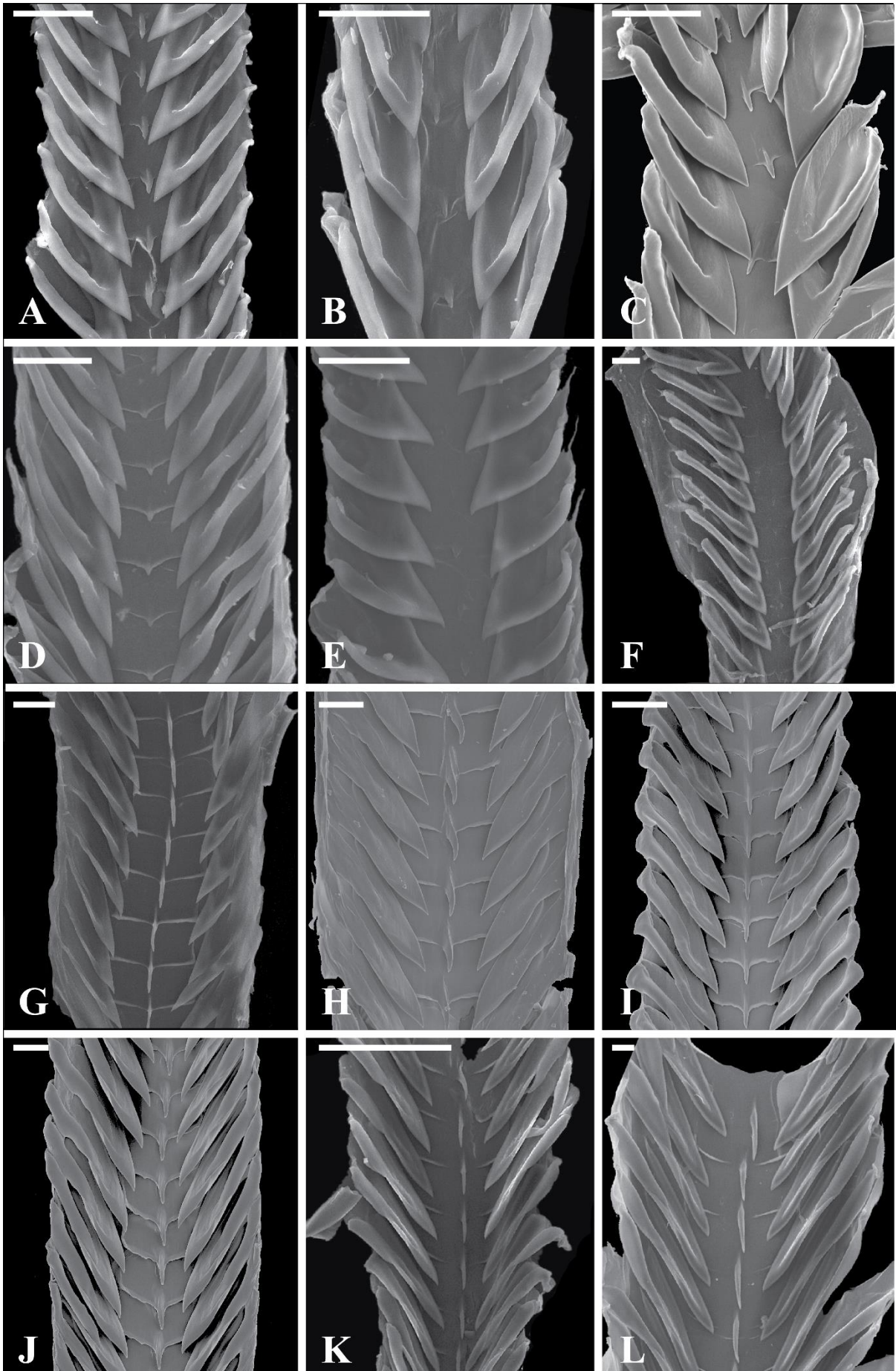
### *Turris*

- A. *Turris babylonia*
- B. *Turris chaldaea*
- C. *Turris crista*, LAGON (DW80)
- D. *Turris hidalgoi*
- E. *Turris spectabilis*
- F. *Turris normandavidsoni*

### *Annulaturris*

- G. *Annulaturris amicta*
- H. *Annulaturris annulata*, NMSA
- I. *Annulaturris kathiewayae*, paratype MNHN-IM-2009-13559, TERRASSES (DW3120)
- J. *Annulaturris pagasa*, MNHN-IM-2009-13353, CONCALIS (DW2949)
- K. *Annulaturris nadaensis*
- L. *Annulaturris undosa*





## PLANCHE C8

### Groupe “sp. 20”

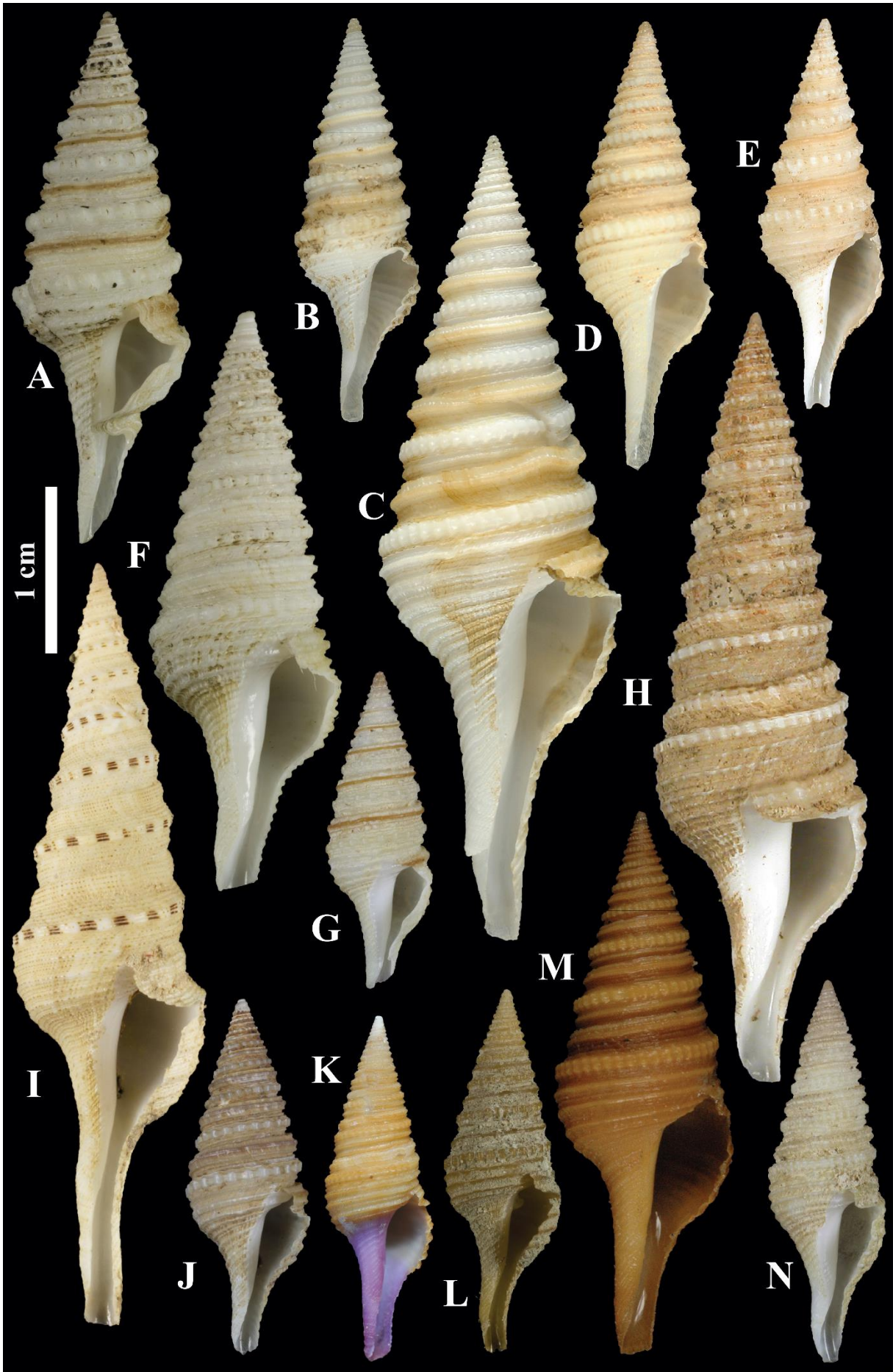
- A. *Gemmula* sp. 20, MNHN-IM-2013-9892, PAPUA NIUGINI (CP4028)

### Groupe “sp. 12”

- B. *Gemmula* sp. 14, MNHN-IM-2007-40944, SALOMON 2 (CP2282)  
C. *Gemmula* sp. 15, MNHN-IM-2007-40598, AURORA 2007 (CP2655)

### Groupe “*albina*”

- D. *Gemmula* sp. 08, MNHN-IM-2009-29008, EXBODI (CP3827)  
E. *Gemmula* sp. 35, MNHN-IM-2007-38228, MIRIKY (DW3244)  
F. *Gemmula* sp. 12, MNHN-IM-2013-18847, PAPUA NIUGINI (CP4049)  
G. *Gemmula* sp. 37, MNHN-IM-2013-68155, KANACONO (DW4703)  
H. *Gemmula* sp. 36, MNHN-IM-2007-38180, MIRIKY (CP3241)  
I. *Gemmula albina*, MNHN-IM-2013-54348, KAVIENG (KR102)  
J. *Gemmula* sp. 09, MNHN-IM-2013-46183, MADEEP (CP4333)  
K. *Gemmula lisajoni*  
L. *Gemmula* sp. 07, MNHN-IM-2013-13391, PAPUA NIUGINI (PR42)  
M. *Gemmula* sp. 34, MNHN-IM-2009-26515, INHACA 2011 (MD15)  
N. *Gemmula* sp. 11, MNHN-IM-2013-68193, KANACONO (DW4714)





## PLANCHE R8

### Groupe “sp. 20”

**A., B.** *Gemmula* sp. 20, MNHN-IM-2013-9892, PAPUA NIUGINI (CP4028)

### Groupe “sp. 12”

**C.** *Gemmula* sp. 41, MNHN-IM-2013-46311, MADEEP (CP4337)

**D., E., F.** *Gemmula* sp. 12, MNHN-IM-2013-9776, PAPUA NIUGINI (CP4023)

**G.** *Gemmula* sp. 15, MNHN-IM-2007-40649, AURORA 2007 (CP2711)

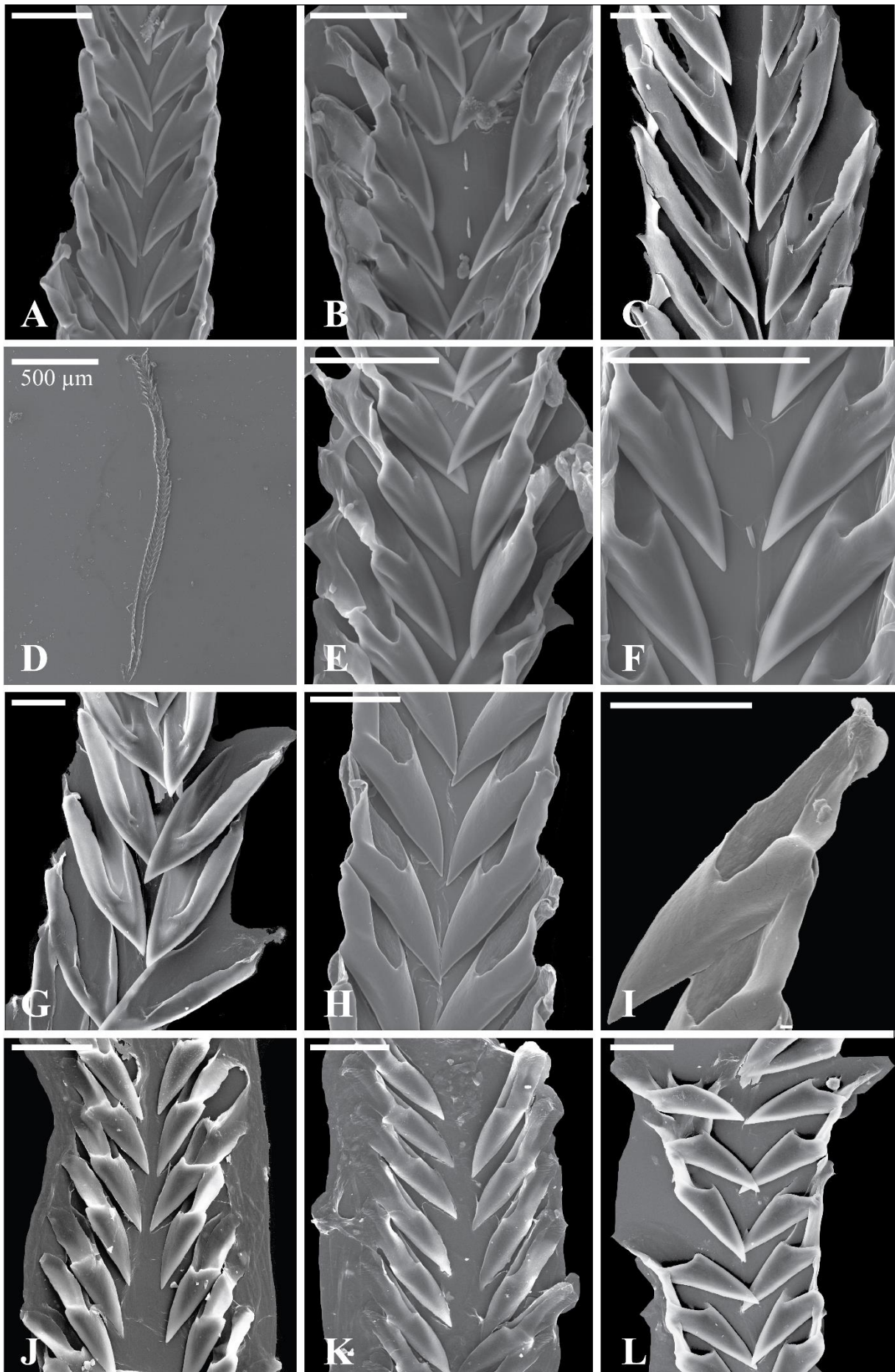
### Groupe “albina”

**H., I.** *Gemmula albina*, MONTROUZIER (1319)

**J.** *Gemmula* sp. 08, MNHN-IM-2013-68012, KANACONO (CP4656)

**K.** *Gemmula* sp. 09, MNHN-IM-2013-46153, MADEEP (CP4333)

**L.** *Gemmula* sp. 11, MNHN-IM-2013-68193, KANACONO (DW4714)

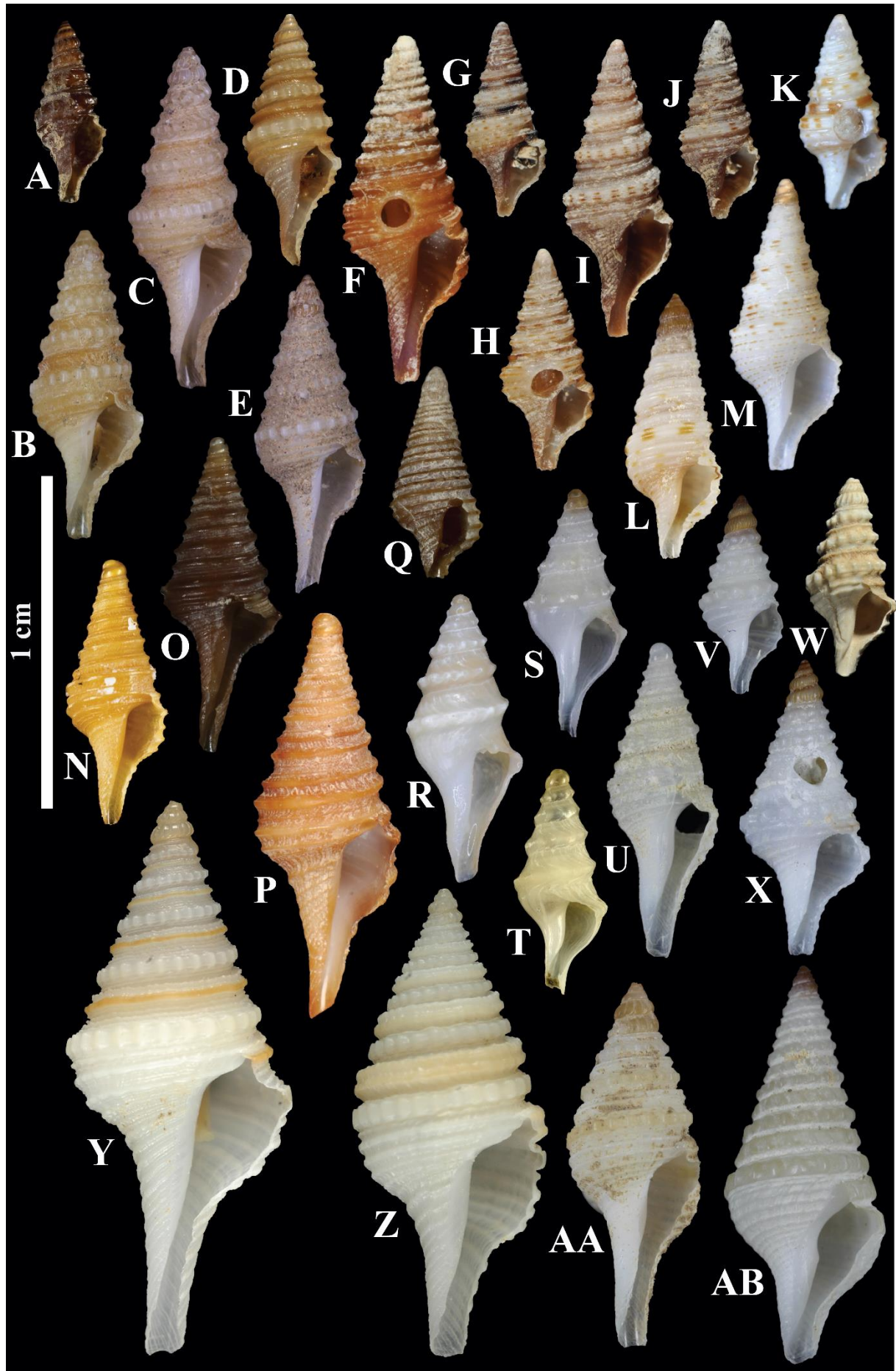




## PLANCHE C9

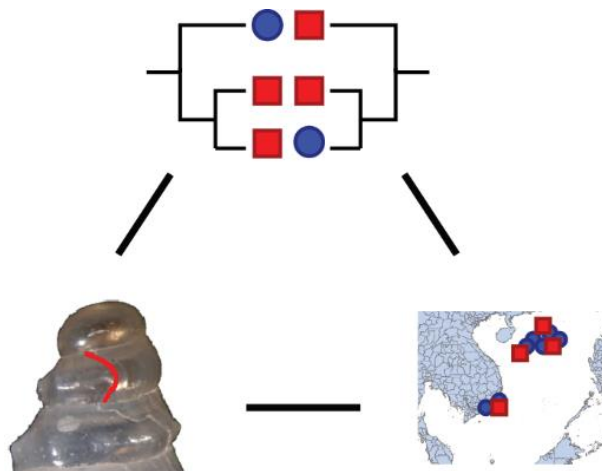
### Juveniles

- A. *Gemmula* sp. 33, MNHN-IM-2013-4356, PAPUA NIUGINI (PD67)
- B. *Gemmula* sp. 43, MNHN-IM-2013-46250, MADEEP (CP4335)
- C. *Eugemmula* cf. *hombroni* 2, MNHN-IM-2013-46207, MADEEP (CP4335)
- D. *Eugemmula* cf. *monilifera* 2, MNHN-IM-2007-41159, SANTO 2006 (AT44)
- E. *Gemmula* cf. *rarimaculata* 1, MNHN-IM-2013-46218, MADEEP (CP4335)
- F. *Gemmula lisajoni*, MNHN-IM-2007-40982, SANTO 2006 (DS04)
- G. *Lophiotoma abbreviata*, MNHN-IM-2007-41322, SANTO 2006 (FS79)
- H. *Lophiotoma brevicaudata*, MNHN-IM-2007-41031, SANTO 2006 (DB48)
- I. *Gemmula albina*, MNHN-IM-2007-41321, SANTO 2006 (FS79)
- J. *Iotyrris cingulifera*, MNHN-IM-2007-41323, SANTO 2006 (FS79)
- K. *Lophiotoma vezzaroi*, MNHN-IM-2007-40988, SANTO 2006 (DS04)
- L. *Iotyrris musivum*, MNHN-IM-2013-53459, KAVIENG (KB36)
- M. *Lophiotoma bratasusa*, MNHN-IM-2007-41333, SANTO 2006 (FS84)
- N. *Polystira* sp. 10, MNHN-IM-2013-9039, KARUBENTHOS 2012 (GD21)
- O. *Unedogemmula* sp. 02, MNHN-IM-2013-14405, PAPUA NIUGINI (PD37)
- P. *Polystira gruneri*, MNHN-IM-2013-71295, MADIBENTHOS (AD261)
- Q. *Turridrupa consobrina*, MNHN-IM-2013-14863, PAPUA NIUGINI (PD48)
- R. *Lucerapex* cf. *carola*, MNHN-IM-2009-18952, PANGLAO 2005 (CP2362)
- S. *Lucerapex* sp., MNHN-IM-2013-68758, KANACONO (DW4690)
- T. *Lucerapex cracens*, MNHN-IM-2013-52064, BIOPAPUA (CP3738)
- U. *Gemmula closterion*, MNHN-IM-2013-68038, KANACONO (DW4661)
- V. *Cryptogemma praesignis*, MNHN-IM-2013-62893, BIOMAGLO (DW4819)
- W. *Pleurotoma rotatilis*, holotype ZMB 60070, Afrique de l'Est
- X. *Cryptogemma unilineata*, MNHN-IM-2013-62916, (BIOMAGLO (DW4833)
- Y. *Gemmula* cf. *cosmoi* 2, MNHN-IM-2007-40737, BOA1 (CP2413)
- Z. *Gemmula* sp. 10, MNHN-IM- 2007-40940, SALOMON 2 (CP2204)
- AA. *Gemmula* sp. 01, MNHN-IM-2013-46228, MADEEP (CP4335)
- AB. *Gemmuloborsonia neocaledonica*, MNHN-IM-2013-45753, MADEEP (DW4309)





**CHAPITRE 1**  
**TAXONOMIE INTEGRATIVE**



L'objectif de ce chapitre est d'identifier les unités opérationnelles qui vont nous servir de terminaux (« tips ») dans l'établissement de la phylogénie présentée le chapitre 3. La démarche utilisée pour délimiter ces entités est une démarche de taxonomie intégrative, présentée ci-après.

Il convient de définir ici l'expression « taxonomie intégrative », présente à la fois dans le titre de la thèse et de ce chapitre. Le terme existait déjà dans la littérature dans les domaines de la psychologie du développement et en management. Mais ce sont indépendamment Dayrat (2005) et Will *et al.* (2005) qui vont donner à cette expression ses lettres de noblesse dans le cadre de la systématique et de la biologie évolutive. Les deux définitions diffèrent un peu : Dayrat présente la taxonomie intégrative comme « la science qui a pour but de délimiter les unités de la diversité du vivant à partir de perspectives multiples et complémentaires (phylogéographie, anatomie comparée, génétique des populations, écologie, développement, comportement, etc.) », tandis que Will *et al.* en parlent comme d'une discipline « utilisant un grand nombre de caractères incluant l'ADN et plein d'autres types de données, dans le but de délimiter, découvrir et identifier des espèces naturelles et des taxons à tous les niveaux pertinents ». On note toutefois que la première définition peut s'appliquer à la taxonomie telle qu'elle a été pratiquée depuis plusieurs siècles : la présentation des Turridae dans l'Introduction en est un exemple, les taxonomistes ont toujours cherché à intégrer différents types de données (morphologie de la coquille, radula, protoconque, anatomie, moléculaire) afin d'établir leur « unités de la diversité du vivant ». On peut alors se demander quelle est la différence entre la définition de la taxonomie intégrative de Dayrat (2005) et la définition de taxonomie. La deuxième définition en revanche fait explicitement référence à l'ADN d'une part et distingue les espèces des autres taxons d'autre part. La référence à l'ADN s'explique par le contexte historique et sociologique : au début des années 2000, la démocratisation des technologies de séquençage ADN fait que le séquençage de l'ADN est un outil de plus en plus utilisé en taxonomie. En 2003, Hebert introduit la technique de barcoding, basée sur le fragment du gène CO1, pour identifier les spécimens à partir d'une base de données de référence. Très vite, l'utilisation du barcoding va dépasser le cadre de l'identification de spécimens inconnus et le CO1 sera de plus en plus utilisé également comme source de caractères pour délimiter les espèces et construire des phylogénies. Dans ce chapitre, le terme barcoding fait référence à l'idée de standardisation du séquençage de spécimens et à leur intégration dans une base de données (voir aussi Puillandre, 2012). Quant à la distinction entre espèce et taxon, elle n'est pas anodine car dans la pratique, l'activité de taxinomie

intégrative consiste surtout à délimiter les espèces, en utilisant surtout des séquences barcode ADN (Pante *et al.*, 2014; Kekkonen & Hebert, 2014; Goldstein & DeSalle, 20110; Hebert *et al.*, 2003). Et dans la théorie, la délimitation d'espèces est tout sauf une affaire anodine.

### **1. La réalité de l'espèce et son importance dans l'histoire de la taxonomie**

*« Le mot espèce est, de tous, celui qui revient le plus souvent dans les livres d'Histoire naturelle. A quelle page ne l'y trouve-t-on pas ? Et quelle notion peut être dite fondamentale à meilleur droit que celle qu'il exprime ? Il est le premier et le dernier mot de l'Histoire naturelle, et le jour où nous en serions complètement maîtres, nous serions bien près de le devenir de la science entière.*

*Malheureusement, que d'interprétations diverses de ce mot ! Que d'acceptions, tantôt vagues et confuses, tantôt plus nettes, mais inconciliables, tour à tour ou simultanément admises : autrefois, faute de définition; aujourd'hui, parce que la science flotte entre des définitions multiples et contraires ! ».* – Isidore Geoffroy Saint-Hilaire (1859)

« Mystère des mystères » pour certains (Herschel, 1836), « le secret que Dieu s'est réservé », disent d'autres (Duruy, 1865), décidément, la notion d'espèce a fait couler beaucoup d'encre. A tel point que certains auteurs considèrent que cette notion est l'une des plus débattue en biologie, au coude à coude avec les notions d'homologie et de caractère (de Pinna, 1999). « Enoncer clairement ses opinions sur la nature de l'*espèce* est pour un naturaliste l'épreuve la plus redoutable de toutes » nous dit Candolle (1865). A force d'exprimer ses opinions, « des définitions d'espèces, il y en a autant que de naturalistes » répond Trémaux (1865). Pour tout systématicien, s'attaquer à la notion d'espèce devient un rite d'initiation, un apprentissage, ou même simplement un prétexte, qui amène à se poser des questions fondamentales sur la nature des objets qu'il étudie. Se poser des questions sur la nature de l'espèce, ou sur la définition de l'espèce, c'est se poser des questions sur le statut ontologique et épistémologique des concepts scientifiques.

### **2. Ontologie et épistémologie**

Je vais partir du principe, un peu naïf, que l'ontologie, dans le domaine des Sciences, revient à se poser la question « Qu'est-ce qu'il y a ? » et que l'épistémologie revient à se poser la question « Comment le savons-nous ? ». Autrement dit l'ontologie est l'étude de ce qui existe dans le monde réel (domaine des choses concrètes, c'est-à-dire des phénomènes), et l'épistémologie est l'étude de l'accès à cette connaissance. Il apparaît très vite que ces deux

questions sont indissolublement liées, et qu'il est nécessaire de se les poser toutes les deux à la fois, construire une synthèse « ontoépistémologique » (Moulines, 1998). En effet, pour reprendre le principe de la subordination de l'ontologie à l'épistémologie cher à Moulines, nous ne pouvons pas savoir de quel genre de chose nous parlons si nous ne savons pas *de quelle manière nous savons* de quoi nous parlons (et inversement). Par exemple si je souhaite établir une ontologie de l'espèce je dois d'abord savoir comment je vais pouvoir savoir si mes espèces sont correctes. Sans cela, mon ontologie pourrait être élégante, avec une certaine cohérence interne, mais sans aucun lien quelconque avec notre accès à la connaissance du monde, et donc ne serait d'aucune utilité scientifique. Pour arriver à une ontologie correcte, il est donc nécessaire d'avoir accès épistémiquement à l'ontologie définie.

Cette intrication obligatoire entre ontologie et épistémologie a des conséquences sur la construction des modèles ontologiques. Les modèles sont des structures permettant d'approcher la réalité du monde et sont composés de deux choses : (i) des objets ou des ensembles d'objets postulés comme réels (ex : des individus, des cellules, des caractères, des points dans l'espace) et (ii) les rapports ou les relations entre ces objets (ex : lien généalogique, distance entre des points dans l'espace). En acceptant que les catégories ontologiques que l'on définit puissent être accessibles épistémiquement, il faut accepter également que lors de la construction d'un modèle ontologique, les objets ou les rapports/relations entre ces objets du modèle ontologique soient de même nature que les objets et les rapports/relations accessibles épistémiquement. Par exemple, dans le modèle théorique de Hardy-Weinberg, il existe un certain nombre d'*objets* de départ postulés, comme des individus, des loci, des génotypes, des allèles, ainsi que des *rapports* de fréquences, par exemple entre les proportions d'allèles. Tous les objets et les rapports du modèle théorique d'Hardy-Weinberg sont accessibles épistémiquement parlant, et donc il est possible par des études empiriques de vérifier si ce modèle est correct ou non.

Dans le cas de sciences dites « historiques », comme c'est le cas de la phylogénétique par exemple, la construction de modèles ontologiques est une tâche d'autant plus complexe que l'accès épistémique aux objets et de leurs relations historiques est très limité. Le débat sur la nature des espèces est probablement aussi ancien que le concept d'espèce lui-même, mais l'établissement de systèmes ontoépistémologiques sur la nature des espèces, notamment en lien avec la théorie de l'évolution, n'a pris son essor qu'à partir de la fin du siècle dernier.

### **3. « Définition » et « critères » d'espèce**

Au fil de la littérature sur l'espèce, de nombreux concepts différents ont commencé à émerger en fonction des propriétés considérées pour les caractériser. Le plus connu étant probablement le concept biologique de l'espèce de Mayr (1942), fondé sur l'inter-reproductibilité : « les espèces sont des groupes de populations naturelles s'entrecroisant qui sont reproductivement isolées d'autres groupes de ce type ». Toutefois le principe d'inter-reproductibilité avait déjà été proposé par des auteurs prédarwinien (voir Trémaux, 1865). Un autre concept, probablement le plus ancien et le plus utilisé, est celui de morpho-espèce, qui utilise tout simplement les différences morphologiques globales entre individus pour distinguer les espèces. Avec les développements conceptuels (ex : les notions de monophylie ou de phénétique) et techniques (ex : acquisition de données moléculaires ou micro-anatomiques), le nombre de concepts d'espèces n'a cessé d'augmenter, en particulier au 20<sup>e</sup> siècle, si bien qu'en 2011 Wilkins recense pas moins de 26 concepts d'espèces. Tous ces concepts ont bien évidemment été utilisés et revendiqués par différentes « écoles » de biologistes et n'ont fait que complexifier les débats sur la nature de l'espèce.

Avec l'accumulation du nombre de concepts, certains auteurs (ex : Baum & Shaw, 1995 ; Mayden, 1997 ; De Queiroz, 1998) vont prendre conscience de la nécessité de distinguer les concepts épistémiques d'espèces, et le concept ontologique d'espèce, ce qui se retrouve dans la littérature moderne sous les expressions respectives de « critères » (de reconnaissance) et « définition » de l'espèce. Ce distinguo a connu un grand succès auprès de la communauté des biologistes de l'évolution ; l'article de 2007 de De Queiroz, qui reprend principalement ses idées de 1998, est l'article de plus lu de la revue *Systematic Biology*, et probablement l'un des articles modernes les plus cités en théorie de la biologie évolutive. La stratégie de De Queiroz a consisté à proposer le concept de lignée spécifique (Lineage Species Concept), un concept très proche du concept internodal d'espèce (Kornet, 1993), comme étant le modèle ontologique du concept d'espèce et de spéciation. Ainsi, « les espèces sont des fragments du réseau généalogique d'organismes, définitivement divergents des autres fragments » (Samadi & Barberousse, 2006). Dans ce système, les autres concepts d'espèces sont des concepts épistémiques, c'est-à-dire qu'ils proposent des propriétés, ou critères, selon les auteurs, permettant d'accéder à la connaissance des lignées spécifiques. Ce système porte aussi le nom de « concept unifié d'espèce » : le modèle (ontologique) de De Queiroz est le modèle qui s'approche de la réalité de l'espèce, les autres modèles (épistémiques) proposent tous des critères permettant d'approcher notre connaissance de l'espèce réelle. Il n'y a donc



plus de raison que subsistent les querelles de taxonomistes pour déterminer quel est le meilleur concept d'espèces, tous servent une même cause : approcher l'espèce ontologique du modèle des lignées spécifiques.

Le succès grandissant de la taxonomie intégrative est donc principalement dû à deux facteurs : (i) la démocratisation et l'automatisation de l'acquisition des données moléculaires ADN, et (ii) l'intégration de critères issues de multiples concepts d'espèces différents pour l'établissement d'hypothèses d'espèces plus robustes.

#### 4. Opérationnalité et hiérarchisation des critères chez les Turridae

Un pipeline de taxonomie intégrative, intégrant plusieurs critères de reconnaissance des espèces a été mis en place pour les Turridae (Fig. 1). Ce pipeline est largement inspiré de celui de Puillandre *et al.*, (2012b), également mis en place pour les Turridae.

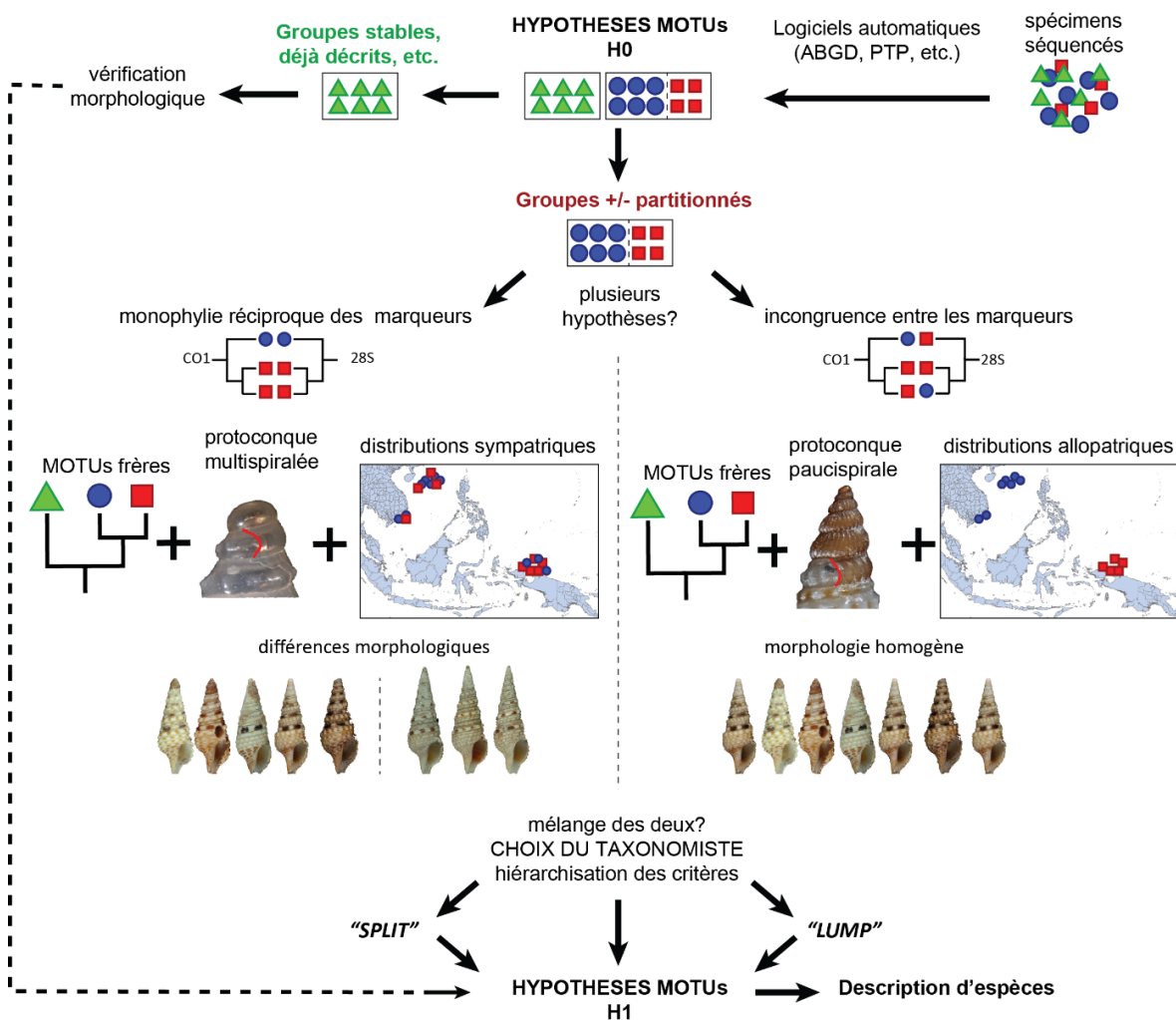


Figure 4 : Pipeline de taxonomie intégrative utilisé dans la thèse. La lecture de ce pipeline commence en haut à droite.

Il est composé de 4 étapes principales :

(i) **Stratégie de séquençage barcoding à l'aveugle.** Cette stratégie repose sur une préparation systématique de tous les spécimens récoltés sur le terrain en vue d'un séquençage du barcode CO1 de retour au laboratoire. Les campagnes TDSB (ex-MUSORSTOM) existent depuis plus de 40 ans, et reposent sur une stratégie d'échantillonnage complète, exhaustive, des habitats (mais sur des zones plus restreintes) par rapport aux campagnes traditionnelles (Bary, 2018). Avec l'arrivée et la démocratisation du séquençage moléculaire de l'ADN, quelques spécimens représentatifs de morpho-espèces ont commencé à être préparés sur le terrain pour du séquençage moléculaire en laboratoire. Ce n'est qu'avec la thèse de Puillandre (2008) que la même stratégie de séquençage exhaustif sur les habitats a aussi été appliquée sur la préparation des spécimens pour du séquençage ADN, conduisant à la découverte d'une diversité cryptique jusqu'alors inquantifiable (Puillandre *et al.*, 2012b). Une illustration de cette diversité cryptique est montée avec les espèces sœurs *Lophiotoma acuta/L. semfala* (Puillandre *et al.*, 2017 - voir Annexe 1).

(ii) **Etablissement des hypothèses primaires d'Unités Taxonomiques Opérationnelles Moléculaires** (MOTU en anglais). Ces hypothèses sont également parfois appelées hypothèses primaires d'espèce dans la littérature (Lorion *et al.*, 2010 ; PSH pour Primary Species Hypotheses en anglais). Il repose sur la création d'une partition du jeu de données de barcode CO1 à l'aide de logiciels automatiques de délimitation d'espèces comme ABGD (Puillandre *et al.*, 2012a), PTP (Zhang *et al.*, 2013) ou GMYC (Monaghan *et al.*, 2009). Cette première étape implique l'application des critères phénétiques (ABGD) ou bien phylogénétiques (PTP, GMYC). Bien souvent dans la littérature (Yeates *et al.*, 2011) la révision nomenclaturale d'un groupe va s'effectuer sur ces hypothèses primaires de MOTUs, avec une discussion et une description des caractères morphologiques soutenant ces hypothèses. Mais l'intégration de critères supplémentaires peut et doit se faire de façon plus intégrative.

(iii) **Utilisation d'autres critères pour rejeter ou corroborer les MOTUs primaires.** Une des priorités est de vérifier si les partitions exprimées par le marqueur barcode CO1 sont congruentes avec d'autres marqueurs. En effet, certaines MOTUs primaires pourraient être la conséquence par exemple d'une structuration géographique, populationnelle, exprimée par le CO1. Pour vérifier cela, les marqueurs doivent être « indépendants », c'est-à-dire localisés dans des régions génomiques ou sur des organites différents que le CO1. Bien souvent donc, on privilégie l'utilisation de marqueurs nucléaires ribosomiques (faciles à séquencer) pour les

mettre en congruence avec le CO1. Chez les Turridae, c'est le marqueur 28S qui a le plus souvent été utilisé (Puillandre *et al.*, 2017 – voir Annexe 1; Puillandre *et al.*, 2012b). Malheureusement, ce marqueur est bien moins variable que le CO1 et, bien souvent, une absence de congruence avec le CO1 pourrait tout aussi bien indiquer une absence de variabilité (et donc d'informativité) pour le 28S qu'un véritable désaccord entre les deux marqueurs. Ces dernières années, la baisse du coût des méthodes de séquençage haut-débit a permis le séquençage de milliers de marqueurs nucléaires indépendants, peu informatifs pris séparément mais permettant des résultats robustes une fois combinés (voir aussi Chapitre 2). Une approche de type RAD-Seq a été utilisée chez les Turridae chez le complexe *Iotyrri/Xenuroturris* (Abdelkrim *et al.*, 2019b - voir Annexe 2). Cependant les méthodes de séquençage haut-débit restent encore très coûteuses et ne peuvent pas (encore ?) être appliquées à l'ensemble des spécimens récoltés sur le terrain. Une stratégie consiste donc à inférer s'il existe potentiellement une structuration populationnelle ou non en combinant deux critères de nature différente : un critère écologique, comme le type de vie larvaire lui-même inféré par le type de protoconque, et un critère géographique (ou bathymétrique) comme la distribution des MOTUs (Fig. 2; mais voir la critique sur l'estimation de la distribution géographique dans un contexte marin dans l'Introduction). La combinaison de ces deux critères est un argument non-moléculaire fort pour décider du rejet ou de la corroboration des MOTUs établies sur la base du CO1. En effet, si la larve est planctotrophe, et que des MOTUs proches génétiquement ont une répartition sympatrique, mais montrent quand même une structuration génétique, alors il y a de fortes chances que les deux MOTUs correspondent à deux espèces distinctes, qui n'échangent plus de matériel génétique. Inversement, si les MOTUs proches ont une larve non-planctotrophe, et que les individus sont répartis dans deux régions géographiques ou bathymétriques (Fig. 2, 3) différentes, alors il y a moins de certitude sur le fait que ces MOTUs constituent des espèces potentielles, car la structure génétique pourrait correspondre en réalité à une structure populationnelle géographique. La validité de ce raisonnement a pu être testée avec le complexe *Iotyrri/Xenuroturris* déjà évoqué plus haut. Chez ce groupe, tous les spécimens possèdent des protoconque multispiralées, indiquant un mode de vie planctotrophe. Le CO1 indiquait deux MOTUs présentant une structuration géographique Ouest-Indien/Centre Indo-Pacifique chez les espèces *X. legitima* et *I. cingulifera*, et deux MOTUs ne présentant pas de structuration géographique chez l'espèce *I. musivum*. L'analyse RAD-seq n'a pas permis de retrouver la structuration géographique chez *X. legitima* et *I. cingulifera*, au contraire de la structuration chez *I. musivum*, aboutissant à la description d'une nouvelle espèce, *I. conotaxis*. Enfin la morphologie est également un critère

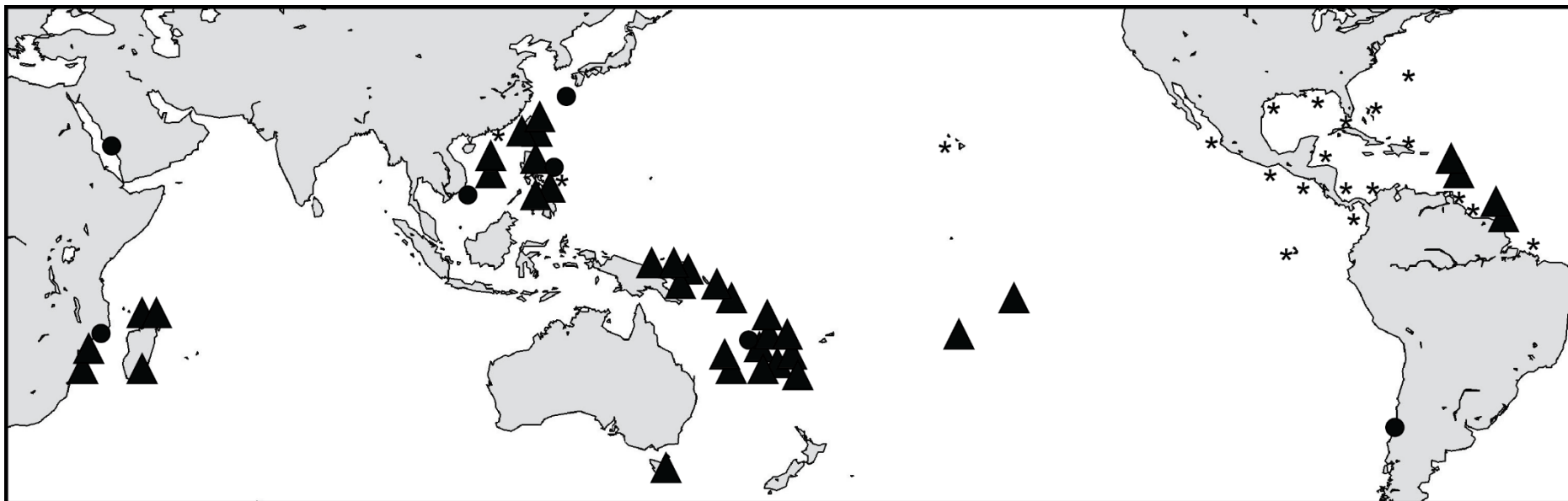


Figure 2 : Carte indiquant l'origine géographique des spécimens séquencés au moins pour le CO1 et utilisés dans cette thèse. Les triangles indiquent une campagne référencée TDSB, alors que les ronds indiquent les spécimens issus de campagnes non-référencées. Les étoiles indiquent les spécimens issus de GenBank ou bien d'autres institutions.



Figure 3 : Graphique en chandelier japonais de 141 MOTUs de Turridae. Les MOTUs font partie uniquement de campagnes référencées, et ont été sélectionnés uniquement s'ils présentaient au moins une valeur minimum et une valeur maximum de profondeur, correspondant aux valeurs de début et de fin d'une opération de dragage/chalutage/plongée. Les barres pleines indiquent l'intervalle de confiance pour lequel nous sommes certains de la profondeur (la valeur maximale de toute les valeurs de profondeur minimale et les valeur minimale de toutes les valeurs de profondeurs maximales). Les traits représentent les valeurs extrêmes de profondeur enregistrées pour chaque MOTU.

utilisé (voir discussion sur *Lophiotoma abbreviata/L. brevicaudata* dans Puillandre *et al.*, 2017 – voir Annexe 1) mais peu fiable (voir Introduction, mais voir également article 1 dans ce Chapitre). Les données anatomiques pourraient être utilisés, mais l'acquisition de ce type d'information est long et fastidieux, contrairement à tous les autres critères évoqués plus haut. De plus, comme présenté dans l'Introduction, ces données peuvent également être non-fiables chez les Turridae. Lorsque tous ces critères ne sont pas suffisants pour trancher sur le rejet ou le maintien d'une MOTU initiale, le taxonomiste doit alors prendre une décision, le plus souvent basée sur son expérience. La stratégie choisie ici est conservatrice (« lumper ») et consiste à ne pas « surdécrire » la biodiversité : s'il y a une incertitude sur la séparation ou non d'un groupe en une ou plusieurs MOTUs, le choix a été fait de n'en garder qu'une. Le choix final de la partition du groupe d'étude par le taxonomiste résulte alors en la proposition de MOTUs secondaires (ou hypothèses secondaires d'espèces).

(iv) **Révision nomenclaturale du groupe d'étude.** Elle consiste en la description des MOTUs délimités et l'attribution de noms portés par des spécimens-types aux hypothèses secondaires de MOTUs proposées par le taxonomiste. Lorsqu'aucun nom n'est attribuable à un MOTU, alors de nouveaux noms d'espèces peuvent être proposés, suivis d'une description de l'espèce, si possible diagnostique. Cette étape n'est pas indispensable à la démarche de taxonomie intégrative et à la recherche de MOTUs. Mais bien que facultative elle reste fortement désirable, pour des raisons de communication entre chercheurs d'autres disciplines, qui basent leurs études sur la base d'espèces nomenclaturales recensées, et non d'études de taxonomie intégrative. Un des meilleurs exemples est le domaine de la conservation, où les espèces en danger ou en voie de disparition sont listées selon la nomenclature en vigueur. Cependant l'attribution des types porte-nom est une étape longue et fastidieuse, car elle demande une expertise et un accès aux collections où sont localisés les types nomenclaturaux (Pante *et al.*, 2014). De plus, l'attribution d'un type à une hypothèse de MOTU est bien souvent une tâche complexe : la plupart de ces types ne sont pas séquençables et/ou les parties molles sont absentes et/ou les coquilles sont dans un mauvais état et/ou les spécimens ont été perdus (il ne reste plus que leur description, parfois minime). Des solutions existent, par exemple via une automatisation de l'acquisition de données morphométriques sur la coquille afin d'attribuer les coquilles de types à un groupe de coquilles séquençées (voir article 1 dans ce Chapitre). Une solution encore plus efficace, mais encore à l'état expérimental, serait le séquençage de l'ADN directement à partir des coquilles (Villanea, Parent & Kemp, 2016; Der Sarkissian *et al.*, 2017).

Au niveau opérationnel, la critique majeure qui peut être faite sur ce pipeline porte sur le point de départ : que se passe-t-il si les séquences CO1 ne sont pas suffisamment variables pour distinguer des MOTUs dans un sous-groupe ? Utiliser un autre marqueur ? Comme dit plus haut, les marqueurs ribosomiques et nucléaires connus sont bien moins variables que le CO1, et les autres marqueurs mitochondriaux ne sont pas plus variables que le CO1. Et comme dit plus haut également, les techniques de séquençage haut-débit ne sont pas encore accessibles économiquement pour tous les spécimens. Une solution intermédiaire consisterait en l'identification et le séquençage d'une dizaine de marqueurs indépendants et le plus informatifs possible, qui seraient obtenus par séquençage haut-débit. Mais la recherche de tels marqueurs chez les Turridae et plus généralement chez les Gastéropodes pose problème, et sera détaillée dans le Chapitre 2.

## **5. Critique**

Nous l'avons vu dans les parties précédentes, l'espèce se construit à différents niveaux.

Au niveau métaphysique, par la création d'un modèle ontologique permettant de contextualiser l'activité de recherche en proposant une représentation (incomplète) de la réalité des patrons et processus biologiques. Le systématicien commence donc son travail de scientifique en se positionnant métaphysiquement (ex : position réaliste sur les espèces). C'est l'étape de définition de l'espèce.

Au niveau méthodologique, pour l'identification de critères et de propriétés permettant d'atteindre cette réalité fixée par le modèle ontologique. C'est l'étape de délimitation de l'espèce.

Au niveau technique, pour l'application des critères définis par la méthode. C'est l'étape de caractérisation de l'espèce.

Cet « axe de la recherche scientifique », ayant pour point de départ un postulat métaphysique aboutissant à une découverte scientifique s'illustre très bien en physique. Un parallèle grossier peut être effectué avec la découverte du Boson de Higgs, dont la détection expérimentale n'a été possible qu'en 2012, mais dont le mécanisme est établi depuis 1964, mécanisme reposant lui-même sur un modèle ontologique qui est le modèle standard de la physique des particules. On peut également citer les prédictions d'Einstein sur les ondes gravitationnelles ou plus loin encore les prédictions de Mendeleïev sur certains éléments du tableau du même nom. La physique, la chimie, et probablement d'autres sciences

mathématisées, expérimentales et ahistoriques se prêtent bien à la recherche scientifique allant dans le sens « ontologie → épistémologie ». Mais il semblerait que ce soit l'inverse en systématique, ou plutôt que les systématiciens soient sans arrêt à la recherche d'une ontologie satisfaisante. C'est le cas des espèces et des taxons, dont certains sont déjà stables et identifiées depuis l'Antiquité (ex : Mammalia, l'Homme), mais dont les critères de reconnaissance n'ont pas toujours été établis, et surtout dont le modèle ontologique était absent, si l'on exclut l'idée de la création divine comme modèle. Pire encore, il apparaît que les modèles ontologiques se soient succédés (ex : théorie de la dégénération de Buffon, transformisme de Lamarck) mais que les entités primairement caractérisées se soient maintenues au fil du temps, ce qui a probablement conduit Jean Rostand à proposer le célèbre aphorisme « le biologiste passe, la grenouille reste ».

Dans le cadre de la taxonomie intégrative, il semblerait que ce soit l'acquisition de nouvelles techniques (l'acquisition de données moléculaires par exemple) qui a conduit à formaliser un peu mieux les critères de reconnaissance des entités à délimiter. Et il semblerait que ça soit l'accumulation de nouveaux critères qui a conduit certains chercheurs à proposer un modèle ontologique pour expliquer les différences entre ces critères et les résultats qu'ils produisent. Le problème, c'est qu'en inversant l'axe de la recherche scientifique, il en est résulté un véritable « capharnaüm ontoépistémologique » que la taxonomie intégrative illustre bien. Des méthodes, issues de théories et d'écoles de pensée différentes, comme la cladistique ou la phénétique, censées donner des résultats différents pour des raisons méthodologiques et théoriques, se côtoient dans les pipelines (ex : ABGD et PTP). Des concepts, comme la monophylie, sont utilisés comme critères alors que le modèle ontologique de De Queiroz ne requiert pas de monophylie (et assume la paraphylie de la lignée ancestrale, puisque par convention l'espèce ancestrale doit s'éteindre pour donner naissance à deux ou plusieurs nouvelles espèces). Enfin, si l'on s'attarde sur les articles qui ont tentés de formaliser un peu mieux la définition de l'espèce (Kornet, 1993 ; Samadi et Barberousse, 2006 ; Alexander *et al.*, 2015), seuls sont considérés comme objets les individus et comme relations les relations généalogiques entre ces individus. Cependant aucun modèle épistémologique (qu'il soit cladistique, phénétique, etc.) n'utilise comme relation les liens généalogiques entre individus, mais tous sont fondamentalement basés sur le lien entre individus et caractères. Or, nous l'avons vu dans la partie 2, il est nécessaire que les objets et les relations entre objets du modèle ontologique soient de même nature que les objets et les relations entre objets de modèles épistémologiques associés. Face à ce constat, quelques voix s'élèvent pour tenter de réinjecter des concepts comme l'écologie ou le phénotype dans la définition d'espèce, sans



pour autant présenter un modèle ontologique formel (ex: Freudenstein *et al.*, 2016). Deux solutions sont alors possibles. Soit les modèles épistémologiques doivent s'adapter au modèle ontologique et dans ce cas il faut développer des méthodes de délimitation d'espèces basées sur les individus et les liens généalogiques. Il existe bien les méthodes de coalescence (Yang & Rannala, 2010 ; Fujita *et al.*, 2012) reposant sur des individus et des liens généalogiques, mais entre séquences, ce qui fait tout de même appel à une notion de caractère. Soit il faut développer un nouveau modèle ontologique, fondé non plus sur les individus, leurs relations généalogiques et le réseau généalogique, mais sur les individus, les caractères, et la nature historique de ces caractères. Là encore, un modèle ontologique de coalescence pourrait être la solution, mais il faudrait ajouter à ce modèle d'autres concepts aux généalogies de séquences tels que la monophylie (et donc par extension les synapomorphies) pour justifier l'activité de taxonomie intégrative, qui utilise (entre autres) les deux approches. Or les approches de coalescence pour délimiter les espèces ne requièrent ni la monophylie réciproque des allèles, ni le concept de synapomorphie (Fujita *et al.*, 2012). L'utilisation de modèles épistémologiques de natures différentes semble donc être un obstacle à la création d'un modèle ontologique adapté à la taxonomie intégrative.

D'un côté donc, je suis tenté de qualifier ma pratique de taxonomie intégrative comme issue de l'école de pensée « it-doesn't-matter-very-much », pour reprendre une expression de Felsenstein (2004) à propos d'une autre critique (sur les classifications), à défaut d'être convaincu par le modèle ontologique proposé. De l'autre, la taxonomie intégrative se trouve être une méthode redoutablement efficace pour proposer des unités opérationnelles, et reste très convaincante, empiriquement parlant. Je propose d'ailleurs de suivre ci-après un cas appliqué sur le genre *Cryptogemma*. Cette étude constitue l'article 1 de ma thèse. J'y propose l'intégration semi-automatisée de données morphométriques sur la coquille, critère qui pourrait être ajouté de manière systématique au pipeline (Fig. 1) à l'avenir.

## **6. Un exemple de l'application du pipeline de taxonomie intégrative chez le genre *Cryptogemma***

**Soumis au Zoological Journal of the Linnean Society**

## **Just the once will not hurt: DNA suggests species lumping over two oceans in deep-sea snails**

Paul Zaharias<sup>1\*</sup>, Yuri I. Kantor<sup>2</sup>, Alexander E. Fedosov<sup>2</sup>, Francesco Criscione<sup>3</sup>, Anders Hallan<sup>3</sup>, Yasunori Kano<sup>4</sup>, Jérémie Bardin<sup>5</sup>, Nicolas Puillandre<sup>1</sup>.

<sup>1</sup>Institut Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 43 rue Cuvier, CP 26, 75005 Paris, France.

<sup>2</sup>A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninski prospect 33, 119071 Moscow, Russian Federation

<sup>3</sup>Australian Museum Research Institute, Australian Museum, Sydney, Australia

<sup>4</sup>Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564

<sup>5</sup>Centre de Recherche en Paléontologie – Paris (CR2P-UMR 7207), Sorbonne Université-CNRS-MNHN, Site Pierre et Marie Curie, 4 place Jussieu, Paris Cedex 05, France.

### Abstract

The practice of species delimitation using molecular data commonly leads to revealing species complexes and increasing the number of delimited species. In few cases, however, DNA-based taxonomy can lead to lumping previously described species. Here, we delimit species in the genus *Cryptogemma* (Gastropoda, Conoidea, Turridae), a group of deep-sea snails with wide geographic distribution, by primarily using the mitochondrial COI gene. Three approaches of species delimitation (ABGD, mPTP and GMYC) were applied to define species partitions. All approaches resulted in eight species. According to previous taxonomic studies and shell morphology, 23 available names potentially apply to the eight *Cryptogemma* species that were recognized herein. Shell morphometrics, radular characters, as well as geographic and bathymetric distributions were used to link type specimens to these delimited species. In all, 23 of these available names are here attributed to seven species, resulting in 16 synonymizations, and one species is described as new: *C. powelli* sp. nov. We discuss the possible causes behind such an apparent ‘over-description’ of species within *Cryptogemma*, which is here shown to constitute a rare case of DNA-based species lumping in the hyperdiversified superfamily Conoidea.

## Introduction

The advent of integrative taxonomy (Dayrat, 2005; Will, Mishler & Wheeler, 2005), mainly driven by the molecular revolution, has led to a great remodeling of the practice of species delimitation, extending the use of standardized approaches in the field of taxonomy (e.g. Lefébure *et al.*, 2006). Several resulting studies, often primarily based on molecular data, have revealed numerous species complexes, thus increasing the number of delimited species within such groups, commonly referred to as “cryptic species” (Bickford *et al.*, 2006; Fiser, Robinson & Malard, 2017). The difficulty in assigning names and/or describe these newly delimited species can constitute an impediment, especially for non-taxonomists. Indeed, when new species are discovered, nearly half of them remain undescribed, at least not described in the publication where they are revealed (Pante, Schoelinck & Puillandre, 2014). Thus, the splitting effect of integrative taxonomy is not accompanied by adequate taxonomic effort, resulting in a plethora of undescribed species. It has been suggested that such lack of appropriate taxonomic procedure can have consequences in various fields, such as ecology and conservation (e.g. Iglésias *et al.*, 2010).

Conversely, there have been cases where integrative taxonomy, by incorporating molecular data, have led to the lumping of species that have been originally described using morphological and anatomical characters only (e. g. Chan, Grismer & Brown, 2018; DeJaco *et al.*, 2016). Such instances may occur when, contrary to the occurrence of “cryptic species”, intra-specific morphological variability is equal to or greater than inter-specific variability (e.g. Puillandre *et al.*, 2010). These phenomena may have biological causes (e.g. sexual dimorphism, allometry during ontogeny, phenotypic plasticity), or in other cases simply be artefacts of a particular taxonomic practice (some taxonomists tend to split more than others). Unfortunately, these cases are difficult to investigate for taxa where morphological characterization is not formalized, genomic resources are poor, and/or known specimens are rare and deposited in different institutions. Nevertheless, investigating whether or not different forms correspond to different species should rely on a combination of modern species delimitation approaches and solid taxonomic expertise in order to clarify both boundaries and taxonomic status of a given taxon.

The Conoidea (Neogastropoda) is an extremely diversified lineage of marine gastropods, traditionally divided into Terebridae, Conidae and Turridae (Bouchet *et al.*, 2011). Recent molecular phylogenies have significantly redefined boundaries within the superfamily (Puillandre *et al.*, 2011; Abdelkrim *et al.*, 2018a), leading to a redefinition of turrids. The family Turridae H. Adams & A. Adams, 1853 (1838) is now a monophyletic

group that encompasses a few morphologically well-defined monophyletic genera, such as *Gemmuloborsonia*, *Polystira*, *Lophiotoma* or the “*Xenuroturrus/Iotyrris* complex”, that have been revised in recent years (Puillandre, Cruaud & Kantor, 2010; Todd & Rawlings, 2014; Puillandre *et al.*, 2017; Abdelkrim *et al.*, 2018b). In some cases, the integrative taxonomy approach has uncovered “cryptic species”, that were described in the same studies in which they were revealed. However, a considerable proportion of turrid species are still left in the paraphyletic genus *Gemmula*, for which it was estimated that less than half the number of species have been described so far (Puillandre *et al.*, 2012). Puillandre *et al.* (2012) also called for the complete revision of *Gemmula*, and suggested that the high number of species and the suspected global morphological stasis of the group are likely explanations as to why this group remains unrevised until now. Moreover, revising the *Gemmula* group would be tantamount to revising the entire Turridae, given that independent “*Gemmula*-like” lineages are distributed all over the turrid tree (Puillandre *et al.* 2012). Finally, while a high number of undescribed species was estimated in *Gemmula*, it is not clear whether all the independent “*Gemmula*-like” lineages equally include undescribed species, or if some lineages will require splitting of previously described species, while others will imply lumping of previously described species, or a combination of both splitting and lumping. Hence, we propose to partition the taxonomic revision by focusing on smaller monophyletic groups that would be revised one by one, in order to progressively resolve the taxonomic predicament of *Gemmula* as it presently stands. The first group we identified is a clade of deep-water species, hereafter named “*Cryptogemma*”, that has already been used in a case-study to illustrate congruent species hypotheses in the integrative taxonomy framework designed in Puillandre *et al.* (2012, Figure 5), comprising five species of Turridae attributed to three genera. We completed the dataset from Puillandre *et al.* (2012) with recent field sampling and material from museums, and applied an integrative taxonomic approach primarily based on the analysis of barcode fragment of *COI* with additional data on the geography and bathymetry included *a posteriori* to help attributing existing names to the molecular species hypotheses. Furthermore, we significantly improved resolution of morphological studies by carrying out a formalized analysis of the morphology of teleoconch and protoconch, and by examining radulae in all sequenced species. The sequencing of a few type specimens and the formalized morphological analysis allowed us to attribute confidently type specimens to molecular species. Our results enabled a comprehensive revision of *Cryptogemma*: rather than finding the usual splitting effect of molecular-based integrative taxonomy, our study led to an unexpected number of new synonymies and only one new species description.

## Material and Methods

### *Sampling*

As a part of our background activity, most neogastropods collected during field expeditions organized by the Muséum national d'Histoire naturelle (MNHN) are processed for DNA extraction and sequencing of the Barcode fragment of the *COI* gene. Among the processed material, all specimens for whom the barcode sequences clustered in the *Cryptogemma* clade were included in this study. These samples were collected during field expeditions in the Indo-Pacific and West Atlantic: EBISCO and KANADEEP in the Chesterfield Islands, CONCALIS, EXBODI, KANACONO, TERRASSES and NORFOLK 2 in New-Caledonia, MIRIKY and ATIMO VATAE in Madagascar, MAINBAZA in Mozambique, BIOMAGLO in Mayotte, BIOPAPUA, KAVIENG, MADEEP and PAPUA NIUGINI in Papua New Guinea, AURORA 2007 and PANGLAO 2004 in the Philippines, SALOMON 2 and SALOMONBOA 3 in the Salomon Islands, TARASOC in the Société-Tuamotu islands, TAIWAN 2013, NANHAI 2014, DONGHA 2014 and ZHONGSHA 2015 in Taiwan, GUYANE 2014 in French Guyana. Additional samples were collected during Japanese expeditions: T/V “*Nagasaki-maru*” cruise N275 and R/V “*Tansei-maru*” cruise KT-12-32 in Japan. We completed this exhaustive sampling with specimens from the Australian museum (AMS), Zoological Museum of Moscow State University (ZMMU) and the National Museum of Natural History (USNM) that could potentially belong to this clade, based on their morphology and the available corresponding sequence in GenBank. Twelve specimens from the AMS were collected during the IN2017\_V03 - Sampling the Abyss cruise (2017). A paratype of *Ptychosyrinx lordhoweensis* Kantor & Sysoev, 1991 (R/V “*Dmitry Mendeleev*”, 16th cruise, st. 1245, N° Lc14353; Kantor & Sysoev, 1991) from the ZMMU was used for DNA extraction. Samples from USNM were also selected for molecular analyses: three paralectotypes of *Gemmula benthima* Dall, 1908 (USS Albatross; Stations 2807, 3360, 3365, USNM N° 96485, 123087, 123092; Kabat, 1996) and one specimen of *Ptychosyrinx carynae* (Haas, 1949), USNM 832922. For USNM samples, *G. benthima* specimens were preserved dry, while the *P. carynae* specimen was probably fixed in formalin before conservation in ethanol.

MNHN specimens processed before 2012 as well as all AMS specimens chosen for molecular analysis were anaesthetized using the isotonic solution of MgCl<sub>2</sub> prior to fixation in 96% ethanol. Additionally, MNHN specimens processed after 2012 were microwaved and fixed in 96% ethanol (Galindo *et al.*, 2014). Shells were subsequently drilled to extract the retracted animals. Japanese specimens were boiled in 70–90°C water for 0.1–0.5 min before

preservation in 99% ethanol. ZMMU specimens were probably fixed and preserved in 75–80% ethanol.

#### *DNA sequencing*

DNA from MNHN samples was extracted using the Epmotion 5075 robot (Eppendorf), following the manufacturers' recommendations. DNA from AMS samples was extracted from small pieces of foot muscle by use of a Bioline Isolate II Genomic DNA extraction kit for animal tissue, following the standard procedure of the manual. DNA from ZMMU and USNM samples was extracted using the EZNA Mollusc DNA Kit (Omega Bio-Tek Inc.), following the manufacturers' recommendations. For most samples, the barcode fragment (658 bp) of the mitochondrial *COI* gene was amplified using the universal primers LCO1490/HCO2198 (Folmer *et al.*, 1994). PCR reactions were performed using a previously well-established protocol (Puillandre *et al.*, 2017). For the ZMMU and USNM samples, the barcode fragment was amplified in two fragments, one with the LCO1490 and a newly designed primer (conoCOIintR: GCNCATGCHGGNGGNTCWGTW) and the other with another newly designed primer (conoCOIintF: TCWTCAGCTGCNGTWGAAAGNGG) and the HCO2198 primer. Both fragments are overlapping over ~50 nucleotides. The PCR mix was the same as for the LCO1490/HCO2198, except that BSA (10 mg/mL) was used instead of DMSO. The amplification procedure was also similar, except that the annealing temperature was 47°C, with 40 cycles instead of 35. A fragment of the mitochondrial *12S* rRNA gene was amplified using the universal primers *12S1/12S3* (Simon, Franke & Martin, 1991) and PCR reactions performed using a previously established protocol (Puillandre *et al.*, 2011). PCR products were purified and sequenced by the Eurofins sequencing facility. *28S* primers used at AMS were C1'/D2 (Dayrat *et al.*, 2001; Jovelin & Justine, 2001). Amplification of *28S* consisted of an initial denaturation step at 94 °C for 4 min, followed by 30 cycles of denaturation at 94 °C for 30 s, with annealing set to 57°C for 45 s, followed by extension at 72 °C for 1 min. The final extension was at 72 °C for 5 min.

#### *Species delimitation*

*COI* sequences were aligned manually using MEGA X (Kumar *et al.*, 2018); *12S* and *28S* sequences were aligned using Muscle (Edgar, 2004) as implemented in MEGA X and then checked by eye. ABGD, GMYC and mPTP and phylogenetic reconstruction methods were applied to the *COI*, plus a concatenation of the *COI*, *12S* and *28S* alignments with one specimen per species (see Results section). For ABGD, the web version (<http://wwwabi.snv.jussieu.fr/public/abgd/>) and the default parameters were used, with a



Kimura (K80) TS/TV model implemented. Maximum likelihood trees were constructed for *COI*, and the concatenated (*COI + 12S + 28S*) alignments using IQ-TREE v1.6.3 (Nguyen *et al.* 2014), with 100 bootstraps. For all alignments, *COI* was consistently divided into three partitions, corresponding to the three codon positions, and *12S* and *28S* considered each as a single partition. The best model for each partition was evaluated using ModelFinder as implemented in IQ-TREE (Kalyaanamoorthy *et al.*, 2017). The mPTP approach of species delimitation (Kapli *et al.*, 2017) was performed on the *COI* maximum likelihood tree, using both ML delimitation with default parameters and MCMC method with a Markov chain of 200,000,000 generations, a sampling frequency each 2,000 generations and a burnin set at 25%. Bayesian trees were reconstructed using BEAST v2.5.0 (Bouckaert *et al.*, 2019), running 200,000,000 generations on the *COI* dataset with a sampling frequency every 4,000 generations and a burnin set at 10%. The *COI* dataset was partitioned by codon position for model evaluation, and prior values were set as in Puillandre *et al.* (2017). The resulting trees were combined using TreeAnnotator as implemented in BEAST2, with node heights corresponding to mean heights of all trees. The “single” method of GMYC was applied using the resulting tree. Finally, a non-ultrametric Bayesian tree was reconstructed for the *COI*, and concatenated datasets using MrBayes v3.2.6 (Ronquist & Huelsenbeck, 2003), each of the two runs consisting of 8 Markov chains and 20,000,000 generations and a sampling frequency each 2,000. The analysis was performed on the Cipres Science Gateway (<http://www.phylo.org/portal2>). The consensus tree was calculated with burnin set at 20%. Four species of Turridae were used as outgroups: *Turris babylonia* (Linnaeus, 1758), *Gemmula kieneri* (Doumet, 1840), *Unedogemmula unedo* (Kiener, 1839) and *Turridrupa* sp.. The MolD program (Fedosov *et al.*, under review) was used to identify unique combinations of diagnostic sites for each delimited species.

### *Shell and radular morphology*

Radulae were prepared by standard methods (Kantor & Puillandre, 2012) and examined by scanning electron microscopes JEOL JSM 840A and Tescan VEGA II LSU at the MNHN. Protoconchs and shells were measured in standard position and the number of protoconch whorls counted according to Bouchet & Kantor (2004).

Variation of shell shapes was analyzed through their outlines. Initially, we made a selection of shells of sequenced specimens that were not broken, truncated or aberrant. To complete this dataset, we obtained photographs of 20 name-bearing types (corresponding to 21 species names) that were deemed high-quality, of which 18 were suitable for

morphometric analysis (nearly complete shells). In total, shape acquisition was performed for 197 sequenced specimens, 18 types and 8 unsequenced specimens of *C. periscelida*.

Shell shape was acquired from photographs using a Photoshop CC 2019 script. Outlines were reconstructed with the package Momocs (Bonhomme *et al.* 2014) from the R software. We used Elliptical Fourier analysis (Giardina and Kuhl, 1977), which presents many advantages over alternative Fourier analyses (Bonhomme *et al.* 2014), to decompose the closed outlines of the shells into periodic functions. The number of harmonics necessary to describe the shape was selected to reach 99.9% of the cumulative harmonic power, ensuring that the majority of shell shape outlines were considered. Shapes were also checked by eye to ensure that enough complexity was captured. The obtained coefficients were normalized to allow the superposition of the first ellipse and make the shapes comparable. We checked that all the reconstructed outlines were appropriately aligned. A principal component analysis (PCA) was performed on the harmonics coefficients in order to capture an optimal shape variation with a minimal number of principal components. Finally, a linear discriminant analysis (LDA) was performed to find the linear combination of principal components that maximizes difference between specimens from species delimited with DNA. Fourier analysis does not take size into account; only shapes are compared. As the size of the shells may provide some discriminatory information, we added the length of each specimen as a variable for the LDA. The appropriate number of principal components to use in the LDA model was carefully chosen to find optimal equilibrium between maximizing the differences and overfitting. A leave-one-out cross-validation procedure was used to find the number of PC axes leading to the best percentage of species predictions (see Figure S1). In order to test if our species attribution to each molecular species partition was correct, type specimens were not included in the LDA training set but subsequently added in the LDA morphospace. Another reason not to include the types in the training set was to extract at best the morphological differences between molecular species without the influences of type specimens morphology.

All the material examined for this study is summarized in Table S1.

## **Abbreviations**

### *Museums and repositories*

AMS: Australian Museum of Sydney, Sydney, Australia.

AWMM: Auckland War Memorial Museum, Auckland, New Zealand

BPBM: Bernice Pauahi Bishop Museum, Honolulu, USA.

FMNH: Field Museum of Natural History, Chicago, USA.

MNHN: Muséum national d'Histoire naturelle, Paris, France.

NHMUK: Natural History Museum of United Kingdom, London, UK.

USNM: National Museum of Natural History, Smithsonian Institution, Washington, DC, USA.

ZMB: Museum für Naturkunde, Humboldt-Universität, Berlin, Germany.

ZMMU: Zoological museum of Moscow University, Moscow, Russian Federation.

ZSIC: Zoological Survey of India, Calcutta, India.

### *Others*

AL: Shell aperture length.

PD: Protoconch diameter.

PL: Protoconch length.

R/V: research vessel.

St.: station.

## **Results**

The USNM and ZMMU samples yielded poor DNA quality, making these difficult to sequence with a classical Sanger sequencing approach. Nevertheless, the *12S* was obtained for the USNM96485 and USNM123087 paralectotypes collected 128 years ago. For the ZMMU paratype, the USNM123092 paralectotype and the USNM83292 lot, the *COI* was obtained using the conoCOIntR and conoCOIntF newly designed primers. The USNM123092 lot had previously been sequenced and published by Todd & Rawlings (2014) for the *COI*, *12S* and 16S fragments (GenBank accession number KM218745, KM218648 and KM218552), despite the fact that the samples were most probably formalin-fixed. Nevertheless, sequencing an additional specimen from the same lot was necessary as the lot might include specimens from different species.

All ABGD, mPTP and GMYC analysis (see Figure S2) on the *COI* dataset resulted in the same partition with eight species hypotheses.

For this partition, we searched for all names available and potentially applicable (see also the Taxonomy section). By consulting with the relevant literature on Turridae (e.g. Powell, 1964; Tucker, 2004) and by comparison of type specimens or their images to shells of sequenced specimens, we found 23 names (see Taxonomy section) that we applied tentatively to this partition. A total of 18 name-bearing types corresponding to 19 names were of sufficient

quality to be included in the morphometric study. The other name-bearing types were discussed and assigned to molecular species using only the shell photographs or drawings and extrinsic parameters such as the bathymetry or geography (see Taxonomy section).

Despite considerable intraspecific morphological heterogeneity, the morphometric analysis yielded instructive results: using shape alone, the leave-one-out cross-validation procedure successfully predicted correct species attribution of 84% of sequenced specimens and confirmed placement of 14 types out of 18 *a priori* attributed types (see Figure S3). By adding shell size, we increased the score of the leave-one-out cross-validation procedure to 90.64% for sequenced specimens, and 15 out of 18 types (Fig. 1). The three type specimens whose placement was incorrectly *a priori* attributed were *Bathybermudia carynae* (only representative of its species), *Pleurotoma praesignis* and *Ptychosyrinx bisinuata japonica* (see also taxonomic section). The most discriminant axis (Fig. 1, axis 1: 71.92% of trace) corresponds to the variation range of the siphonal canal length and the width of the shell, while variation along the second axis (15,57% of trace) mainly corresponds to the intensity of the curvature of the outer aperture lip.

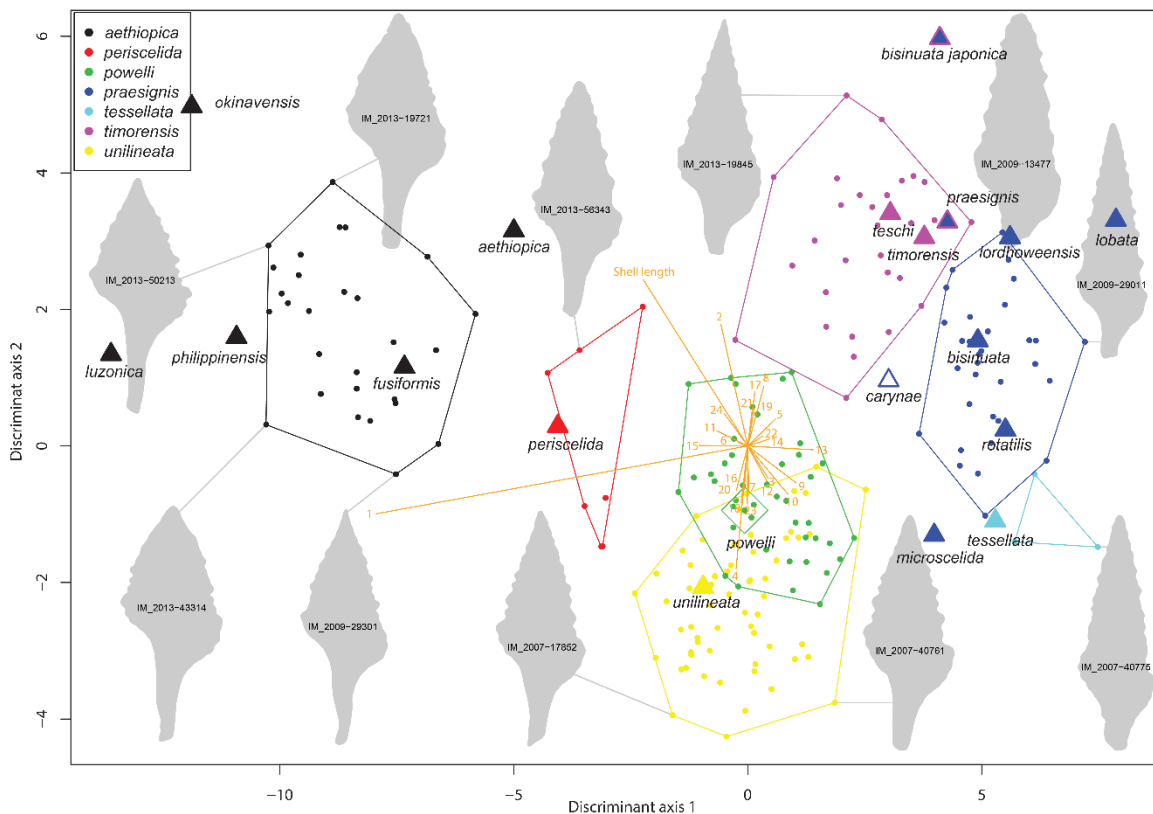


Figure 5 : Projection of specimens along the two first discriminant axes. Convex hulls are drawn to easily observe the range of each species in the morphospace. Triangles are type specimens, with names of species indicated; the background color corresponds to the species attribution from the literature and the border color to the LDA prediction. Type specimens were not included in the LDA training set. Orange segments correspond to the projection of coefficients multiplied by the standard deviation of each raw variable (PC axes and shell length). Outlines of most extreme shapes are drawn with their respective MNHN numbers.

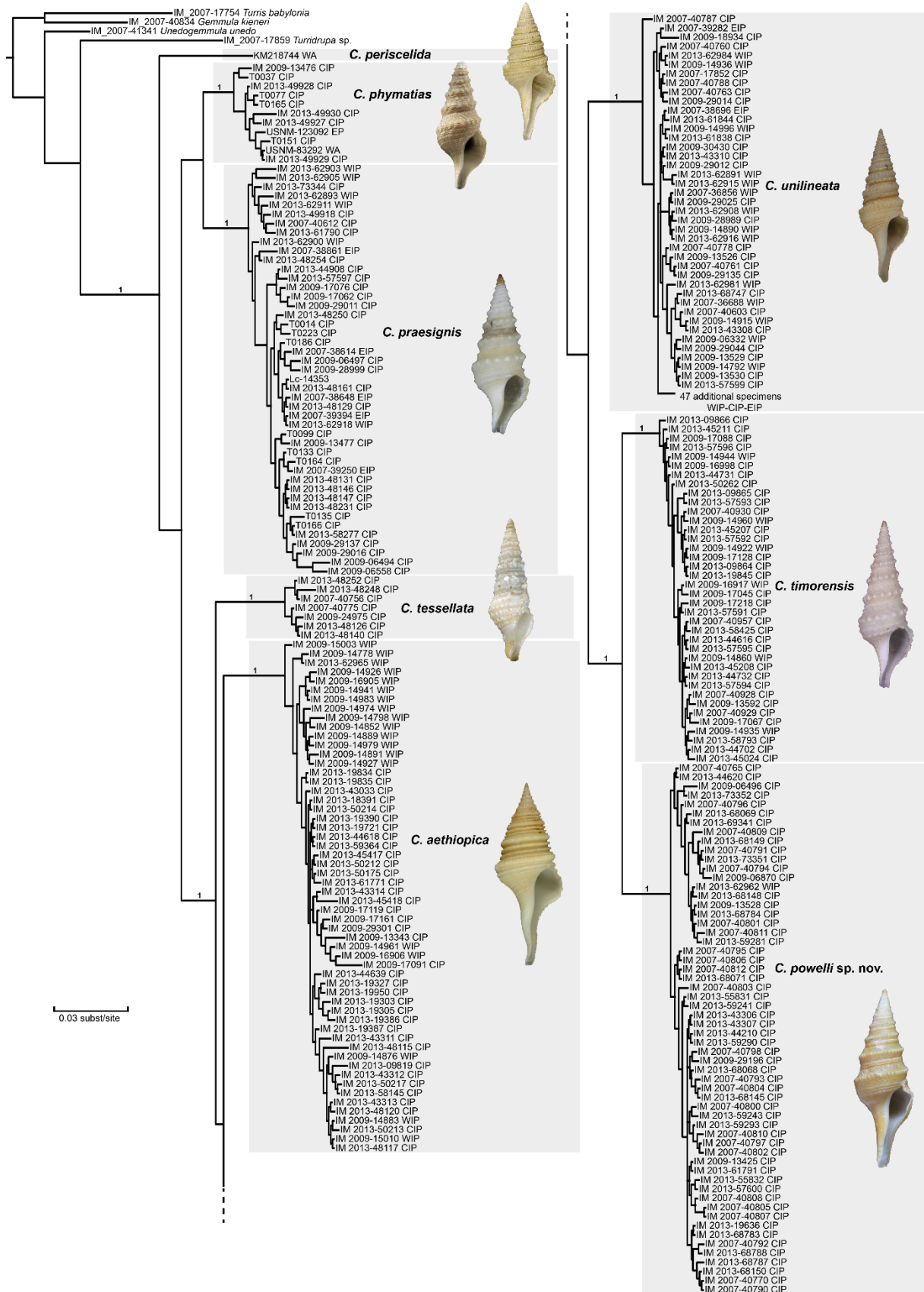


Figure 2: Bayesian Tree (MrBayes) for the COI. Posterior probabilities (>0.95) are only shown for specific and inter-specific nodes. Abbreviations next to each specimen number refer to main marine provinces: WIP: Western Indo-Pacific; CIP: Central Indo-Pacific; EIP: Eastern Indo-Pacific; EP: Eastern Pacific; WA: Western Atlantic. Alignment provided in Appendix 1.

The phylogenetic trees resulting from the *COI* (Fig. 2) and concatenated-gene analyses (Fig. 3) showed poorly supported interspecific relationships, except for the *Cryptogemma phymatias*/*C. praesignis* sister relationship and the (*C. unilineata*, *C. timorensis*, *C. powelli*) clade. Thus, the two molecular species PSH 11 and PSH 14 in the study of Puillandre *et al.* (2012), which were attributed to the genus *Ptychosyrinx* Thiele, 1925, did not form a monophyletic group. The genera *Cryptogemma* Dall, 1918 (type species: *Gemmula benthima* Dall, 1908), *Ptychosyrinx* Thiele 1925 (type species: *Pleurotoma (Subulata) bisinuata* Martens, 1901) and *Pinguiggemmula* McNeil, 1961 (type species: *Pinguiggemmula okinavensis* McNeil, 1960), despite being easily recognizable morphologically, showed a level of molecular divergence among them that was similar to, or even lower than, the level of divergence found among species in most other turrid genera, thus suggesting synonymization of the two latter genera with *Cryptogemma*.

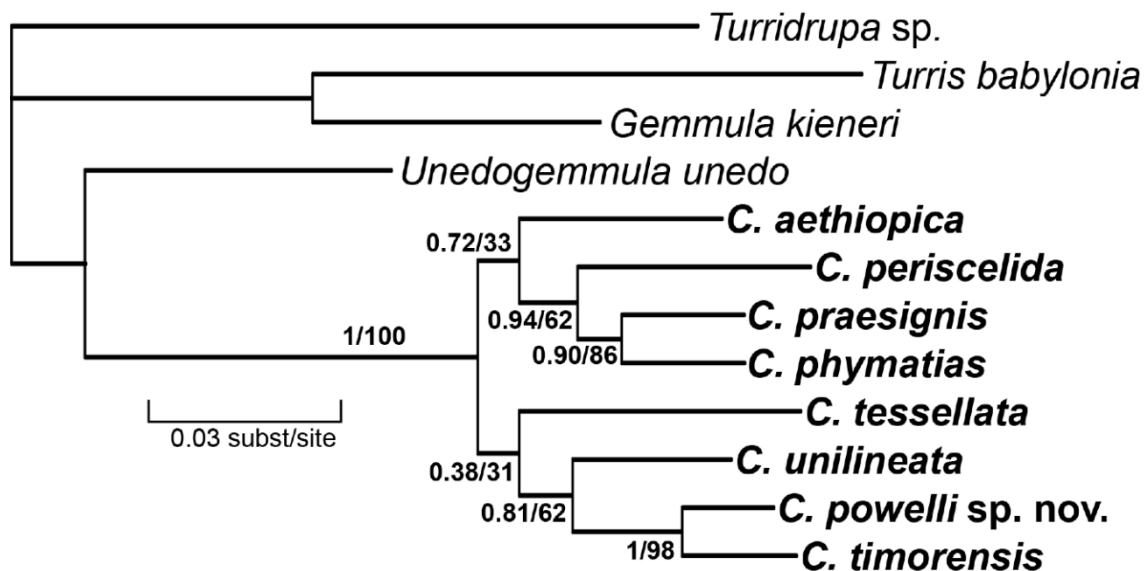


Figure 3: Bayesian Tree (MrBayes) of the *COI*, 12S and 28S concatenated. Posterior probabilities and bootstrap values are shown for each node. Alignment provided in Appendix 2.

All species showed large, oceanic or transoceanic, sympatric distributions. Despite having large sympatric distribution, the *C. phymatias*/*C. praesignis* sister species are found in different bathymetric zones, ~1400–3000 m and ~300–1400 m respectively, with the exception of two specimens, one for each species, found at the same station CP2752 (1378–1436 m) during the AURORA 2007 expedition, off Luzon Island, in Philippines (see Table S1). All species geographic (Fig. 4) and depth (Fig. 5) ranges were represented using only sequenced specimens, except for *C. periscelida*. For the latter, the *COI* sequencing did not work for the 13 MNHN samples available but information on geography and bathymetry was included nonetheless.



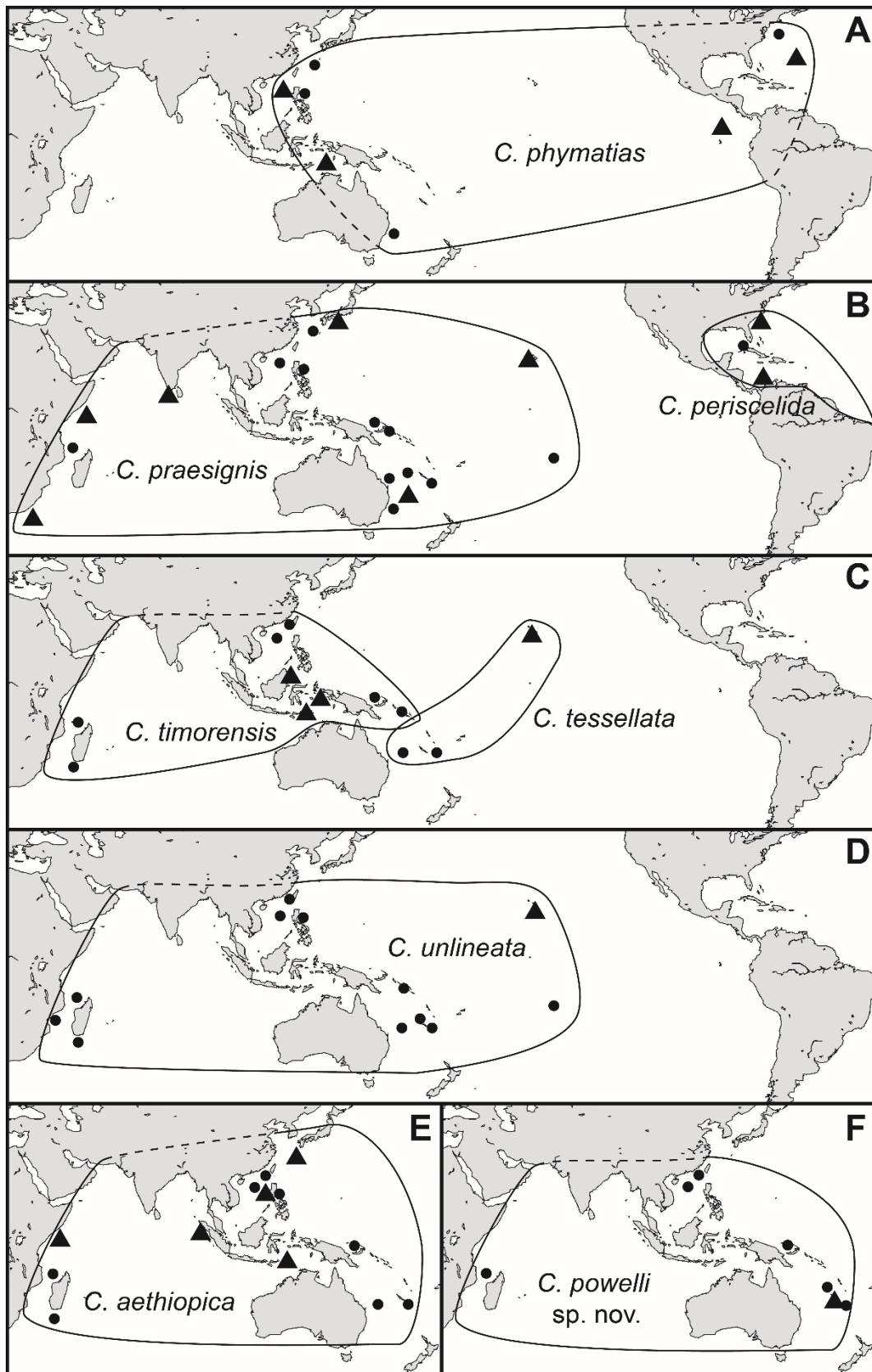


Figure 4: Map showing the species distributions. Filled circles represent localities of sequenced specimens; full triangles represent localities of type specimens.

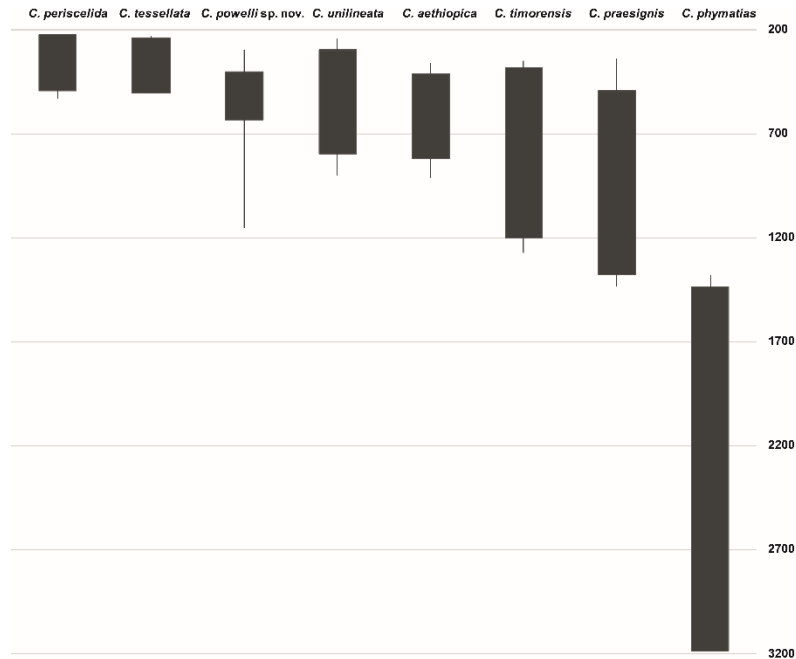


Figure 5: Candlestick chart of depth distributions of the eight species. The high and low value correspond respectively to the minimum and maximum values of depth recorded for each species, while the open and close values correspond to the confirmed minimum and maximum values of depth. Confirmed minimum and maximum values of depth correspond respectively to the highest depth value of ending of trawl and the highest depth value of beginning of trawl.

## Taxonomy

### Superfamily Conoidea Fleming, 1822

#### Family Turridae H. & A. Adams, 1853 (1838)

#### Genus *Cryptogemma* Dall, 1918

Type species: *Gemmula benthima* Dall, 1908 (OD)

*Ptychosyrinx* Thiele, 1925 (type species *Pleurotoma bisinuata* Martens, 1901 – OD)

*Bathybermudia* Haas, 1949 (type species *Bathybermudia carynae* Haas, 1949 – OD)

†*Pinguiggemula* McNeil, 1960 (type species †*Pinguiggemula okinavensis* McNeil, 1960 – OD)

*Included species:* *Cryptogemma aethiopica* (Thiele, 1925); *C. periscelida* (Dall, 1889); *C. phymatias* (Watson, 1886); *C. powelli* sp. nov.; *C. praesignis* (Smith, 1895); *C. tessellata* (Powell, 1964); *C. timorensis* (Tesch, 1915); *C. unilineata* (Powell, 1964).

*Remarks:* According to WoRMS (checked July 2019), *Cryptogemma* comprises 11 species, and has never been revised. A consultation of available type photos and published images on

*Cryptogemma* species indicated that all of 11 species except the type species *C. benthima* and *C. phymatias*, the senior synonym of *C. benthima* should be excluded from the genus, as the majority of them lack key characters such as the narrow fusiform shell and the well-marked peripheral anal sinus. Moreover, studied samples from Japan most certainly corresponding to the *Cryptogemma corneus* as pictured in Hasegawa (2009: figs 335–338) showed closer affinity to other conoidean families (e.g. Horaiclavidae) than to the Turridae, based on both, the sequence of the barcode fragment of *COI*, and radular morphology (results not shown). The eight species of *Cryptogemma* listed herein have wide distributions, most of them covering the whole Indo-Pacific tropical region, and are not found at depths shallower than 200m.

Protoconch consisting of 4 to 5.25 whorls. Shell shape of the teleoconch combined with size is an important factor to distinguish species (see Results section), despite high intra-specific variability. Widest variability range observed in shell proportions, with shells with wide last whorl and long siphonal canal being at one extreme of the range, and more elongated shells with short siphonal canal on the other.

### ***Cryptogemma phymatias* (Watson, 1886)**

(Fig. 6)

*Pleurotoma phymatias* R. B. Watson, 1886: 283, pl. 26, fig. 2. 1920 m, Philippine Islands, 16°42'N, 119°22'E (Expedition H.M.S. *Challenger* on November 13, 1874, st. 205).

*Gemmula benthima* Dall, 1908: 267, pl. 1, fig. 7, pl. 13, fig. 4. 2323 m, Gulf of Panama.

*Gemmula benthina* – Dall, 1918: 318 (lapsus calami); Powell, 1964: 279 (lapsus calami).

*Pleurotoma truncata* Schepman, 1913: 40(404), pl. 26, figs. 1a, b. 2798 m, Banda Sea, Indonesia, 6°24'S, 124°39'E.

*Bathybermudia carynae* Haas, 1949: 4, figs. 3, 4. 3109 m, off Bermuda, 32°08.2'N, 64°33'W.

*Remarks:* This species was not present in the sampling of Puillandre *et al.* (2012) but was included in the phylogeny of Puillandre *et al.* (2011) under the name *Ptychosyrinx carynae*.

Due to the frequent loss of the shell apex, there are very few available protoconchs for this species. Only the eroded protoconch of the type specimen of *Bathybermudia carynae* could be measured, with PD= 1.3 mm and PL > 1.6 mm, consisting of more than 3 whorls.

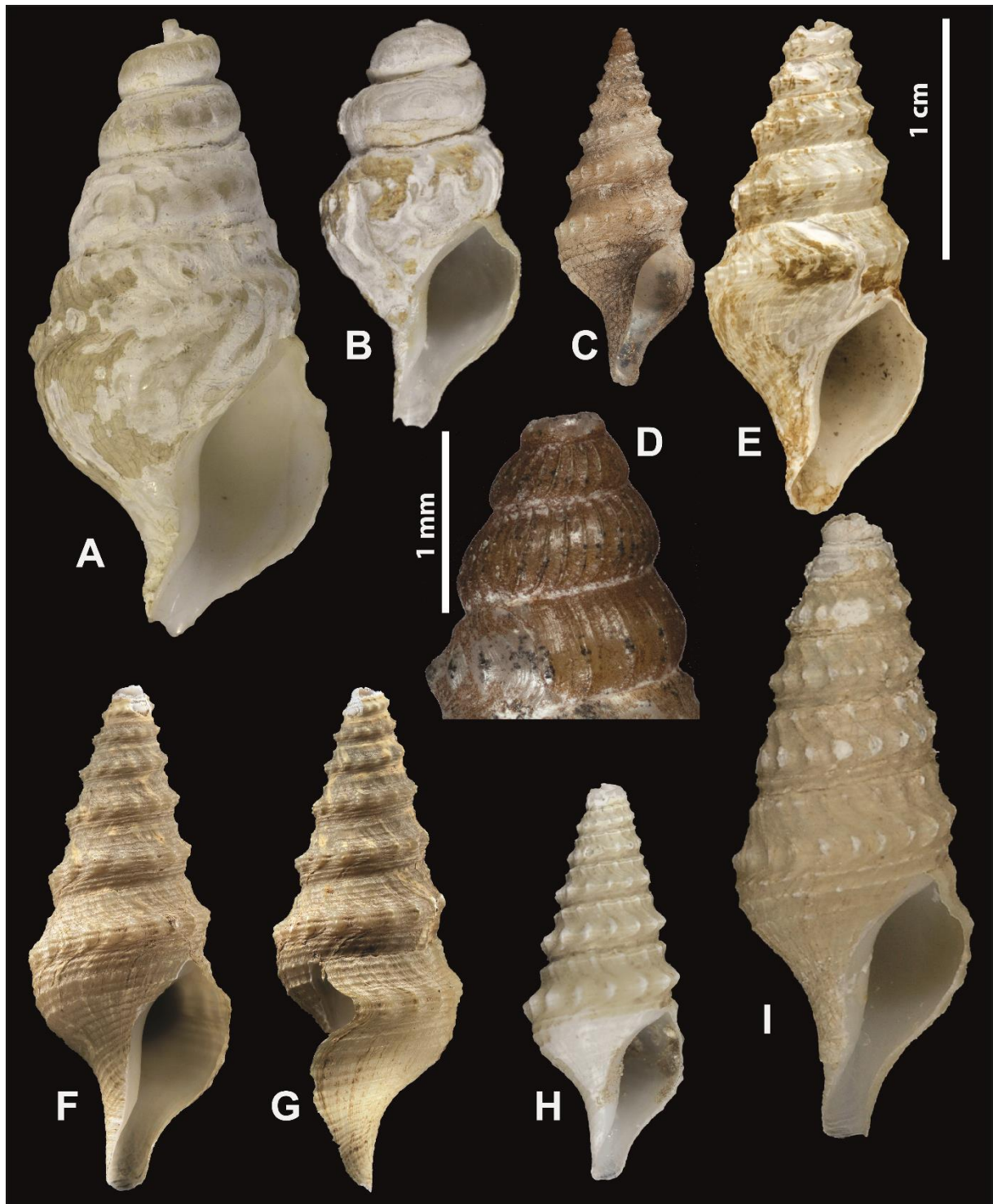


Figure 6: *Cryptogemma phymatias* (Watson, 1886). (A) Lectotype of *Gemmula benthima* Dall, 1908, USNM 123089, Gulf of Panama. (B) Paralectotype of *Gemmula benthima* Dall, 1908, USNM 123087, Gulf of Panama. (C) Holotype of *Bathybermudia carynae* Haas, 1949, FMNH 31656, Bermuda. (D) Lateral view of the protoconch of *B. carynae* (E) Holotype of *Pleurotoma phymatias* Watson, 1886, NHMUK 1887.2.9.957, Philippines. (F-G) AMS T0077, IN2017\_V03 - Sampling the Abyss, Jervis Commonwealth Marine Reserve. (H) MNHN-IM-2013-49928, R/V Tansei-maru, KT-12-32, Okinawa. (I) MNHN-IM-2009-13476, AURORA 2007, Philippines.

Radula medium long, about 2.6 mm in length (0.44 of AL), composed of 74 transverse rows of teeth. Marginal teeth 118–128  $\mu\text{m}$  long (mean 120  $\mu\text{m}$ ,  $n = 5$ ), duplex. Anterior (inner) 1/3 of tooth length solid, very narrow in dorsal view, pointed, in posterior 2/3 major and accessory limbs broadly bifurcating, accessory limb with clear constriction at about half tooth length, slightly shorter than major limb. Central formation with distinct narrow carinated cusp and lateral inconspicuous flaps with indistinct lateral and anterior margins (Fig. 7B–C).

Although this species can be distinguished from its congeners by the frequent loss of the first whorls, and from its sister species *Cryptogemma praesignis* by the slightly larger diameter of the shell, the bathymetry is its most distinct characteristic. *C. praesignis* is found between ~300 and ~1400 m depth, while *Cryptogemma phymatias* is the only Turridae to our knowledge to be exclusively found below ~1400 m deep. Moreover, it is the only Turridae so far, and possibly the first reported benthic gastropod, to have a Pacific–Atlantic distribution (Fig. 4; see Discussion). Most studied type specimens have very eroded shells, but the variability of the last whorls, siphonal twist and aperture among them is in agreement with the variability in the sequenced specimens. Noteworthy, the etymology of this species combines “*Cryptogemma*” (hidden gems) and “*phymatias*” (one who has tubercles), resulting in a confusing combination.

*List of COI diagnostic sites (position: character state):* [379: T, 493: C, 622: C]. Counting starts right next to the 3' end of the LCO1490 primer (from 1 to 658).

*Distribution:* From the Central Indo-Pacific (Banda Sea) to the North Atlantic Ocean (Bermuda Is.) (Fig. 4A), from ~1400 to ~3000 meters depth (Fig. 5). This species is expected to be found in the Indian Ocean, as is the case with other Indo-Pacific species of *Cryptogemma*, but it has not been documented there, due presumably to a lack of sampling at bathyal depths.

### ***Cryptogemma praesignis* (Smith, 1895)**

(Fig. 8A–K)

*Pleurotoma praesignis* Smith, 1895: 4, pl. 1, fig. 4. 1234 m, off Colombo, Ceylon.

? *Pleurotoma microscelida* Dall, 1895: 677. 642 m, South of Oahu Island Albatross st. 3475



*Pleurotoma (Subulata) bisinuata* Martens, 1901: 17. 1134 m, off East Africa, 1°49'N, 45°29'E.

*Pleurotoma rotatilis* Martens, 1902: 239. 1134 m, off Mogadishu, Somalia, East Africa, 1°49'N, 45°29'E. von Martens, 1904: 78, pl. 1, fig. 3.

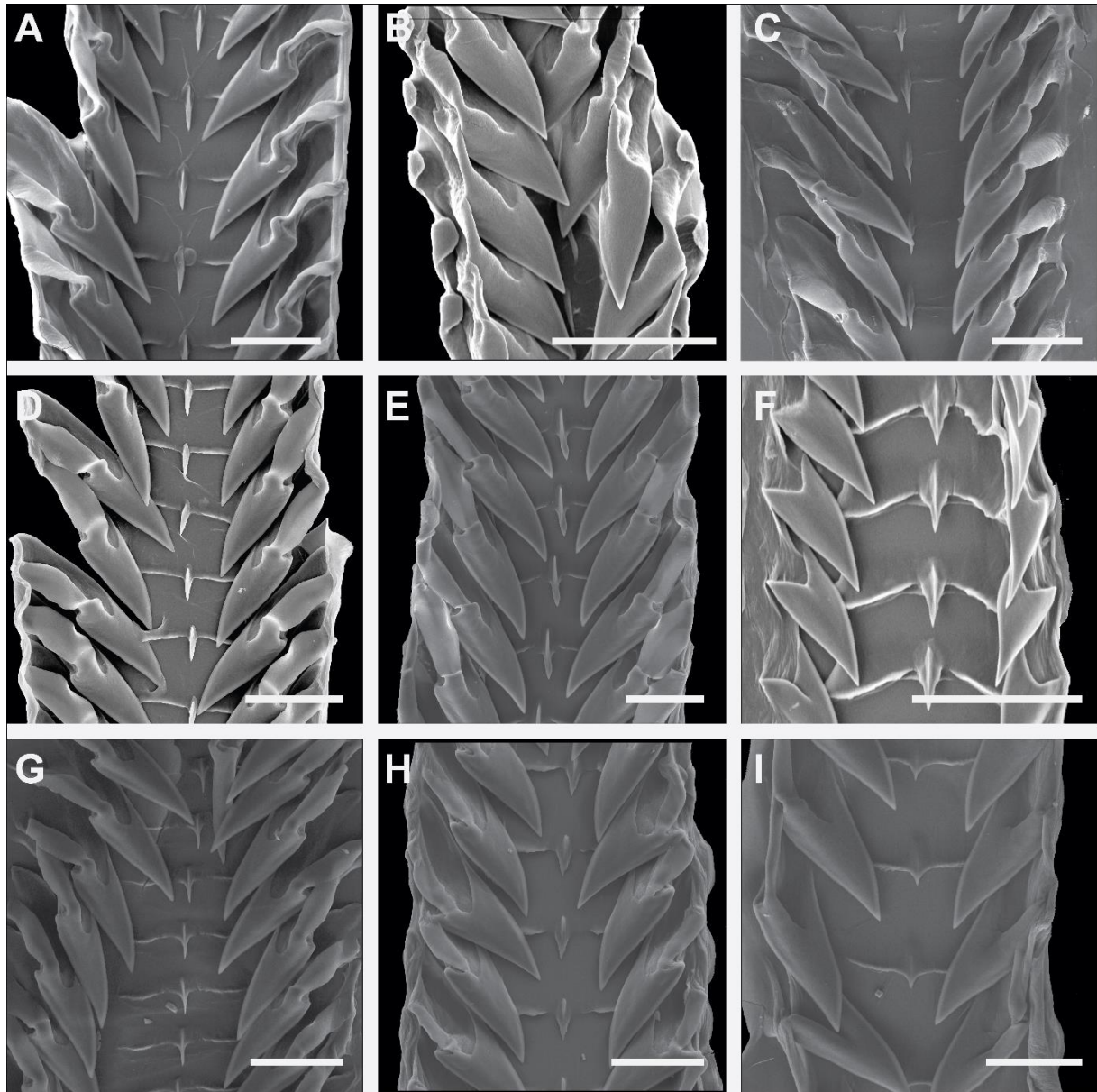


Figure 7: Radulae of studied *Cryptogemma*. (A) *Cryptogemma praesignis* (Smith, 1895). (B-C) *Cryptogemma phymatias* (Watson, 1886). (B) MNHN uncatalogued, Biocal, New-Caledonia, st. CP23 (C) USNM 857019, Venezuela. (D) *Cryptogemma periscelida* (Dall, 1889), R/V Pelican, 28°23.935'N, 89°22.508'W, 675-765 m. (E) *Cryptogemma aethiopica* (Thiele, 1925), MNHN-IM-2013-50175, DONGSHA 2014, Taiwan. (F) *Cryptogemma tessellata* (Powell, 1964), MNHN-IM-2007-40775, EBISCO, Chesterfield Islands. (G) *Cryptogemma unilineata* (Powell, 1964), MNHN-IM-2013-61844, ZhongSha 2015, Taiwan. (H) *Cryptogemma powelli* sp. nov., MNHN-IM-2013-68787, KANACONO, New Caledonia. (I) *Cryptogemma timorensis* (Tesch, 1915), MNHN-IM-2013-09864, PAPUA NIUGINI, Papua. Scale bars 50  $\mu$ m.

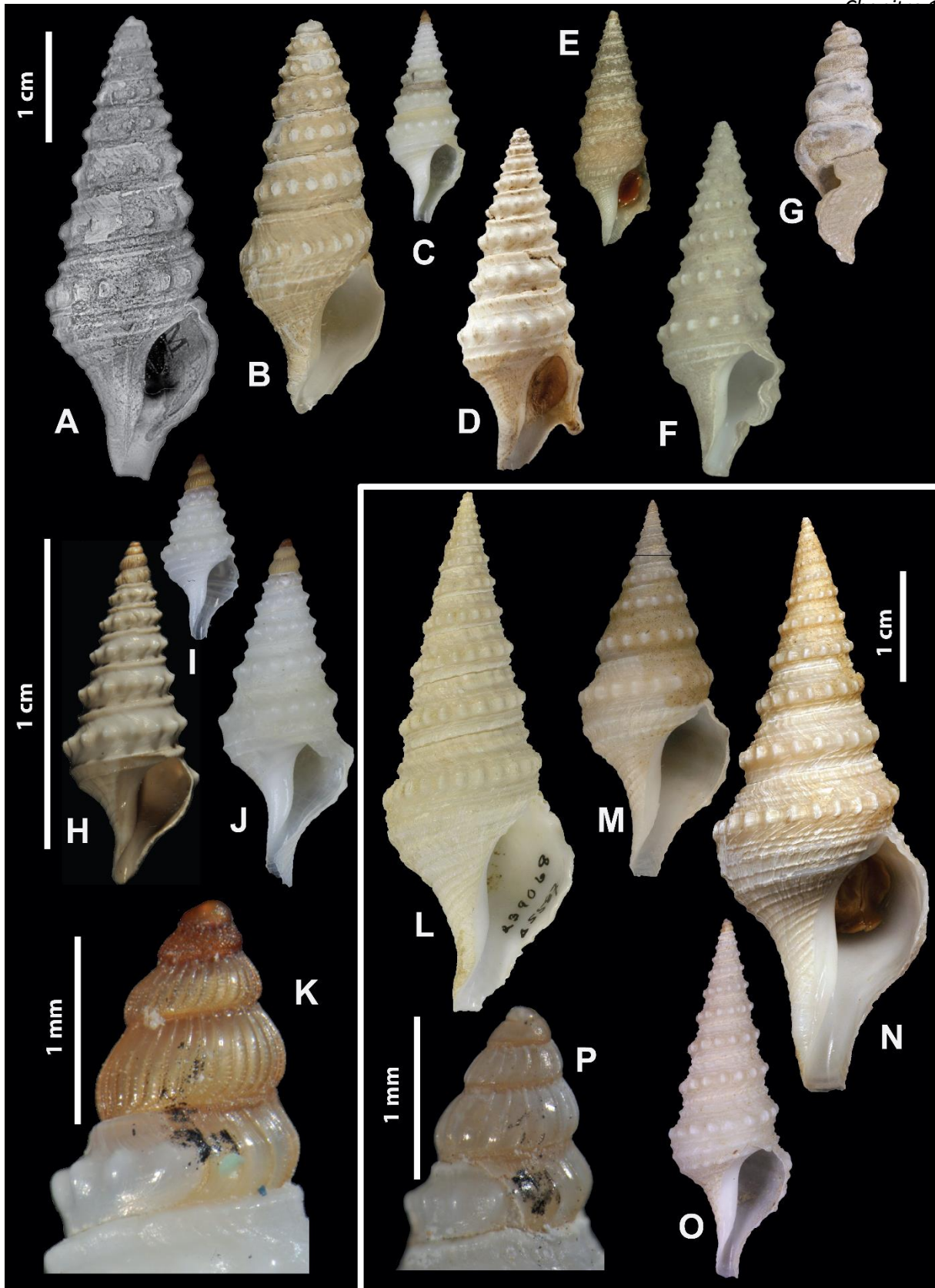


Figure 8: (A-K) *Cryptogemma praesignis* (Smith, 1895) and (L-P) *Cryptogemma timorensis* (Tesch, 1915). (A) Holotype of *Pleurotoma praesignis* Smith, 1895, ZSIC, Sri Lanka. (B) MNHN-IM-2009-13477, AURORA-2007, Philippines. (C) MNHN-IM-2013-62903, BIOMAGLO, Iles Glorieuses. (D) Holotype of *Pleurotoma lobata* Sowerby III, 1903, NHMUK 1903.7.27.49, South Africa. (E) Holotype of *Ptychosyrinx lordhoweensis*, ZMMU Lc-14532, Lord Howe Rise. (F) Syntype of *Pleurotoma bisinuata* von Martens, 1901, ZMB 683, off East Africa. (G) Holotype of *Pleurotoma microscelida* Dall, 1895, USNM 127122, Hawaiian Islands. (H). Holotype of *Pleurotoma rotatilis* von Martens, 1902, ZMB 60070, off East Africa. (I) MNHN-IM-2013-62893, BIOMAGLO, Iles Glorieuses. (J) MNHN-IM-2013-62900, BIOMAGLO, Iles Glorieuses. (K) Lateral view of the protoconch of MNHN-IM-2013-62893. (L) Holotype of *Ptychosyrinx timorensis teschi* Powell, 1964, USNM 239068, Borneo. (M) MNHN-IM-2009-14860, ATIMO VATAE, South of Madagascar. (N) MNHN-IM-2007-40928, SALOMON 2, Salomon Islands. (O) MNHN-IM-2013-58793, KAVIENG, Papua. (P) Lateral view of the protoconch of MNHN-IM-2009-14935.



*Pleurotoma (Surcula) lobata* G. B. Sowerby III, 1903: 213, pl. 4, fig. 9. 805 m, off Cape Natal, Durban and 567 m, off Buffalo River, East London, S. Africa. E. A. Smith, 1906a.

*Ptychosyrinx bisinuata japonica* Okutani, 1964: 425, pl. 4, fig. 9. 620 m, Sea of Enshu-nada, Japan, 34°25.7'N, 137°58.5'E.

*Ptychosyrinx lordhoweensis* Kantor & Sysoev, 1991:205, figs. 1, 2. 1210 m, Lord Howe Rise, off eastern Australia, 30°24'S, 161°51'E.

*Remarks:* This species corresponds to PSH 11 in Puillandre *et al.* (2012).

. Anterior (inner) 1/3 of tooth length solid, narrow in dorsal view, pointed, in posterior 2/3 major and accessory limbs broadly bifurcating, accessory limb with clear constriction and bent at about half tooth length, shorter than major limb. Central formation with distinct narrow carinated cusp and lateral inconspicuous flaps with indistinct lateral and anterior margins (Fig. 7A).

Radula long, about 3.4 mm in length (0.44 of AL), formed of 72 transverse rows of teeth. Marginal teeth 142–151 µm long (mean 149 µm, n = 5, or 1.9% of AL), duplex. Anterior (inner) 1/3 of tooth length solid, narrow in dorsal view, pointed, in posterior 2/3 major and accessory limbs broadly bifurcating, accessory limb with clear constriction and bent at about half tooth length, shorter than major limb. Central formation with distinct narrow carinated cusp and lateral inconspicuous flaps with indistinct lateral and anterior margins (Fig. 7A).

This species is morphologically close to *Cryptogemma timorensis*, although it differs from that species in its generally smaller size (~15–35mm SL), in having a narrower last whorl, and the presence of a tertiary apertural notch in mature females. *Cryptogemma praesignis* also differs from *C. timorensis* in having a much longer radula (0.44 vs 0.2 of AL) and relatively longer marginal teeth (1.9% of AL vs 1.35%). Anatomical examination of Australian material (C571714 and C571704) shows that mature males (SL 29.4 and 34.9 mm respectively) can obtain an extremely large and muscular penis, with a width almost equal to the width of the animal itself, and with a very large, long lateral appendage situated distally. The penial tip is rather blunt, with the opening situated distally. A third, smaller Australian specimen (C571757; SL 25.7 mm), did not possess a well-developed papilla, suggesting that such a feature may develop with increasing maturity and therefore not be present in subadults. Furthermore, the shells of C571714 and C571704 did not possess a tertiary apertural notch. The observation of C571714 and C571704 supplements the findings of Kantor & Sysoev (1991) that the mature females of this species develop a tertiary apertural notch. The authors hypothesized that this structure was possibly connected with the process of fertilization. Such

a feature has been observed in *C. aethiopica* specimens as well (Fig. 9G, I), but not in other *Cryptogemma* species.

The general morphology, and more specifically the presence of a tertiary apertural notch in *Pleurotoma bisinuata* and *Ptychosrynx bisinuata japonica*, supports the synonymization of these names with *C. praesignis*. Because the holotype of *P. bisinuata japonica* is a comparatively large specimen (39 mm), the result of the LDA and the leave-one-out cross-validation procedure attributed it to *C. timorensis*. Possibly for this same reason, the holotype of *P. praesignis* (42 mm) also fell into the *C. timorensis* range. The result of the LDA analysis and the leave-one-out cross-validation procedure when considering only shape predicted the two holotypes as *C. praesignis*, (see Figure S3) thus justifying the attribution of *P. praesignis* to this molecular species. The sequencing of a paratype of *P. lordhoweensis* confirmed its conspecificity with *C. praesignis*. The name *Pleurotoma lobata*, although conchologically very similar to *P. bisinuata*, has never been synonymized based on the reportedly different morphology of the radula: the absence of a central tooth (= central formation as understood herein) in the specimens of *P. lobata* examined by Barnard (1958), contrasting with the large unicuspid rectangular-based central tooth described in *P. bisinuata* by Thiele (1929, p. 359). Powell (1964) suggested that Thiele could have mixed his radular preparations, but in light of the radulae type of *Cryptogemma* (Fig. 8), all composed of a unicuspid central formation, it is more likely that either it was Barnard, who mixed up the radula preparation, or that the specimens examined by Barnard are not *Cryptogemma* specimens. The holotype of *Pleurotoma rotatilis* is almost identical conchologically to a sequenced juvenile from the East African coast (MNHN-IM-2013-62900; Fig. 8J). Some doubts remain as to the status of *Pleurotoma microscelida*, but the size and overall last whorl morphology of the type specimen show a stronger resemblance to *C. praesignis* than to any other *Cryptogemma* species.

*List of COI diagnostic sites (position: character state):* [232: A, 319: T, 407: C, 409: T]

*Distribution:* This species is known to occur from the West Indian Ocean to the East Pacific Ocean (Fig. 4B), from ~300 to ~1400 m of depth (Fig. 5).

***Cryptogemma timorensis* (Tesch, 1915)**

(Fig. 8L-P)

†*Pleurotoma timorensis* Tesch, 1915:27, pl. 77(5), figs. 52, 53. Timor. Pliocene.

†*Pleurotoma ktolemandoënsis* K. Martin, 1933: 20, pl. 2, figs. 14, 14a. Ktolemando. Miocene.

*Ptychosyrinx timorensis teschi* Powell, 1964: 291 (22–853), pl. 223, figs. 5, 6. 759 m, NW off Sipadan Island, Borneo.

*Type locality:* Pliocene of Timor Island.

*Remarks:* This species corresponds to PSH 14 of Puillandre *et al.* (2012).

Protoconch commonly eroded, with PD = 1–1.2 mm and PL = 1.3–1.675 mm and number of whorls varying from 4 to 4.5.

Radula medium short, about 2.5 mm in length (0.2 of AL), formed of 45 transverse rows of teeth. Marginal teeth 162–172  $\mu\text{m}$  long (mean 168  $\mu\text{m}$ ,  $n = 5$ , or 1.35% of AL), duplex. Anterior (inner) 1/3 of tooth length solid, medium broad in dorsal view, lanceolate, in posterior 2/3 major and accessory limbs broadly bifurcating, accessory limb with clear constriction and bent at about half tooth length, nearly the same length as major limb. Central formation with short and obtuse cusp and lateral inconspicuous flaps with indistinct lateral and anterior margins. Posterior margin thickened and uninterrupted along its length (Fig. 7I).

This type of this species, *Pleurotoma timorensis*, is a fossil specimen from Timor island (Indonesia) described by Tesch, and Martin described another very similar specimen from Buton island (Indonesia), *Pleurotoma ktolemandoënsis*, without reference to the work of Tesch. However, Martin furtherly recognized the affinities between *P. timorensis* and *P. ktolemandoënsis* (Martin, 1935). Finally, Robba *et al.* (1989) synonymized *P. ktolemandoënsis* with *P. timorensis*. Powell (1964) described the subspecies *Ptychosyrinx timorensis teschi* based on an extant specimen, which Sysoev (1996) elevated to species rank. Both authors' arguments for separating the fossil species from the extant was the 'much broader' shell of the extant *Ptychosyrinx timorensis teschi* specimen. Examination of the morphological variability of the sequenced specimens showed that dimensions and shape of fossil specimens is close to the inferred average breadth of the species; some specimens show broader aperture and a taller last whorl than others, despite having the same number of teleoconch whorls. The results of the LDA showed that *P. timorensis* and *P. timorensis teschi* fall into the variability of the sequenced specimens. The broader last whorl and the

generally larger size (~30-55mm) are characteristic for *C. timorensis* when compared to *C. praesignis*.

*List of COI diagnostic sites (position: character state):* [73: G, 214: G, 334: G, 511: G]

*Distribution:* Found in the West Indian Ocean to the Central Indo-Pacific (Fig. 4C), from ~300 to ~1200 meters depth (Fig. 5). Curiously, this species has not been found in New-Caledonia, despite considerable sampling effort in this region.

### ***Cryptogemma aethiopica* (Thiele, 1925)**

(Fig. 9)

*Pleurotoma aethiopica* Thiele, 1925:174(208), pl. 22(34), fig. 25. 638 m, off Somalia, East Africa, 0°27'S, 42°47.3'E. (Expedition Deutschen Tiefsee, st. 253).

*Pleurotoma fusiformis* Thiele, 1925:176(210), pl. 22(34), fig. 24. 614 m, Nias-südkanal, 0°15.2'N, 98°08.8'E.

*Gemmula thielei* Finlay, 1930:47. (nom. nov. for *Pleurotoma fusiformis* Thiele, 1925).

†*Pleurotoma trincincta* Martin, 1935:113, pl. 2, figs. 2, 2a. Buton Island, SE Celebes. Oligocene.

†*Pinguiggemula okinavensis* McNeil, 1960:104, pl. 9, figs. 12–14. Okinawa. Shinzato Tuff Member, Miocene or Pliocene.

*Pinguiggemula luzonica* Powell, 1964 1964:278(22–790), pl. 215, figs. 3, 4. 326m, off Hermana, Menor Island, Luzon Island, Philippines.

*Pinguiggemula philippinensis* Powell, 1964:278(22–790), pl. 215, figs. 5, 6. 512m, off Santiago, west Luzon Island, Philippines.

*Remarks:* This species was not included in the study by Puillandre *et al.* (2012). Protoconch commonly eroded, with PD = 1.05-1.25, PL = 1.375-1.75 and number of whorls ranging from 4.2 to 5.

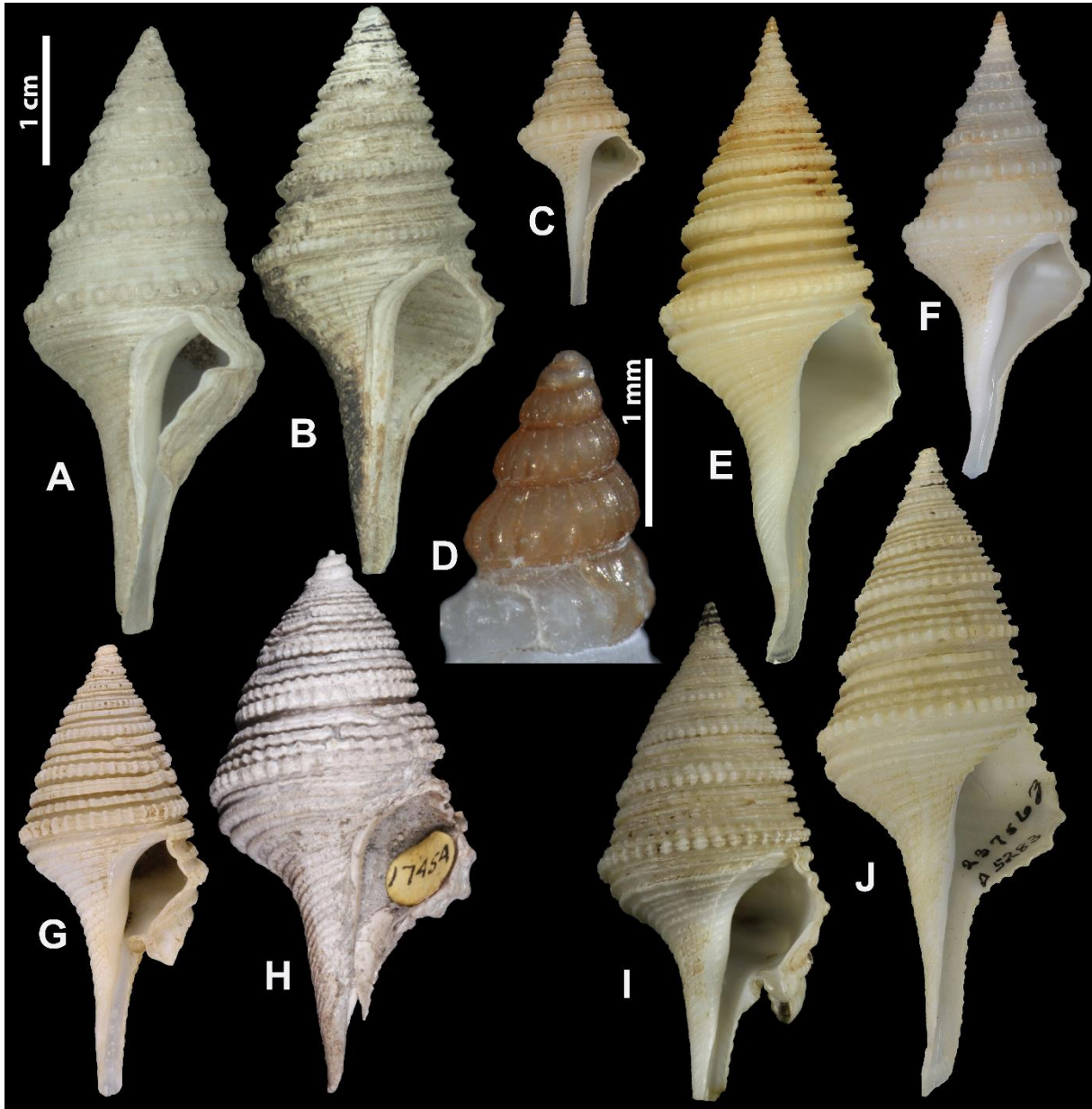


Figure 9: *Cryptogemma aethiopica* (Thiele, 1925). (A) Holotype of *Pleurotoma aethiopica* Thiele 1925, ZMB 109376, off East Africa. (B) Holotype of *Pleurotoma fusiformis* Thiele, 1925, ZMB 109388, off Sumatra. (C) MNH- IM-2009-14979, ATIMO VATAE, South of Madagascar. (D) Lateral view of the protoconch of MNHN-IM-2013-48115, KANADEEP, Chesterfield Islands. (E) MNHN-IM- 2009-29301, EXBODI, New Caledonia. (F) MNHN-IM- 2009-14778, ATIMO VATAE, South of Madagascar. (G) Holotype of *Pinguiggemula luzonica* Powell, 1964, USNM 237784, Philippines. (H) Holotype of *Pinguiggemula okinawensis* McNeil, 1960, UNSM MO 562851, Miocene or Pliocene of Okinawa. (I) MNHN-IM-2013-19950, PAPUA NIUGINI, Papua. (J) Holotype of *Pinguiggemula philippinensis* Powell, 1964, USNM 237563, Philippines.

Radula long, about 3.4 mm in length (0.37 of AL), composed of 77 transverse rows of teeth. Marginal teeth 150-159  $\mu\text{m}$  long (mean 155  $\mu\text{m}$ ,  $n = 5$ , or 1.66% of AL), duplex. Anterior (inner) 0.4 of tooth length solid, narrow in dorsal view, awl-shaped, in posterior part major and accessory limbs broadly bifurcating, accessory limb with clear constriction and bent at about half tooth length, thin, nearly same length as major limb. Central formation with

long, very narrow, sharp, carinated cusp and lateral flaps with distinct posterior and antero-lateral margins. Flaps not completely fused with cusp (Fig. 7E).

The former *Pinguiggemmula* genus is easily distinguishable from *Gemmula* in having a broadly conic spire, a strongly constricted base and a long, straight siphonal canal (Powell, 1964). Although expressing a wide variety of forms, the molecular analysis resulted in a single species hypothesis. Several species were described based on the sculptural details such as the number of gemmate cords (1, 2 or none). The sculpture seems to correlate with geography, with the forms from West Indian Ocean having smooth inter-suture sculpture, while the forms from East Indian-West Pacific usually possess two or three gemmate cords. Some large specimens of this species possess a similar 'tertiary notch' to that of *C. praesignis*, suggesting sexual dimorphism also for this species (Kantor & Sysoev, 1991).

*List of COI diagnostic sites (position: character state):* [115: C, 307: A, 418: G]

*Distribution:* Found in East Africa to Central Indo-Pacific (Fig. 4E), from ~400 to ~850 meters of depth (Fig. 5).

### ***Cryptogemma tessellata* (Powell, 1967)**

(Fig. 10A-D)

*Gemmula tessellata* Powell, 1967:439(22–734a), pl. 315. 183–219m, off Waikiki, Oahu Island, Hawaiian Islands (Collected by Dr. Pat Burgess)

*Remarks:* This species corresponds to PSH 12 in Puillandre *et al.* (2012). The protoconch of the holotype is 4.5 whorls according to Powell (1967) and PD = 1.22 mm and PL = 1.64 mm according to measurements inferred from photographs of the holotype. Protoconch with high variability, with PD = 1.175–1.3, PL = 1.525–1.975 and consisting of 4 to 5.5 whorls.

Radula medium long, about 1.2 mm in length (0.29 of AL), composed of 51 transverse rows of teeth. Marginal teeth 86–91 µm long (mean 89 µm, n = 2, or 2.1% of AL), duplex. Anterior (inner) 1/2 of tooth length solid, medium broad in dorsal view, triangular, accessory limb weak, thin, without constriction, much shorter than major limb, but nearly of same width (the marginal teeth on Fig. 8F are not fully sclerotized). Central formation with long, sharp and carinated central cusp, slightly curved in profile, and with lateral conspicuous flaps with



distinct margins. Flaps not completely fused with cusp. Anterior margin of central formation strongly concave (Fig. 7F).

The ‘light form’ of *C. tessellata* has been misidentified in Puillandre *et al.* (2012) as *Xenuroturrus gemmuloides* Powell, 1967 as it shares superficially similar features such as the white-yellowish shell punctuated with regular brown-orange spots, and the small size (~15-25mm). The ‘brown-orange form’ of *C. tessellata*, closer to the holotype, is quite different from the ‘light form’ not only in coloration, but also in having a stouter outline and a more tuberculate subsutural fold (Powell, 1964). Although readily distinguished from the other *Cryptogemma* species due to its color pattern and small size, LDA analysis indicates that this species more closely resembles small adults of *C. praesignis*.

*List of COI diagnostic sites (position: character state):* [46: C, 61: A, 316: T]

*Distribution:* The sequenced specimens were found in New Caledonia only and the holotype is from the Hawaiian Islands (Fig. 4C), from ~200 to ~500 meters depth (Fig. 5). The protoconch characteristics, similar in shape, size and number of whorls to *C. praesignis*, imply that the species may have a much broader range than what is currently documented, possibly covering the entire Central Indo-Pacific.

### ***Cryptogemma unilineata* (Powell, 1967)**

(Fig. 10E-H)

*Gemmula congener unilineata* Powell, 1967:437(22–716a), pl. 313. 366 m, off Waikiki, Oahu Island, Hawaii (Expedition Pele, June 13 1964).

*Remarks:* This species corresponds to PSH 13 in Puillandre *et al.* (2012).

Protoconch commonly retained, with PD = 1.075-1.275, PL = 1.375-1.925 and number of whorls ranging from 4.75 to 5.5.

Radula long, about 3 mm in length (0.31 of AL), composed of 99 transverse rows of teeth. Marginal teeth 121-127  $\mu\text{m}$  long (mean 125  $\mu\text{m}$ , n = 5, or 1.27% of AL), duplex. Anterior (inner) 0.4 of tooth length solid, very narrow in dorsal view, awl-shaped, in posterior part major and accessory limbs broadly bifurcating, accessory limb with clear constriction and bent at about half tooth length, slightly shorter than major limb. Central formation with long

sharp carinated cusp and lateral flaps with distinct posterior and antero-lateral margins. Flaps not completely fused with cusp (Fig. 7G).

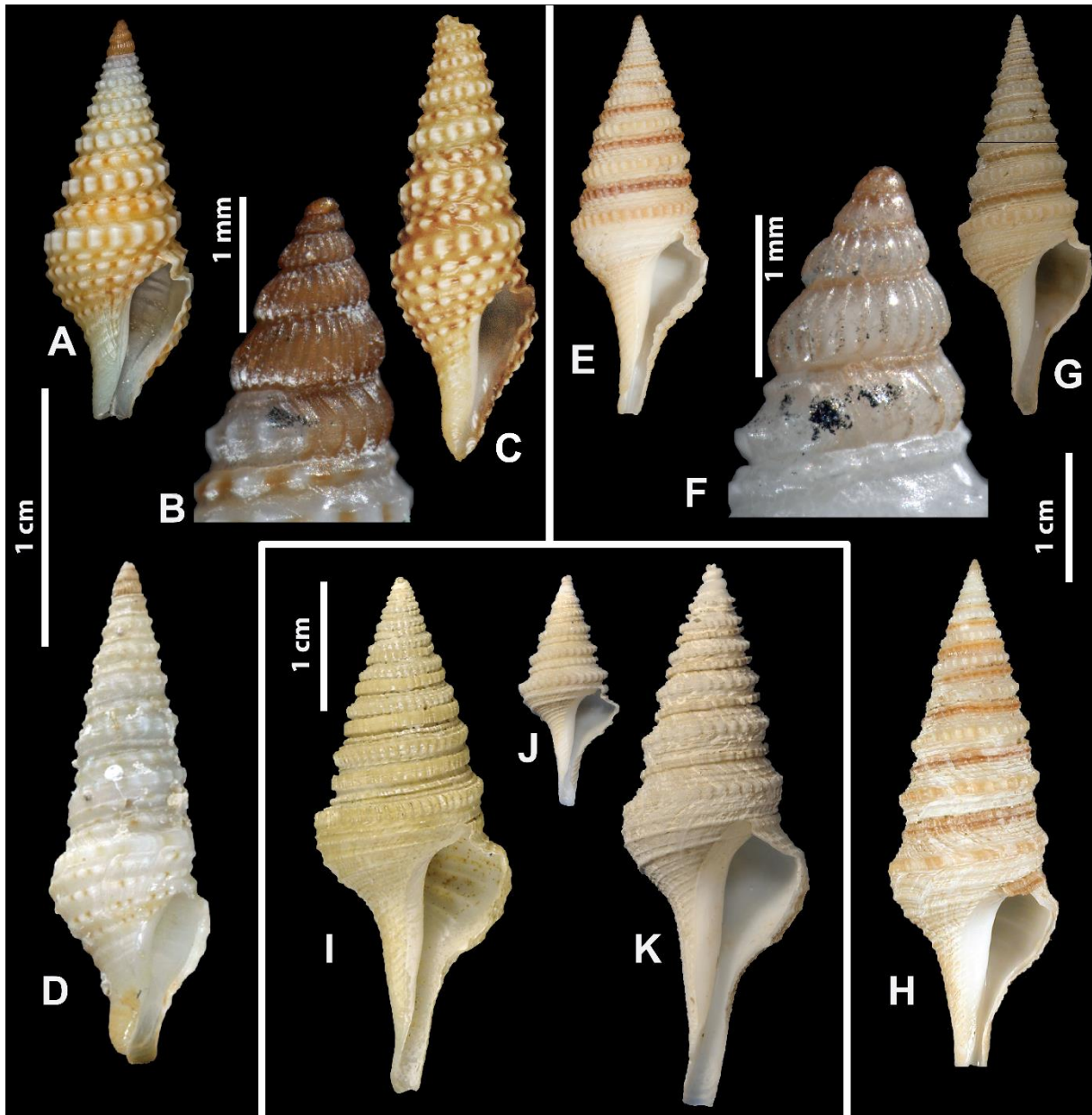


Figure 10: (A-D) *Cryptogemma tessellata* (Powell, 1964), (E-H) *Cryptogemma unilineata* (Powell, 1964) & (I-K) *Cryptogemma periscelida* (Dall, 1889). (A) Holotype of *Gemmula tessellata* Powell, 1964, AWMM MA71042, Hawaiian Islands. (B) Lateral view of the protoconch of MNHN-IM-2013-48252, KANADEEP, Chesterfield Islands. (C) MNHN-IM-2009-24975, EXBODI, New Caledonia. (D) MNHN-IM- 2007-40775, EBSICO, Chesterfield Islands. (E) Holotype of *Gemmula congener unilineata* Powell, 1964, BPBM 8929, Hawaiian Islands. (F) Lateral view of the protoconch of MNHN-IM-2007-40786, NORFOLK 2, New Caledonia. (G) MNHN-IM-2009-14953, ATIMO VATAE, South of Madagascar. (H) MNHN-IM-2007-38340, MAINBAZA, Mozambic channel. (I) Lectotype of *Pleurotoma periscelida* Dall, 1889, USNM 87391, Columbia. (J) MNHN-IM- 2013-56285, GUYANE 2014, French Guyana. (K) MNHN-IM- 2013-56283, GUYANE 2014, French Guyana.

This species has been named based on the characteristic carinated brown-orange subsutural band. The brown-orange subsutural band has been found in *C. powelli* n. sp. and in some specimens of *C. praesignis*, as well as in other Turridae such as *Gemmula cosmoi* (Sykes, 1930), but it is generally smoother and thinner to that of *C. unilineata*. The LDA analysis indicated that *C. unilineata* has generally a more angulated shape on the outer aperture lip of the last whorl, indicating a more concave subsutural ramp in comparison to other *Cryptogemma*. However, we note that the convex hulls (Fig. 1) of *C. unilineata* and *C. powelli* sp. nov. are greatly overlapping, so the strong carinated brown-orange subsutural band remains the best suited character for distinguishing *C. unilineata* from *C. powelli*.

*List of COI diagnostic sites (position: character state):* [172: G, 307: T, 361: G, 370: C]

*Distribution:* From East Africa to Central Pacific (Hawaiian and Society islands) (Fig. 7D), from ~300 to ~800 meters of depth (Fig. 5).

### ***Cryptogemma periscelida* (Dall, 1889)**

(Fig. 10I-K)

*Pleurotoma periscelida* Dall, 1889:74, pl. 32, fig. 2. 283m., (USS Albatross expedition, st. 2143, collected on March 23 1884) and 196m, off Hatteras, North Carolina. Dall, 1889b.

#### *Remarks:*

The protoconch is unknown, as it is consistently eroded in all examined material, even in younger specimens.

Radula medium long (part of youngest section lost), more than 1.9 mm in length ( $>0.22$  of AL), composed of more than 50 transverse rows of teeth. Marginal teeth 107-116  $\mu\text{m}$  long (mean 110  $\mu\text{m}$ ,  $n = 5$ , or 1.25% of AL), duplex. Anterior (inner) 0.4 of tooth length solid, narrow lanceolate in dorsal view, in posterior part major and accessory limbs medium broadly bifurcating, accessory limb with clear constriction and bent at less than half tooth length, thin, nearly same length and width as major limb. Central formation with long, narrow, sharp carinated cusp and lateral flaps with distinct posterior and postero-lateral margins. Flaps completely fused with cusp (Fig. 7D).

This species has been described from the Gulf of Mexico, although specimens from Suriname and from the MNHN recent expedition GUYANE 2014 (French Guyana) were also collected. This is the only ‘*Gemmula*’-like species to be found exclusively in the Atlantic Ocean, and it occurs in a 200-800 m depth range.

*List of COI diagnostic sites (position: character state):* [352: C, 412: G, 655: C]

*Distribution:* Found in the Gulf of Mexico and further south to French Guyana (Fig. 4B), from ~200 to ~500m of depth (Fig. 5).

***Cryptogemma powelli* sp. nov.**

(Fig. 11)

*Type material:* Holotype MNHN-IM-2013-68787; Paratype 1, MNHN IM-2007-40795, paratype 2, MNHN-IM- 2007-40765, all live collected and processed for DNA extraction.

*Type locality:* New-Caledonia, South-West of Ile des Pins, 22°48'S, 167°15'E, 449-465 meters depth, sand and debris (Expedition KANACONO, st. DW4697).

*Etymology:* Named after the New Zealand malacologist A.W.B. Powell (1901-1987), who contributed enormously to the systematics of Conoidea, and in particular, who revised the family Turridae in 1964 and described several species of *Cryptogemma*.

*Description (holotype):* Shell narrowly fusiform, with high spire, and medium-long siphonal canal. Protoconch conical, eroded, with a diameter of 1.125 mm, of about 4 convex whorls. Teleoconch of 9 whorls; suture shallow, impressed. Shell height 38 mm, aperture height 10.9 mm and shell diameter 12.5 mm. Whorls strongly shouldered, with slightly concave shoulder slope, very weakly convex, nearly cylindrical periphery. Spiral sculpture of fine, broadly spaced, wavy subequal spiral cords on subsutural ramp, 10 on last whorl, uppermost being much more pronounced than others, colored light orange-brown. Sinus cord strongly gemmate, with gemmules clearly bisected (39 on the body whorl). Spiral cords becoming thicker on whorl periphery, sometimes nodulose, some notably wider than others. Axial sculpture of fine growth lines. Last whorl shortly constricted to long siphonal canal, with 32

cords below sinus, of which about 20 on canal. Aperture irregular oval, outer lip thin, simple. Anal sinus moderately deep, U-shaped. Shell color straw-yellow, subsutural cord orange-brown, gemmae slightly lighter than background.

Radula medium long, around 3.1 mm in length (0.28 of AL), composed of 75 transverse rows of teeth. Marginal teeth 129-135  $\mu\text{m}$  long (mean 133  $\mu\text{m}$ ,  $n = 5$ , or 1.22% of AL), duplex. Anterior (inner) 1/3 of tooth length solid, lanceolate in dorsal view, in posterior part major and accessory limbs broadly bifurcating, accessory limb with clear constriction at half tooth length, thin, about half length and width of major limb. Central formation with medium long and medium broad sharp carinated cusp and lateral flaps with distinct posterior and posterolateral margins. Flaps not completely fused with cusp (Fig. 7H).

*Remarks:* This species corresponds to PSH 15 in Puillandre *et al.* (2012). *Cryptogemma powelli* sp. nov. exhibits similar intraspecific variation in shell shape to that of its congeners, based on its convex hull area (Fig. 1). Studied specimens vary in shades of shell color, from light yellowish to light orange, occasionally with a patchy pattern. In terms of both size and shape, *C. powelli* can in some cases be undistinguishable from *C. unilineata* (Fig. 1). The morphometric analysis shows considerable overlap of the measured shell features between the two species, with the main distinction between them expressed by the second axis. Interpretation of this result suggests that the *C. powelli* on average possess a more concave curvature of the outer aperture lip, compared to *C. unilineata*. Besides, distinction of *C. powelli* and *C. unilineata* mostly relies on the degree of pronunciation of the subsutural cord – it is typically not significantly stronger than the succeeding cords in the former species, but is always thick and gemmate in the latter. The subsutural cord is, however, thicker in younger specimens, rendering the distinction of juvenile *C. powelli* and *C. unilineata* specimens problematic.

*List of COI diagnostic sites (position: character state):* [112: G, 322: T, 433: A]

*Distribution:* The confirmed distribution of this species based on sequenced specimens ranges from the Indian Ocean to the Central Indo-Pacific (Fig. 4F), from ~400 to ~600m of depth.

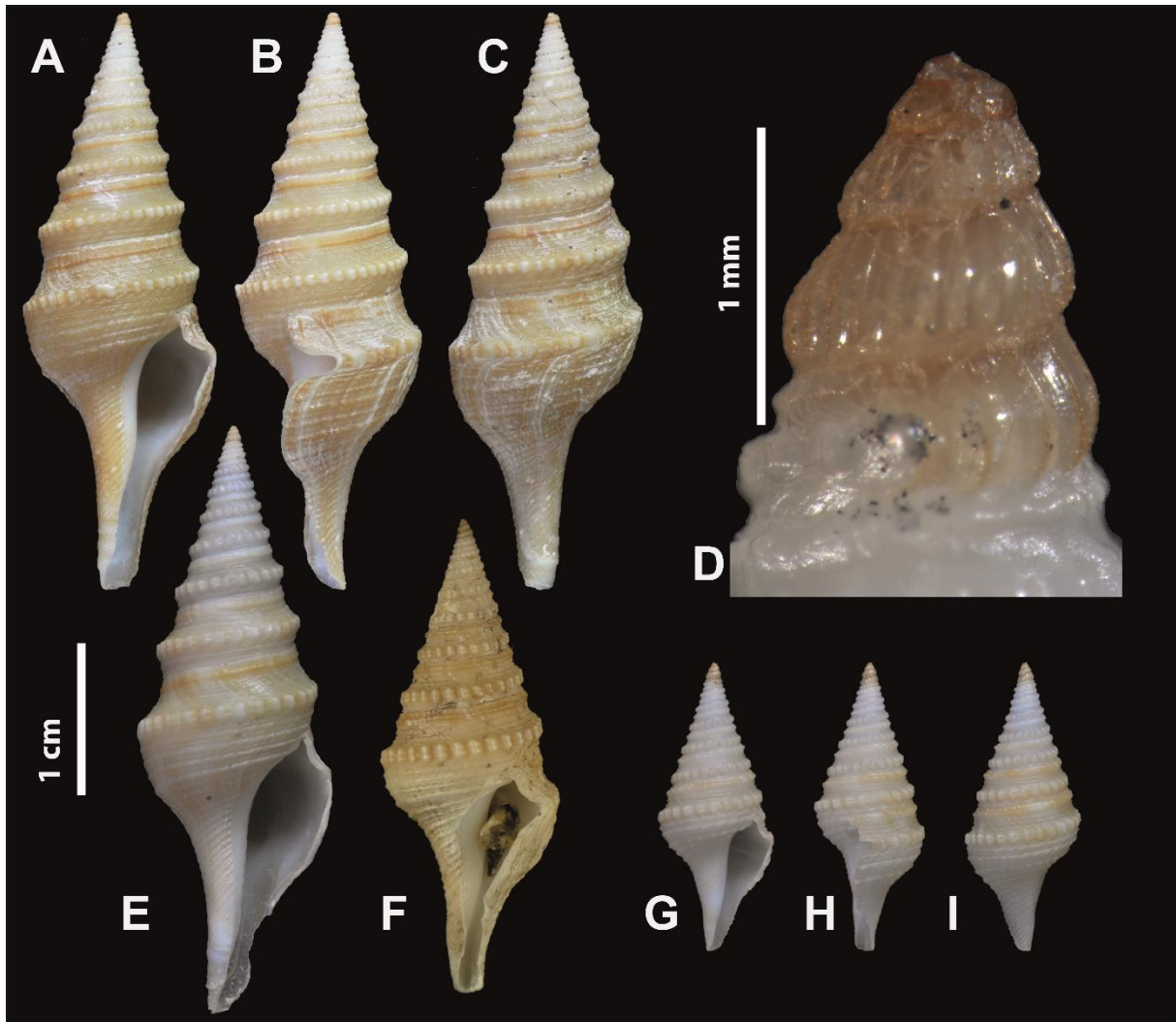


Figure 11: *Cryptogemma powelli* sp. nov. (A-C) Holotype, MNHN-IM- 2013-68787, KANACONO, New Caledonia. (D) Lateral view of the protoconch of the paratype 1 MNHN-IM- 2007-40765, EBISCO, Chesterfield Islands. (E) Paratype 2, MNHN-IM-2007-40795, Norfolk 2, New Caledonia. (F) MNHN-IM-2013-55832, TAIWAN 2013, Taiwan. (G-I) Paratype, MNHN-IM- 2007-40765, EBISCO, Chesterfield Islands.

## Discussion

Using a combination of dna-based species delimitation methods and, more classical, morphological taxonomic methods, we reevaluated the *Cryptogemma* group studied in Puillandre *et al.* (2012) with new sequences – in some cases from type material – resulting in eight delimited species. Although both *COI* and *28S* proved, in this clade, to be appropriate markers for species delimitation, confirming the five *COI*-based species hypothesis (*C. praesignis*, *C. timorensis*, *C. tessellata*, *C. unilineata* and *C. powelli* sp. nov.) in Puillandre *et al.* (2012), *28S* sequences were not produced here. The *28S* sequencing of *C. aethiopica* and *C. periscelida* failed, but these two species are well recognizable morphologically and are not sister to a single species but to a group of species, making it even less probable that the *COI*-



based species hypotheses are incorrect than if there were sister species. The case of *C. praesignis* and *C. phymatias* could be considered problematic as they are found to be sister species and are separated bathymetrically. However, the single sequence of 28S recovered for *C. phymatias* (C571715) showed a divergence with the available 28S sequences of *C. praesignis* comparable to the divergence between other sister species (results not shown herein; see Puillandre *et al.*, 2012; Puillandre *et al.*, 2017).

Among the eight delimited species, *C. phymatias*, *C. praesignis*, *C. timorensis* and *C. aethiopica* comprise 20 of the available species names, with *C. powelli* the only one described as new. Our results contrast with the usual tendency in this family toward the discovery of new species (e.g. Puillandre *et al.*, 2017; Abdelkrim *et al.*, 2018b). All the synonymized species names in this group were established before the ‘molecular era of taxonomy’, reinforcing the recommendation – yet not always followed (Bouchet & Strong, 2010) – of systematically relying on molecular analysis, combined with other data, to delimit species. However, our study also shows that using only molecular delimitation of species while ignoring some key specimens such as types would limit knowledge of intra-specific morphological variability, thus obscuring also the ability to relate such variability to extrinsic parameters, such as geographic and bathymetric ranges. In this study, morphological study and taxonomic expertise on types resulted in the extension of intra-specific variability (Fig. 1; e.g. *C. aethiopica* types). This might be explained because, when describing a new species, there can be a strong subjective effect on the choice of the type specimens (usually the largest of a lot). The recent and promising success in sequencing DNA obtained from empty shells (Villanea, Parent & Kemp, 2016; Der Sarkissian *et al.*, 2017) should assist in attributing names to molecularly-defined species by sequencing type specimens. However, sequencing shells implies destructive sampling, and will therefore possibly not be accepted by many museum curators. Furthermore, identifying and attributing material to species within a species complex still requires considerable taxonomic investigation.

By exploring the intraspecific variability of the shell, in isolation or in relation to interspecific variability, geographical and bathymetric ranges, or ecological data, we here emphasize a few remarks that may provide directions for the future study of this group.

*Geographical intra-specific variability.* in *C. aethiopica*, for which many species were synonymized, there are significant geographical variations in the spire whorls. Specimens from the West Indian Ocean were found to have very smooth spire whorls, while specimens from the central Indo-Pacific were found with either two or three well-marked gemmate cords. Considering the few available discrete shell features, these distinguishable

characteristics were used by authors to justify the creation of several species names. No such intraspecific variation in shell sculpture was found in the other species of *Cryptogemma*, although other geographically-structured morphological variability has been observed in the Turridae genus *Lophiotoma* (Puillandre *et al.* 2017), with shell color varying from one side of the Indo-Pacific to the other.

*Ecological variability.* It is shown (e.g. Bierne, Bonhomme & David, 2003; Hauquier *et al.*, 2017) that some benthic species have clear preferences with regards to the substrate on which they thrive (e.g. sandy/muddy vs. rocky bottom). However, other species are generalists with respect to the substratum, as seems to be the case for the *Cryptogemma* species. All species were found both on hard and soft substrate, with or without organic matter (e.g. leaves or sunken wood). A remarkable case is *C. tessellata*, for which the three available trawl images showed three very different environments (Fig. 12), and, even more remarkably, correlated with shell morphology (Fig. 10C-D); specimens found in sandy or spongy environments exhibited a light pale color ground while the individual found in a volcanic rock bottom has a reddish color ground. However, the holotype of this species shows an intermediate pattern and was collected “in mud and sand” (Powell, 1964). More material is needed to properly study the potential effects of substrate on shell color.

*Sexual dimorphism.* *C. praesignis* is the first Turridae turrid species for which apparent sexual dimorphism of the shell has been documented (Kantor & Sysoev, 1991). The (mature) females possess one or several tertiary notches that is probably linked to fertilization according to the authors, since mature males possess an extremely large and muscular penis. The authors proposed the hypothesis that these notches appear periodically during the breeding period, mainly based on the observation of several successive notches on the shells.

*Bathymetric zonation.* Bathymetric segregation

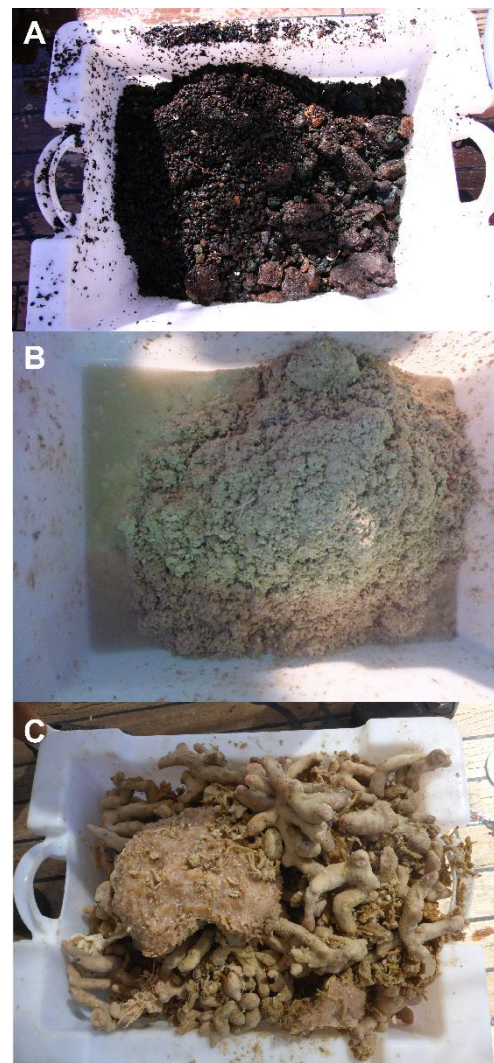


Figure 12: Pictures of trawls from stations where *Cryptogemma tessellata* (Powell, 1964) specimens were found.

between sister species has been shown in molluscan species (e.g. Clague *et al.*, 2011). In the Turridae, it has been found that a pair of sister species, *Lophiotoma abbreviata* and *L. brevicaudata*, show distinct bathymetric preferences, with the former observed at an average depth of 3 m and the latter at an average depth of 15 m (Puillandre *et al.*, 2017). The present study demonstrates that, at a much wider scale, a similar case was found in the *C. praesignis/C. phymatias* species pair, with the former species occurring between ~300 to ~1400 m and the latter between ~1400 to ~3000m depth (Fig. 5). This example could be added to the list of models used for studying speciation processes in benthic species with high dispersal capabilities and large geographic distributions, for which a topographic barrier would not satisfactorily explain genetic isolation (see Etter *et al.*, 2005; Zardus *et al.*, 2006).

*Geographical ranges.* While an Indo-Pacific distribution is common in benthic species, but so far no – molecularly supported - cases of Pacific-Atlantic distribution have been found in benthic gastropods. There are cases of cosmopolitan pelagic gastropods: *Glaucus atlanticus* (Churchill *et al.*, 2014), *Clione limacina* (Jennings *et al.*, 2010) and *Fiona pinnata* (Trickey *et al.*, 2016). To this extent, *C. phymatias* could be the first documented, molecularly supported, case of a benthic gastropod with a Pacific-Atlantic distribution (Fig. 4A). Cases of such molecularly confirmed cosmopolitan benthic species have been found in other phyla such as Annelida (Meyer *et al.*, 2008; Sun *et al.*, 2016; Tomioka *et al.*, 2016), Arthropoda (Casanova *et al.*, 1998; Havermans *et al.*, 2013) and Algae (Macaya & Zuccarello, 2010). However, human-mediated introduction was strongly suggested for some of these cases, e.g. annelids. Nothing is known of the biological features of *C. phymatias* that would enable such a wide dispersal of larvae.

*Protoconch variability.* As reported in Bouchet & Warren (1979; 1994), many deep-water gastropods' larval shells indicate a planktotrophic larval development, which is generally associated with good dispersal abilities and gene flow. The *Cryptogemma* species are no exception to this, with protoconch varying from 4 to 5.25 whorls. However, no correlation was found between the number of whorls of protoconch in a specimen and its associated depth, implying that, from a certain threshold (e.g. 4), the number of whorls is not the only factor determining the ability to disperse. Furthermore, in other neogastropod species complex for which there is a comparable number of protoconch whorls (e.g. Sanders *et al.*, 2017) the study showed a strict ocean-delimited (e.g. Indian only, Pacific only) distribution. As for *C. phymatias*, the (eroded) protoconch of the holotype of *B. carynae* is to our knowledge the only one preserved (but slightly eroded) in this species, and it shows no major difference

(diameter/height/number of whorls) from other *Cryptogemma* species with regards to this character that could explain its very large Pacific-Atlantic distribution.

*Morphological stasis in teleoconch.* *C. praesignis* and *C. timorensis*, previously assigned to *Ptychosyrinx* in Puillandre *et al.* (2012), show considerable similarity in shell sculpture despite not being sister species, and the two taxa would be difficult to distinguish without their significant difference in size. These two species represent a clear case of morphological homogeneity in Conoidea, for which other examples have been reported (Kantor, Fedosov & Puillandre, 2018). Large-scale morphometrics correlated with phylogeny may serve to better characterize and understand the conserved ‘*Gemmula*-like’ form across geological time. Some specimens from the Mio-Pliocene have been attributed to species of *Cryptogemma* (see ‘Taxonomy’), however, an incommensurable number of ‘*Ptychosyrinx*-like’ or ‘*Gemmula*-like’ Turridae from the Paleogene will probably remain unattributed to a clade until a more comprehensive understanding, characterization and formalization of the family is achieved.

*Radular morphology.* The radula is an important morphological character in conoidean taxonomy. In Turridae, the radula is seemingly quite similar between congeneric species with few exceptions, such as for *Iotyrris* (Abdelkrim *et al.*, 2018b). *Cryptogemma* presently is one of few genera of Turridae for which the radula has been examined for every representative species. Morphology of the marginal duplex teeth is rather similar among all species, except for slight variation in the development of the accessory limb; either shorter or nearly the same in length as the major limb (*C. timorensis*), or much narrower or nearly of the same width as the major limb (*C. tessellata*). However, we note this with the caveat that the shape of the duplex teeth as inferred from SEM images may differ considerably due to the position of the tooth, thus rendering comparison of different specimens and species imperfect. Additionally, there are large variations between species in the morphology of the central formation in terms of the shape of the central cusp and the degree of development and shape of the lateral flaps, in some cases with very distinct elevated edges (*C. tessellata*), in others with the flaps being barely distinguishable (*Cryptogemma phymatias*). Another potentially important character is the relative length of the radula, which may differ greatly between otherwise morphologically close species (e.g. *C. praesignis* and *C. timorensis*), number of transverse rows of teeth (in *Cryptogemma* varying from 45 to 99) and the relative length of the marginal teeth (from 1.22% of AL in *C. powelli* to 1.9% in *C. praesignis*). Unfortunately, little has been published on relative tooth size in different Turridae lineages and other groups of Conoidea to date.

The above-mentioned characteristics of the *Cryptogemma* species have been revealed by the morphological and molecular delimitation of species, a practice strongly advised for future studies on the Turridae. In particular, it is shown here that morphological differences associated with geographical or ecological diversity within the distribution range of a species may lead to erroneous taxonomic conclusions. Unexpectedly wide geographic distributions in particular, such as those documented herein, further add to the challenges of conoidean taxonomy, notably where morphologically heterogeneous taxa are concerned. This is because material potentially seen as too widely separated geographically to be considered relevant (or where geographically well-separated taxa have not even been considered), may in fact represent conspecific material, a notion supported by the numerous synonymizations herein. Combined with the fact that sampling of the deep sea is infrequent and geographically severely incomplete (potentially masking the presence of morphological clines which results in the over-emphasis on few, morphologically discrete forms), future studies on deep-sea Conoidea must take such wide distributions into account when embarking upon taxonomic investigation.

Finally, the study of shell variability may serve to produce hypotheses regarding the evolutionary processes that lead to such observed diversity. Indeed, the fact that the actual, molecular-based *Cryptogemma* species were previously split by ante-molecular taxonomists in several species each, has led us to a re-evaluation and re-interpretation of the intraspecific morphological diversity that we could have neglected otherwise. In this regard, the words of Darwin (1857) "It is good to have hair-splitters & lumpers." make sense: the splitter reminds lumpers that intraspecific variability should not be ignored, and the lumper reminds splitters that intraspecific variability should not be over-interpreted.

### **Acknowledgments**

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) and 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). The contribution of Y. Kantor and A. Fedosov was supported by the grant from Russian Science Foundation no. 16-14-10118-II (PI Y. Kantor), F. Criscione and A. Hallan supported by a grant from the Australian Biological Resources Study (ABRS grant RF217-57 to PI F.

Criscione), and Y. Kano by the Japan Society for the Promotion of Science (KAKENHI grant nos 15H04412 and 18H02494).

The material in this paper 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, GUYANE 2014, PAPUA NIUGINI, KAVIENG 2014; PI P. Bouchet)) and/or by MNHN and Institut de Recherche pour le Développement (IRD) as part of the Tropical Deep-Sea Benthos programme (BIOMAGLO, TARASOC, MIRIKY, AURORA 2007, EBISCO, TAIWAN 2013, DongSha 2014, NANHAI 2014, BIOPAPUA, SALOMON 2, SALOMONBOA 3, CONCALIS, EXBODI, NORFOLK 2, TERRASSES, ZHONGSHA 2015, KANACONO, KANADEEP, MADEEP; PIs Bertrand Richer de Forges, Philippe Bouchet, Sarah Samadi, Nicolas Puillandre, Wei-Jen Chen, Tin-Yam Chan). Scientific partners included the University of Papua New Guinea (UPNG); National Fisheries College, Kavieng, Papua New Guinea; the National Museum of the Philippines; Institut d'Halieutique et Sciences Marines (IH.SM), Université de Tuléar, Madagascar; University of Taipei and University of Keelung, Taiwan; 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. 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 Virginie Héros, Barbara Buge, and Julien Brisset (MNHN) for their help in curating the vouchers; Alexander Sysoev (ZMMU) for photographs and tissue loan; Christine Zorn (ZMB) for photographs and type specimens loan; Ellen Strong, Mark. S. Florence and Kathy Hollis (USNM) for photographs; Jochen Gerber (FMNH) for photographs; Philippe Maestrati (MNHN) for photographs; Regina Kawamoto & Norine Yeung (BPBM) for photographs; Severine Hannam (AWMM) for photographs; Andreia Salvador (NHMUK) for photographs; Jun Hashimoto (Nagasaki University) for organizing the Nagasaki-maru cruise N275; the Plateau Technique de Microscopie Électronique (PtME) of MNHN for SEM pictures. Posthumous thank goes to Richard Kilburn for photograph of *P. praesignis*.



## References

- Abdelkrim J, Aznar-Cormano L, Buge B, Fedosov A, Kantor Y, Zaharias P, Puillandre N. 2018a. Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Molecular Ecology* 27: 4591–4611.
- Abdelkrim J, Aznar-Cormano L, Fedosov AE, Kantor YI, Lozouet P, Phuong MA, Zaharias P, Puillandre N. 2018b. Exon-Capture-Based Phylogeny and Diversification of the Venomous Gastropods (Neogastropoda, Conoidea). *Molecular Biology and Evolution* 35: 2355–2374.
- Barnard KH. 1958. Contribution to the knowledge of South African marine Mollusca. Part 1. Gastropoda; Prosobranchiata: Toxoglossa. *Annals of the South African Museum. Annale van die Suid-Afrikaanse Museum*. 44: 73–163.
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* 22: 148–155.
- Bierne Nicolas, Bonhomme François, David Patrice. 2003. Habitat preference and the marine-speciation paradox. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: 1399–1406.
- Bonhomme V, Picq S, Gaucherel C, Claude J. 2014. Momocs: outline analysis using R. *Journal of Statistical Software* 56: 1–24.
- Bouchet P, Kantor YI. 2004. New Caledonia: The major centre of biodiversity for volutomitrid molluscs (Mollusca: Neogastropoda: Volutomitridae). *Systematics and Biodiversity* 1: 467–502.
- Bouchet P, Kantor YI, Sysoev A, Puillandre N. 2011. A new operational classification of the Conoidea (Gastropoda). *Journal of Molluscan Studies* 77: 273–308.
- Bouchet P, Strong E. 2010. Historical Name-Bearing Types in Marine Molluscs: An Impediment to Biodiversity Studies? In: Polaszek A, ed. *Systema Naturae 250 - The Linnaean Ark*. CRC Press, 63–74.
- Bouchet P, Waren A. 1979. Planktotrophic larval development in deep-water gastropods. *Sarsia* 64: 37–40.
- Bouchet P, Warén A. 1994. Ontogenetic migration and dispersal of deep-sea gastropod larvae. *Reproduction, larval biology, and recruitment of the deep-sea benthos. Columbia University Press, New York*: 98–118.

- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, Maio ND, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, Plessis L du, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* 15: e1006650.
- Casanova JP, De Jong L, Faure E. 1998. Interrelationships of the two families constituting the Lophogastrida (Crustacea: Mysidacea) inferred from morphological and molecular data. *Marine Biology* 132: 59–65.
- Chan KO, Grismer LL, Brown RM. 2018. Comprehensive multi-locus phylogeny of Old World tree frogs (Anura: Rhacophoridae) reveals taxonomic uncertainties and potential cases of over- and underestimation of species diversity. *Molecular Phylogenetics and Evolution* 127: 1010–1019.
- Churchill CKC, Valdés Á, Foighil DÓ. 2014. Molecular and morphological systematics of neustonic nudibranchs (Mollusca : Gastropoda : Glaucidae : *Glaucus*), with descriptions of three new cryptic species. *Invertebrate Systematics* 28: 174–195.
- Clague GE, Jones WJ, Paduan JB, Clague DA, Vrijenhoek RC. 2012. Phylogeography of *Acesta* clams from submarine seamounts and escarpments along the western margin of North America. *Marine Ecology* 33: 75–87.
- Darwin Correspondence Project, “Letter no. 2130,” accessed on 31 July 2019, <http://www.darwinproject.ac.uk/DCP-LETT-2130>
- Dayrat B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85: 407–417.
- Dayrat B, Tillier A, Lecointre G, Tillier S. 2001. New Clades of Euthyneuran Gastropods (Mollusca) from 28S rRNA Sequences. *Molecular Phylogenetics and Evolution* 19: 225–235.
- Dejaco T, Gassner M, Arthofer W, Schlick-Steiner BC, Steiner FM. 2016. Taxonomist’s Nightmare ... Evolutionist’s Delight : An Integrative Approach Resolves Species Limits in Jumping Bristletails Despite Widespread Hybridization and Parthenogenesis. *Systematic Biology* 65: 947–974.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- Etter R, Rex MA, Chase MR, Quattro JM. 2005. Population Differentiation Decreases with Depth in Deep-Sea Bivalves. *Evolution* 59: 1479–1491.
- Fišer C, Robinson CT, Malard F. 2018. Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* 27: 613–635.

- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*: 6.
- Galindo LA, Puillandre N, Strong EE, Bouchet P. 2014. Using microwaves to prepare gastropods for DNA barcoding. *Molecular Ecology Resources* 14: 700–705.
- Giardina CR, Kuhl FP. 1977. Accuracy of curve approximation by harmonically related vectors with elliptical loci. *Computer Graphics and Image Processing* 6: 277–285.
- Hauquier F, Leliaert F, Rigaux A, Derycke S, Vanreusel A. 2017. Distinct genetic differentiation and species diversification within two marine nematodes with different habitat preference in Antarctic sediments. *BMC Evolutionary Biology* 17: 120.
- Havermans C, Sonet G, d’Acoz C d’Udekem, Nagy ZT, Martin P, Brix S, Riehl T, Agrawal S, Held C. 2013. Genetic and Morphological Divergences in the Cosmopolitan Deep-Sea Amphipod *Eurythenes gryllus* Reveal a Diverse Abyss and a Bipolar Species. *PLOS ONE* 8: e74218.
- Iglésias SP, Toulhoat L, Sellos DY. 2010. Taxonomic confusion and market mislabelling of threatened skates: important consequences for their conservation status. *Aquatic Conservation: Marine and Freshwater Ecosystems* 20: 319–333.
- Jennings RM, Bucklin A, Ossenbrügger H, Hopcroft RR. 2010. Species diversity of planktonic gastropods (Pteropoda and Heteropoda) from six ocean regions based on DNA barcode analysis. *Deep Sea Research Part II: Topical Studies in Oceanography* 57: 2199–2210.
- Jovelin R, Justine JL. 2001. Phylogenetic relationships within the polyopisthocotylean monogeneans (Platyhelminthes) inferred from partial 28S rDNA sequences. *International Journal for Parasitology* 31: 393–401.
- Kabat AR. 1996. Molluscan types of the Albatross expeditions to the Eastern Pacific described by WH Dall (1908): Bulletin of the Museum of Comparative Zoology.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587–589.
- Kantor YI, Fedosov AE, Puillandre N. 2018. New and unusual deep-water Conoidea revised with shell, radula and DNA characters. *Ruthenica* 28.
- Kantor YI, Puillandre N. 2012. Evolution of the Radular Apparatus in Conoidea (Gastropoda: Neogastropoda) as Inferred from a Molecular Phylogeny. *Malacologia* 55: 55–90.
- Kantor YI, Sysoev AV. 1991. Sexual Dimorphism in the Apertural Notch of a New Species of *Gemmula* (Gastropoda: Turridae). *Journal of Molluscan Studies* 57: 205–209.

- Kapli P, Lutteropp S, Zhang J, Kobert K, Pavlidis P, Stamatakis A, Flouri T. 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 33: 1630–1638.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* 35: 1547–1549.
- Lefébure T, Douady CJ, Gouy M, Gibert J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* 40: 435–447.
- Macaya EC, Zuccarello GC. 2010. Dna Barcoding and Genetic Divergence in the Giant Kelp *Macrocystis* (laminariales)1. *Journal of Phycology* 46: 736–742.
- Martin K. 1935. *Oligocaene Gastropoden von Buton*. Rijksmuseum van Geologie en Mineralogie.
- Meyer A, Bleidorn C, Rouse GW, Hausen H. 2008. Morphological and molecular data suggest a cosmopolitan distribution of the polychaete *Proscoloplos cygnochaetus* Day, 1954 (Annelida, Orbiniidae). *Marine Biology* 153: 879–889.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* 32: 268–274.
- Pante E, Schoelinck C, Puillandre N. 2015. From Integrative Taxonomy to Species Description: One Step Beyond. *Systematic Biology* 64: 152–160.
- Powell AWB. 1964. The family Turridae in the Indo-Pacific. Part 1. The subfamily Turrinae. *Indo-Pacific mollusca*. 1: 227–346.
- Puillandre N, Cruaud C, Kantor YI. 2010. Cryptic species in *Gemmuloborsonia* (Gastropoda: Conoidea). *Journal of Molluscan Studies* 76: 11–23.
- Puillandre N, Fedosov AE, Zaharias P, Aznar-Cormano L, Kantor YI. 2017. A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess. *Zoological Journal of the Linnean Society* 181: 243–271.
- Puillandre N, Kantor YI, Sysoev A, Couloux A, Meyer C, Rawlings T, Todd JA, Bouchet P. 2011. The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). *Journal of Molluscan Studies* 77: 259–272.
- Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S. 2012. Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* 21: 2671–2691.

- Robba E, Sartono S, Violanti D, Erba E. 1989. Early Pleistocene gastropods from Timor. *Memorie di Scienze Geologiche* 41: 61–113.
- Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Sanders MT, Merle D, Bouchet P, Castelin M, Beu AG, Samadi S, Puillandre N. 2017. One for each ocean: revision of the *Bursa granularis* species complex (Gastropoda: Tonnoidea: Bursidae). *Journal of Molluscan Studies* 83: 384–398.
- Sarkissian CD, Pichereau V, Dupont C, Ilsøe PC, Perrigault M, Butler P, Chauvaud L, Eiríksson J, Scourse J, Paillard C, Orlando L. 2017. Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past. *Molecular Ecology Resources* 17: 835–853.
- Simon C, Franke A, Martin A. 1991. The Polymerase Chain Reaction: DNA Extraction and Amplification. In: Hewitt GM,, In: Johnston AWB,, In: Young JPW, eds. *NATO ASI Series. Molecular Techniques in Taxonomy*. Berlin, Heidelberg: Springer Berlin Heidelberg, 329–355.
- Sun Y, Wong E, Tovar-Hernández MA, Williamson JE, Kupriyanova EK. 2016. Is *Hydroides brachyacantha* (Serpulidae : Annelida) a widespread species? *Invertebrate Systematics* 30: 41–59.
- Sysoev A. 1996. Deep-sea conoidean gastropods collected by the John Murray Expedition, 1933-34. *Bulletin of the Natural History Museum*. 62: 1–30.
- Thiele J. 1929. *Handbuch der systematischen weichtierkunde*. Jena: G. Fischer.
- Todd JA, Rawlings TA. 2014. A review of the Polystira clade—the Neotropic’s largest marine gastropod radiation (Neogastropoda: Conoidea: Turridae sensu stricto). *Zootaxa* 3884: 445–491.
- Tomioka S, Kondoh T, Sato-Okoshi W, Ito K, Kakui K, Kajihara H. 2016. Cosmopolitan or Cryptic Species? A Case Study of *Capitella teleta* (Annelida: Capitellidae). *Zoological Science* 33: 545–554.
- Trickey JS, Thiel M, Waters JM. 2016. Transoceanic dispersal and cryptic diversity in a cosmopolitan rafting nudibranch. *Invertebrate Systematics* 30: 290–301.
- Tucker JK. 2004. Catalog of Recent and fossil turrids (Mollusca: Gastropoda). *Zootaxa* 682: 1–1295.
- Villanea FA, Parent CE, Kemp BM. 2016. Reviving Galápagos snails: ancient DNA extraction and amplification from shells of probably extinct endemic land snails. *Journal of Molluscan Studies* 82: 449–456.

Will KW, Mishler BD, Wheeler QD. 2005. The Perils of DNA Barcoding and the Need for Integrative Taxonomy. *Systematic Biology* 54: 844–851.

Zardus JD, Etter RJ, Chase MR, Rex MA, Boyle EE. 2006. Bathymetric and geographic population structure in the pan-Atlantic deep-sea bivalve *Deminucula atacellana* (Schenk, 1939). *Molecular Ecology* 15: 639–651.

### **Supplementary Material**

The supplementary figures cited in the text are available as electronic supplementary material.



## **7. Résultat du pipeline de taxonomie intégrative sur l'ensemble des Turridae**

Le nombre de spécimens et de séquences CO1, dépendants des missions d'exploration et des moyens mis en œuvre pour séquencer les spécimens, ainsi que la délimitation même de la famille Turridae, a évolué au cours de cette thèse. Par exemple, le genre *Gemmuloborsonia*, placé par Bouchet *et al.* (2011) chez les Clavatulidae, a été déplacé chez les Turridae après l'étude phylogénomique d'Abdelkrim *et al.* (2018a - voir Annexe 2) sur les Conoidea. L'analyse du pipeline de taxonomie intégrative effectuée pendant mon stage de Master 2, sur l'ensemble des Turridae (2296 séquences CO1) à l'exception des genres *Gemmuloborsonia*, *Lucerapex*, *Kuroshioturris* et '*Ptychosyrinx*' *chilensis*, indique qu'en fonction de l'attitude du taxinomiste, *lumper* ou *splitter*, le nombre de MOTUs secondaires chez les Turridae peut passer de 173 à 196, ce qui peut avoir des conséquences non négligeables sur les hypothèses de spéciation (cf. Coyne et Orr, 2004) et l'échantillonnage d'espèces en amont d'analyses macroévolutives (Chapitre 3).

Les résultats du pipeline (Fig. 1) à l'issue de cette thèse indiquent 190 MOTUs sur l'ensemble des Turridae (~2900 séquences CO1) du MNHN. Ce résultat est à mettre en perspective avec les quelques 208 espèces de Turridae décrites (Worms, 2019). Les deux principaux logiciels utilisés sur l'ensemble du jeu de données CO1 sont ABGD et PTP (car ce sont les deux plus rapides). Les principaux critères utilisés pour valider sont (dans l'ordre) la monophylie des hypothèses primaires (ABGD peut proposer des groupes paraphylétiques), la confirmation de cette monophylie par le marqueur 28S (quand il est disponible), et les combinaisons du type de protoconque, de la répartition géographique et bathymétrique chez les MOTUs frères. Toutefois ces résultats sont à prendre avec beaucoup de précautions, car le processus de taxonomie intégrative est plus abouti chez les groupes pour lesquels une étude a été publiée (par exemple *Iotyrris/Xenuroturris* – Annexe 3, *Lophiotoma* – Annexe 1, *Cryptogemma* – voir article 1 dans ce Chapitre) que chez les autres groupes. De plus, l'étape (iv) de révision nomenclaturale et de description n'a pas été effectuée pour les groupes n'ayant pas fait l'objet de publication, pour des raisons déjà évoquées plus haut.

Malgré tout, ces unités opérationnelles restent certainement plus fiables que les hypothèses issues de l'approche traditionnelles (morpho-espèces) chez les gastéropodes (voir Introduction), et nous permettent quelques considérations sur la magnitude de la diversité chez les Turridae. Ainsi, il apparaît que la majorité des MOTUs (68,6%) sont représentés par moins de 10 spécimens, et 20,7% par des singletons, tandis que seulement 13,3% des MOTUs sont représentés par plus de 30 spécimens (Fig. 4). Des résultats différents de ceux trouvés par

Bouchet *et al.* (2009) où le pourcentage d'espèces représentées par des singletons est de 41% tandis que celui des espèces représentées par plus de 30 spécimens n'est que de 4,5%. Toutefois l'étude a été effectuée sur les Turridae s.l. et seulement dans la région de Nouvelle-Calédonie. Si l'on restreint notre étude sur les spécimens de la région Nouvelle-Calédonie uniquement, alors la proportion d'espèces représentées par des singletons passe à 38,8% (comparable à l'étude de Bouchet), mais 10,2% des MOTUs sont représentés par plus de 30 spécimens. Loin de permettre de tirer des conclusions sur l'impact de la délimitation des espèces dans l'étude de la magnitude de la biodiversité, la comparaison amène plusieurs questions : Les Turridae s. s. sont-ils plus abondants que le reste des Conoidea ? Ou bien les espèces des autres clades de Conoidea sont-elles tout simplement moins abondantes seulement en Nouvelle-Calédonie ? Ou encore, il y-a-t-il des espèces pour lesquelles on ne retrouve que des coquilles vides (déplacées) et que nos méthodes d'échantillonnage ne permettent pas de capturer à l'état vivant ?

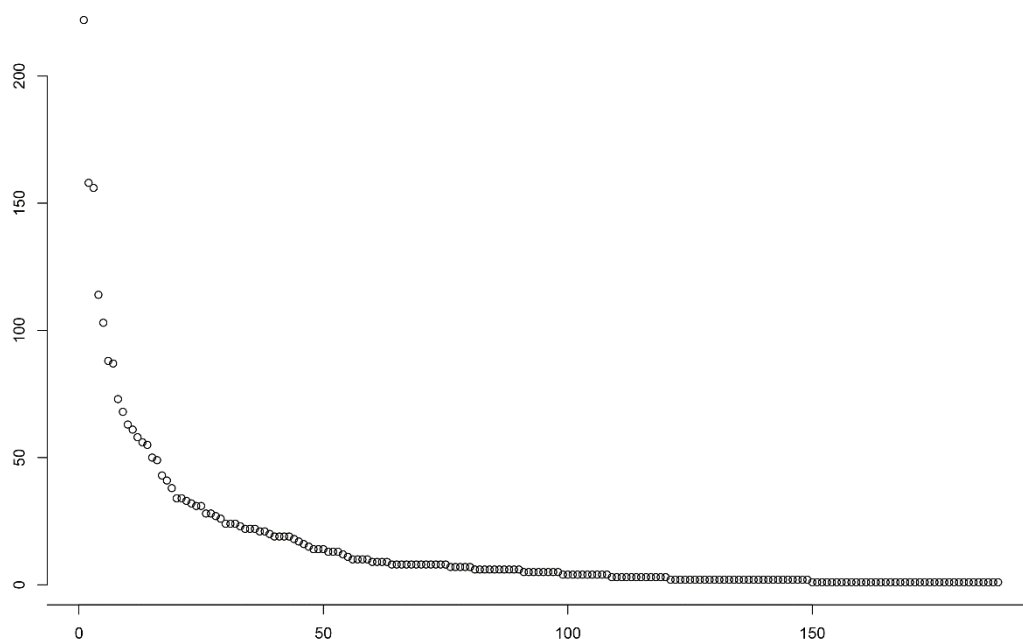


Figure 4: Courbe abondance-rang représentant le nombre de spécimens (ordonnée) par MOTU (abscisse)

L'évolution du nombre d'espèces nouvelles à chaque nouvelle campagne permet aussi d'évaluer la complétude de l'échantillonnage de la diversité spécifique des Turridae. On observe un ralentissement du nombre de nouvelles espèces délimitées à la suite des cinq dernières campagnes étudiées dans cette thèse (Fig. 5). Toutefois ces résultats ne prennent pas en compte les spécimens non collectés pendant les campagnes TDSB (voir par exemple l'article *Xenuroturris* – Annexe 3), ainsi que les spécimens issus des bases de données en

ligne comme GenBank. Il est à noter également que les 87 MOTUs trouvés par Puillandre *et al.* (2012b) – dont l'échantillonnage s'arrête à la mission AURORA 2007 – ne correspondent pas aux 104 MOTUs délimités dans cette thèse pour les mêmes campagnes. Plusieurs raisons peuvent être invoquées. La première est que l'ajout de nouveaux spécimens et d'espèces peuvent modifier les hypothèses d'espèces pré-établies. Mais surtout, le séquençage de l'ensemble des spécimens d'une mission n'aboutit parfois qu'après plusieurs années. Beaucoup de spécimens, notamment issus de la campagne SANTO 2006 n'ont été séquencés qu'après la publication de l'article par Puillandre *et al.* Cette remarque suggère que cette courbe peut être modifiée de façon rétroactive, au gré des financements et de l'effort de séquençage des missions précédentes. Cependant, nous pouvons supposer que nous avons collecté des spécimens de la majorité des espèces présentes dans la zone échantillonnée et qu'il est peu probable que le nombre d'espèces de Turridae aille au-delà du double de celui estimé actuellement. Des indices simples d'estimation du nombre d'espèces totales, comme l'indice Chao1 (Chao, 1984 ; Gotelli & Chao, 2013), permettent de confirmer cette intuition. L'indice Chao1 utilise l'information du nombre de singletons et de doubletons (espèces représentées par deux spécimens) pour estimer la limite inférieure de l'asymptote des courbes de raréfaction (voir figure 4). Pour l'échantillonnage utilisé dans la figure 4, l'indice Chao1 estime la diversité totale des Turridae à environ 214 espèces. Il faut toutefois garder à l'esprit que ce chiffre est probablement sous-estimé car nous n'avons pas inclus tous les spécimens/espèces séquencés (notamment de Genbank) dans cet échantillonnage. Dans l'Indo-Pacifique, il est peu probable que l'échantillonnage de zones encore non-explorées par les campagnes TDSB, comme la Mer d'Oman ou le Golfe du Bengale, aient un impact très important sur le nombre d'espèces total évalué, surtout si l'on considère que la majorité des espèces de Turridae sont planctotrophes. Toutefois la question se pose pour les genres pour lesquels beaucoup d'espèces possèdent des larves non-planctotrophes (*Lucerapex*, *Polystira*, *Turridrupa*). L'exploration plus complète de certaines localités pourrait révéler une diversité mal évaluée chez ces groupes, comme le suspectent certains auteurs (Todd & Rawlings, 2014).

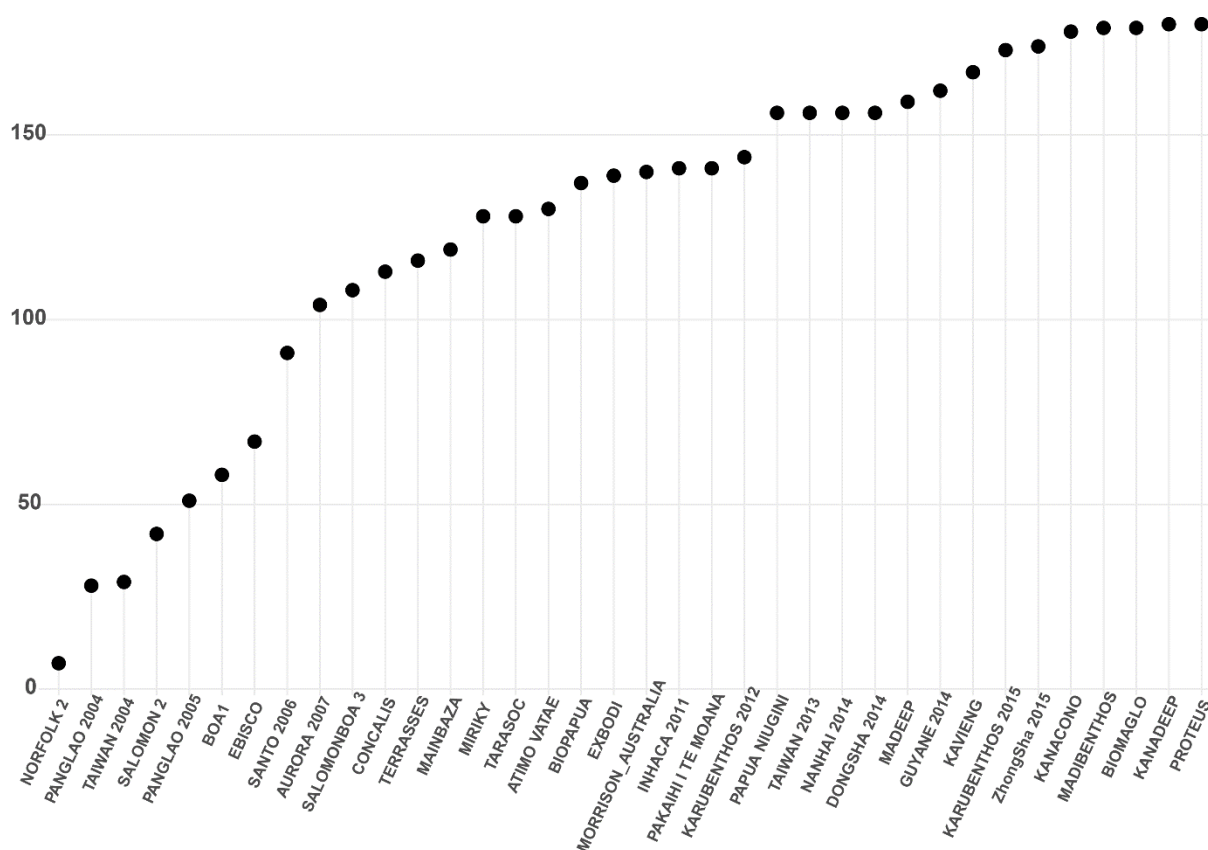


Figure 5: Courbe cumulative du nombre de nouveaux MOTUs (ordonnée) par campagne TDSB référencée (abscisse)

Si la courbe d'accumulation des espèces semble atteindre un plateau, cela n'est pas le cas pour l'échantillonnage par région. Les courbes cumulatives de sept régions explorées au moins trois fois ne semblent pas atteindre un plateau, hormis peut-être pour les Caraïbes (Fig. 6g) et pour la Nouvelle-Calédonie (Fig. 6a), même si le nombre de campagnes est trop faible pour véritablement en tirer des conclusions. Au niveau régional donc, contrairement au niveau global, la plupart des espèces n'ont pas encore été trouvées, ou du moins pas encore séquencées. Ce constat vient appuyer une conclusion déjà proposée dans l'Introduction : l'échantillonnage actuel est insuffisant pour répondre à des questions de distribution géographique, d'écologie ou d'abondance des espèces. L'effort d'échantillonnage doit donc se poursuivre, même dans les zones géographiques déjà explorées par plusieurs missions.

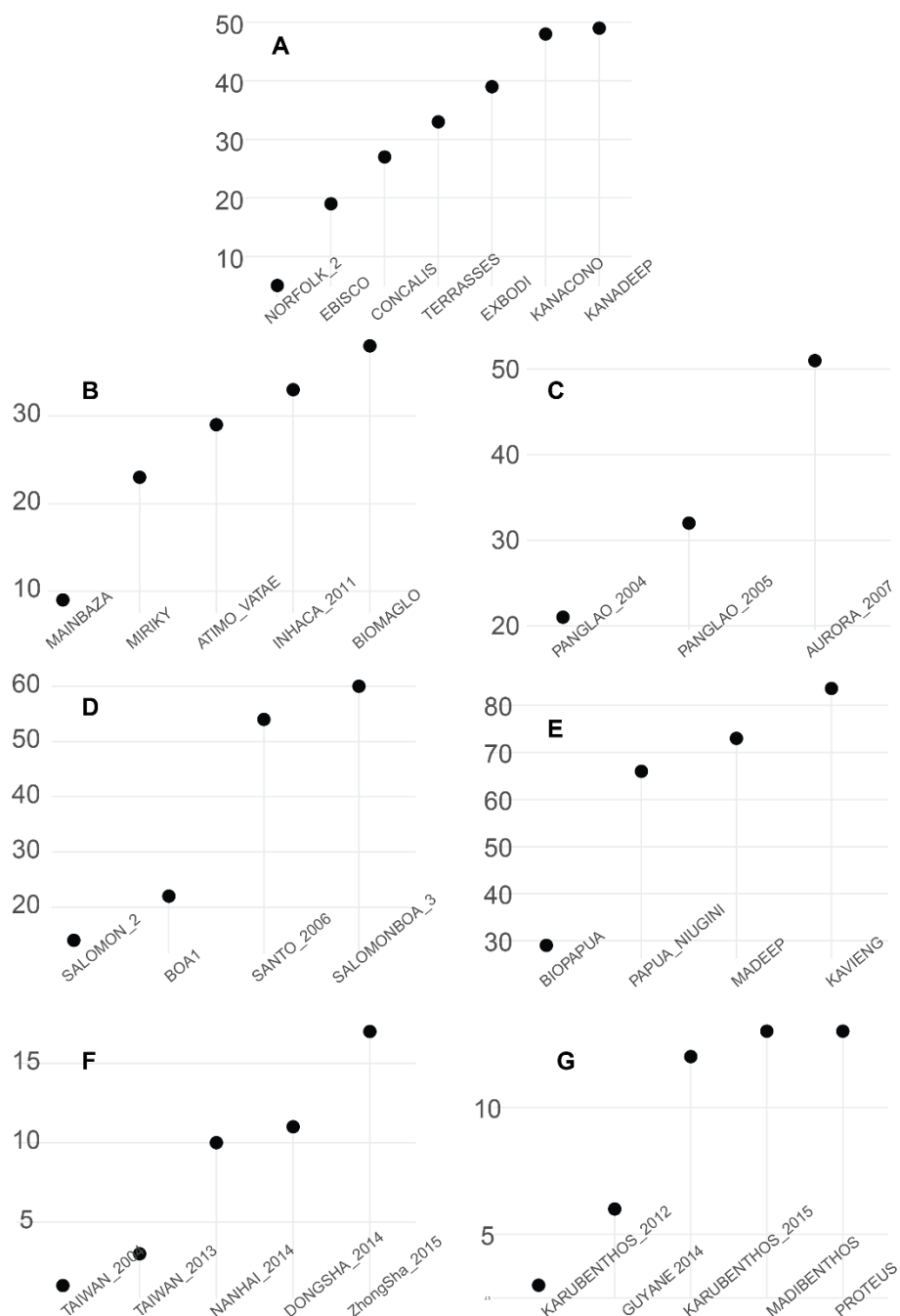


Figure 6: Courbe cumulative du nombre de nouveaux MOTUs (ordonnée) par région géographique. A. Nouvelle-Calédonie. B. Madagascar-Mozambique C. Philippines. D. Salomon-Vanuatu. E Papouasie. F. Taïwan. G. Caraïbes.

## 8. Arbre phylogénétique des MOTUs

Un spécimen par MOTU a été sélectionné afin de reconstruire un arbre phylogénétique des Turridae sur la base du marqueur CO1. Aux 190 terminaux obtenus, j'ai ajouté 14 représentants d'autres espèces (principalement côtières) que je ne possédais pas dans mon échantillonnage. Les séquences sont principalement issues de l'étude de Fedosov *et al.* (2011)

sur le genre *Turris* et de Todd & Rawlings (2014) sur le genre *Polystira*. Un arbre de maximum de vraisemblance a été construit en utilisant le logiciel IQ-tree (Nguyen *et al.*, 2014). L'alignement CO1 a été partitionné par position de codon et le meilleur modèle d'évolution pour chaque partition a été choisi en utilisant le logiciel ModelFinder, tel qu'implémenté dans IQ-tree (Kalyaanamoorthy *et al.*, 2017) et en suivant un critère BIC. Ensuite, 1000 « ultrafast » bootstraps (Hoang *et al.*, 2017) ont été calculés pour obtenir des valeurs de support pour chaque nœud. Enfin, les nœuds qui étaient supportés à moins de 95% ont été « collapsés », afin de ne retenir que les clades les mieux supportés (Fig. 7).

L'arbre obtenu permet de retrouver les groupes « connus », pour lesquels des études récentes ont été réalisées par notre équipe (*Cryptogemma*, *Iotyrris/Xenuroturris*, *Gemmuloborsonia*, *Lophiotoma*) ou des groupes étudiés récemment par d'autres (*Polystira*, *Turris*). En revanche, pas moins de 24 lignées de *Gemmula* (supportées à 95%) sont trouvées. Etant données la très probable polyphylie de ce genre, et la très probable monophylie des autres genres déjà étudiés, notre stratégie d'échantillonnage d'unités terminales pour la phylogénie (Chapitre 3) a consisté à prioriser les représentants de lignées de *Gemmula*, et à échantillonner un peu moins les autres genres déjà connus (*Lophiotoma*, *Gemmuloborsonia*, etc.). Ainsi, la phylogénie finale permettra la validation ou non de la monophylie des différentes lignées de *Gemmula* identifiées par cet arbre CO1. Au besoin, les relations au sein des autres lignées sous échantillonnées (*Lophiotoma*, *Gemmuloborsonia*, etc.) pourront être complétés *a posteriori* avec un échantillonnage réduit de loci.

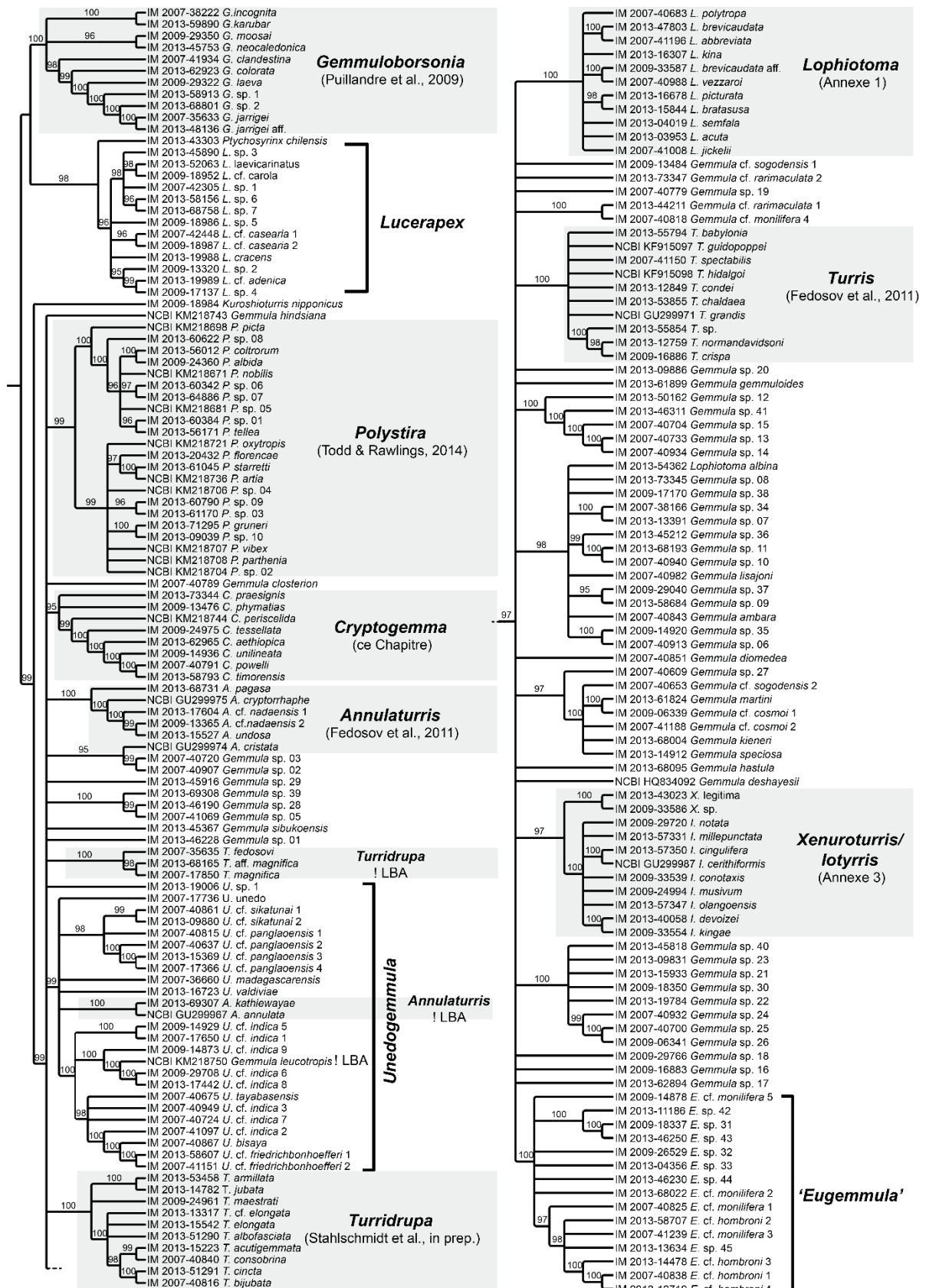


Figure 7: Arbre ML sur la base du marqueur CO1 sur 204 MOTUs de Turridae. Les cases grisées correspondent à des groupes pour lesquels des études récentes impliquant des données moléculaires existent. L'indication « !LBA » indique que ces MOTUs ne devraient pas se trouver à cet endroit de l'arbre et que leur placement est certainement dû à un effet d'attraction de longues branches (Long Branch Attraction en anglais). Les groupes indiqués en accolade sont des groupes qui n'ont pas été étudiés récemment mais qui sont homogènes morphologiquement et pour lesquels il y a un fort soupçon de monophylie.



## 9. Conclusion

D'un point de vue empirique, le pipeline de taxonomie intégrative mis en place pour les Turridae fonctionne de manière optimale, permettant de délimiter rapidement des dizaines de MOTUs qui peuvent maintenant être utilisés pour des études macro-évolutives (voir Chapitre 3) et/ou faire l'objet de révisions nomenclaturales. La stratégie suivie ici est développée dans mon équipe d'accueil depuis plusieurs années, et permet de proposer des hypothèses d'espèces plus robustes et précises. Toutefois, elle reste limitée par l'échantillonnage de certains MOTUs (singletons ou doubletons), le manque de marqueur nucléaire pour un certain nombre d'entre eux et la nécessité d'obtenir des données écologiques supplémentaires (par exemple, sur l'habitat).

L'étude plus poussée du genre *Cryptogemma* a permis de mettre en évidence un cas peu documenté en taxonomie intégrative des gastéropodes de *lumping* de noms d'espèces. Plus important : l'étude a mis en évidence des distributions très larges et recouvrantes chez les espèces de ce groupe, dont une distribution inédite (en tous cas chez les gastéropodes) Pacifique-Atlantique. Ce constat renforce l'idée de spéciation écologique comme modèle privilégié chez les Turridae, même si nous l'avons critiqué dans l'Introduction. Le couple d'espèce *Cryptogemma praesignis*/*C. phymatias*, avec sa séparation bathymétrique à 1500m, pourrait servir de modèle pour tester une hypothèse bathymétrique ou écologique de spéciation, en intégrant des données sur les proies et le venin.

La stratégie de délimitation d'espèces utilisée dans ce travail permet également de construire un arbre exploratoire (avec le marqueur CO1) qui permet de définir un jeu de données plus homogène pour les analyses phylogénétiques. Beaucoup d'études phylogénétiques moléculaires n'intègrent qu'un ou deux spécimens séquencés par morphoespèce ou par espèce décrite (par exemple Phuong *et al.*, 2019). Cette stratégie peut créer des jeux de données déséquilibrés s'il existe des biais d'identification morphologique ou des cas de sur-descriptions d'espèces dans le groupe d'étude (voir le cas de *Lautoconus* dans Phuong *et al.*, 2019). La stratégie exploratoire utilisée dans ce travail garantit un peu plus l'homogénéité de l'échantillonnage d'unités terminales pour la phylogénie car nos hypothèses a priori sont fondées sur un échantillonnage des individus à l'aveugle. De plus, l'arbre exploratoire CO1 permet d'anticiper les cas problématiques (par exemple les longues branches) et d'envisager certaines précautions (par exemple, intégrer plus ou moins de représentants pour les lignées problématiques). Nos 190 MOTUs constituent maintenant les unités de bases de notre étude. Cela ne veut pas dire pour autant qu'elles ne peuvent pas être

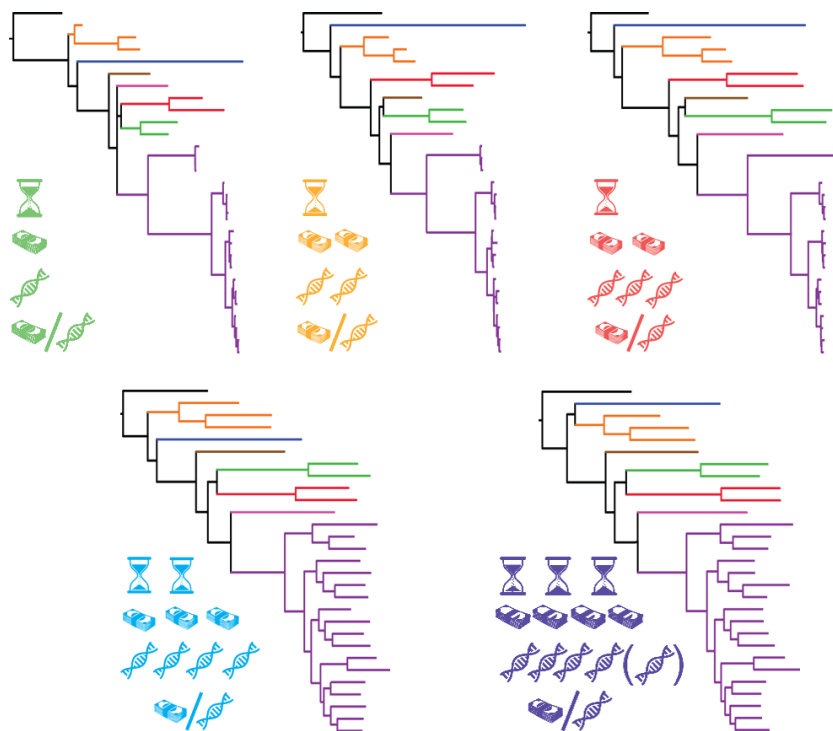
## Chapitre 1

remises en cause. Le RAD-Seq (Abdelkrim *et al.*, 2019b – Annexe 3) a remis en perspective l'utilisation de seulement quelques marqueurs pour délimiter les espèces et a pointé le besoin impérieux de corroborer les hypothèses CO1/28S avec d'autres marqueurs nucléaires, qu'il faut identifier. De plus, il nous reste à évaluer la stratégie d'échantillonnage des données moléculaires pour reconstruire la phylogénie du groupe. Ces deux objectifs ont été explorés durant cette thèse et sont présentés dans le Chapitre suivant.

# CHAPITRE 2

## UTILISATION EXPERIMENTALE DES

### TRANSCRIPTOMES EN SYSTEMATIQUE



Comme discuté dans le chapitre précédent, une des limitations du pipeline de taxonomie intégrative proposé est le faible nombre de marqueurs moléculaires indépendants pour tester les hypothèses d'espèces basées sur le CO1. La recherche de tels marqueurs était un des objectifs de cette thèse, et ce chapitre devait initialement être consacré à la recherche et aux tests *in silico* et empiriques de nouveaux marqueurs moléculaires pour les Neogastropoda. Ces tests devaient être effectués à partir de plusieurs dizaines de transcriptomes dont une grande partie ont été préparés, séquencés et assemblés durant ma thèse. Malheureusement, ces recherches n'ont pas abouti, en parti à cause de l'absence de génomes complets assemblés et annotés chez les Neogastropoda.

Cependant, l'utilisation expérimentale des transcriptomes a inspiré une autre étude, phylogénétique cette fois, sur l'identification de la meilleure stratégie pour établir des phylogénies moléculaires. Cette étude, détaillant également la façon dont les transcriptomes ont été échantillonnés et assemblés, sera présentée dans la première partie de ce chapitre. Dans la deuxième partie, une partie des résultats négatifs de la recherche de nouveaux marqueurs pour la délimitation d'espèces sera présentée.

### **1. Données, Temps, Argent : à la recherche du meilleur compromis pour construire des phylogénies moléculaires**

Il n'est pas rare de lire dans les premières lignes d'un article contemporain de biologie ou d'entendre dans la présentation d'un conférencier que l'acquisition de données de séquençage haut-débit (HTS pour high throughput-sequencing en anglais), par opposition au séquençage « traditionnel » de type Sanger, a révolutionné la pratique de la systématique moderne. Pourtant, depuis plus de 15 ans que les techniques HTS existent et se démocratisent, la majorité des études continuent d'utiliser et de produire des données de type Sanger, en particulier pour les études se focalisant sur des taxons non-modèles (voir Article 2 ci-après). Plusieurs causes, comme le manque de moyens ou d'expérience dans certaines équipes de chercheurs, pourraient expliquer le maintien de l'utilisation de marqueurs Sanger, mais une autre raison pourrait être qu'il n'est pas forcément nécessaire d'obtenir des milliers de marqueurs afin d'obtenir des arbres phylogénétiques robustes et utilisables pour des analyses méta-phylogénétiques, ou simplement pour répondre à des questions d'évolution.

L'objectif de ce travail est donc de proposer une synthèse sur les stratégies d'obtention des données pour l'établissement de phylogénies moléculaires, afin d'offrir un contexte et des

éléments de réponse sur les pratiques des phylogénéticiens. Pour cela nous avons utilisé 32 transcriptomes (dont 29 de Turridae) et reconstruit des phylogénies en utilisant différents jeux de données moléculaires extraits *in silico* de ces transcriptomes et représentatifs des jeux de données utilisés classiquement pour reconstruire des phylogénies. Les coûts, en temps et en argent, et la robustesse des arbres ont été évalués et intégrés dans la discussion. Une discussion sur les données manquantes, la robustesse des arbres et les méthodes utilisées est également proposée.

Ces résultats ont été publiés dans la revue *Molecular Phylogenetics and Evolution* (Article 2).

Cette étude a permis de valider et de justifier l'utilisation des méthodes de captures comme stratégie optimale pour l'établissement de phylogénies moléculaires, en particulier chez les Turridae, tout en maintenant l'idée que les approches Sanger restent intéressantes en approche préliminaire (voir Chapitre 1). Il y est proposé d'intégrer systématiquement les marqueurs traditionnels et les génomes d'organites dans les études utilisant des méthodes de capture, dans l'optique de combiner plus facilement des phylogénies entre elles.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)

## Data, time and money: evaluating the best compromise for inferring molecular phylogenies of non-model animal taxa

Paul Zaharias <sup>a</sup>, Eric Pante <sup>b</sup>, Delphine Gey <sup>c</sup>, Alexander E. Fedosov <sup>d</sup>, Nicolas Puillandre <sup>a</sup><sup>a</sup> Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 43 rue Cuvier, CP 26, 75005 Paris, France<sup>b</sup> Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS, La Rochelle Université, 2 rue Olympe de Gouges, 17042 La Rochelle, France<sup>c</sup> Acquisition et Analyses de Données pour l'histoire naturelle (2AD) UMS 2700, Muséum National d'Histoire Naturelle, Paris, France<sup>d</sup> A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninsky prospect 33, 119071 Moscow, Russian Federation

## ARTICLE INFO

## Keywords:

Phylogenomics  
Transcriptomics  
High throughput sequencing  
Sanger sequencing  
Non-model taxa  
Turridae

## ABSTRACT

For over a decade now, High Throughput sequencing (HTS) approaches have revolutionized phylogenetics, both in terms of data production and methodology. While transcriptomes and (reduced) genomes are increasingly used, generating and analyzing HTS datasets remain expensive, time consuming and complex for most non-model taxa. Indeed, a literature survey revealed that 74% of the molecular phylogenetics trees published in 2018 are based on data obtained through Sanger sequencing. In this context, our goal was to identify the strategy that would represent the best compromise among costs, time and robustness of the resulting tree. We sequenced and assembled 32 transcriptomes of the marine mollusk family Turridae, considered as a typical non-model animal taxon. From these data, we extracted the loci most commonly used in gastropod phylogenies ( *cox1*, 12S, 16S, 28S, *h3* and 18S), full mitogenomes, and a reduced nuclear transcriptome representation. With each dataset, we reconstructed phylogenies and compared their robustness and accuracy. We discuss the impact of missing data and the use of statistical tests, tree metrics, and supertree and supermatrix methods to further improve phylogenetic data acquisition pipelines. We evaluated the overall costs (time and money) in order to identify the best compromise for phylogenetic data sampling in non-model animal taxa. Although sequencing full mitogenomes seems to constitute the best compromise both in terms of costs and node support, they are known to induce biases in phylogenetic reconstructions. Rather, we recommend to systematically include loci commonly used for phylogenetics and taxonomy (i.e. DNA barcodes, rRNA genes, full mitogenomes, etc.) among the other loci when designing baits for capture.

## 1. Introduction

For over a decade now, high throughput sequencing (HTS) data has allowed not only the production of a substantial amount of DNA sequences relevant for phylogenetics, but also triggered many discussions on phylogenetic reconstruction methods (e.g. [Edwards, 2009](#); [Lemmon and Lemmon, 2013](#); [Leaché et al., 2015a](#); [Leaché and Oaks, 2017](#) ). Most authors concluded in the superiority of HTS approaches for reconstructing trees at all phylogenetic scales, especially in terms of robustness of the tree but also in the context of studying biological processes (e.g. introgression or horizontal transfer). Nevertheless, there is still a considerable amount of recent studies presenting trees obtained using first generation sequencing (chain-termination sequencing based on the incorporation of dideoxynucleotides, herein referred to as “Sanger sequencing” – e.g. [Heather and Chain, 2016](#) ). This technique

has typically been used to sequence a few loci amplified by PCR. Although the first commercial HTS technology was introduced in 2004 ([Mardis, 2008](#) ), phylogenetic studies using this technology were not exceeding 12% of the total molecular phylogenetic studies up until 2016 ([Fig. 1](#) ). In 2018, only 26% of molecular phylogenetic studies were based on HTS data. The simple, rapid, and stable standard protocols for producing and analyzing datasets based on Sanger sequencing data may explain why they are still primarily used in phylogenetic studies, while HTS-based studies remain costlier and more complex (both in terms of library preparation and data analysis). Thus, the sustained attractiveness of the Sanger sequencing approach to phylogenetics, indisputable in terms of number of published studies to date, contrasts with the premise that HTS data will allow us to “achieve phylogenomic Nirvana” ([Faircloth, 2013](#) ).

Multiple studies have explored tree robustness through the

Corresponding author.

E-mail address: [paul.zaharias@edu.mnhn.fr](mailto:paul.zaharias@edu.mnhn.fr) (P. Zaharias).<https://doi.org/10.1016/j.ympev.2019.106660>

Received 26 August 2019; Received in revised form 17 October 2019; Accepted 17 October 2019

Available online 19 October 2019

1055-7903/© 2019 Elsevier Inc. All rights reserved.

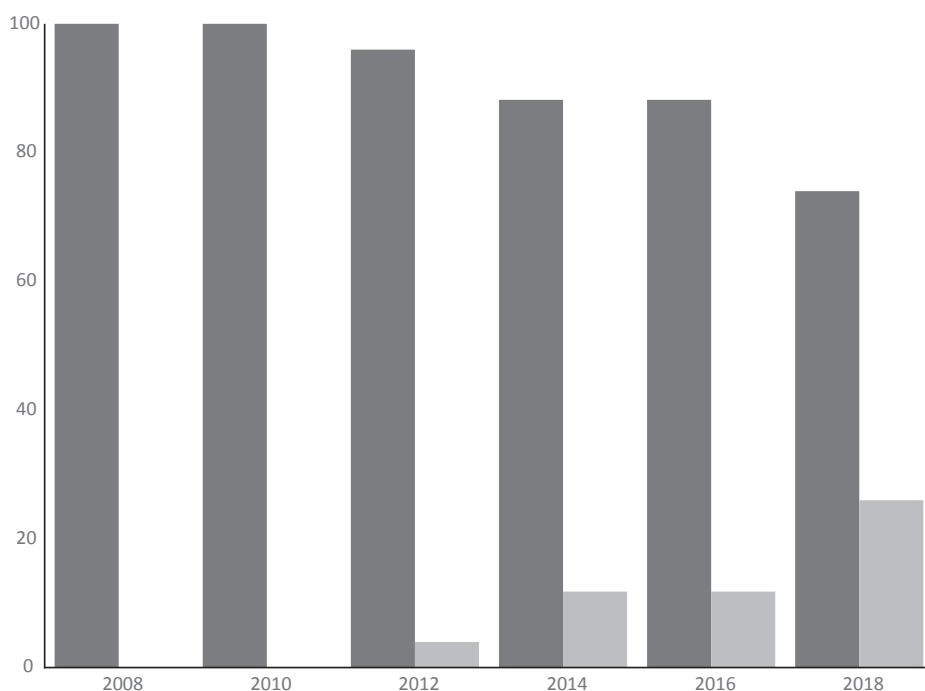


Fig. 1. Proportion of articles over time that used Sanger sequencing (dark grey) or HTS (light grey) to reconstruct a phylogeny. Articles were extracted using the Web of Science “Basic Search”, every two years from 2006 to 2018, and using the keyword “Phylogen\*” in TITLE only. Only the first 50 articles of the list with newly produced genetic data with one of the two methods were screened and categorized as “Sanger” or “HTS”.

dissection of a particular HTS dataset – the recently defined practice of “phylogenomic subsampling” (reviewed in Edwards, 2016). This practice mostly focuses on “the study of the information content of phylogenomic matrices of different sizes,” implying an *in silico* subsampling of loci (Edwards, 2016). One way of extending the practice of phylogenomic subsampling beyond its quantitative aspects would be to take into account the nature of particular loci. This approach would evaluate the phylogenetic significance of the use of particular genome regions (e.g. coding vs. non-coding sequences; Chen et al., 2017). The development of orthology assessment pipelines (e.g. UPhO; Ballesteros and Hormiga, 2016) also enabled their comparison with the same dataset, usually of transcriptomic nature (Washburn et al., 2017; Cuhna and Giribet, 2019). Other studies also focused on the sequencing method used to extract a particular set of loci, mostly leading to studies comparing HTS vs. Sanger sequencing methods to recover phylogenetic datasets and draw conclusions on the superiority of one dataset type over another (e.g. Ruane et al., 2015; Lee et al., 2018).

In addition to data exploration alone, the computational time needed to analyze various datasets - including the phylogenetic reconstruction - can be calculated (e.g. Leaché et al., 2015b). The time for sample preparation and sequencing can also be estimated (e.g. Lemmon et al., 2012), but this information is more often reported in HTS method description articles, and rarely compared among methods of data acquisition. Even fewer studies evaluated monetary costs (reagents, library preparation and sequencing) for a particular phylogenetic dataset (e.g. Moreau and Wray, 2017) or more generally for a taxonomic group (e.g. McKain et al., 2018). Finally, we only found two studies that evaluated both time and money in relation with the preparation of a specific phylogenetic dataset (Lemmon et al., 2012; Cruaud et al., 2014). These studies led to the conclusion that HTS data will produce more data and more robust trees, justifying the higher costs.

However, such studies are often conducted on so-called “model taxa” (e.g. Primates in Collins and Hrbek, 2018), for which genomic data is already abundant and the range of possible data acquisition methods is not limited by the absence of annotated genomes. However, >99% of the biodiversity is constituted of non-model taxa, i.e. taxa for which no or little genomic and/or transcriptomic data are available, and for which orthologous loci databases are information-poor. For those taxa, alternative strategies have been developed such as

exon-capture (Bi et al., 2012), Anchored hybrid Enrichment (Lemmon et al., 2012) or Ultraconserved elements (McCormack et al., 2012), referred herein as “sequence capture”. These strategies allowed phylogenomicists to utilize very distant genomic resources for specific non-model groups (e.g. Haddad et al., 2018). Still, lacking whole genome data limit the possibilities to a handful of loci (the highly conserved ones), exclude non-coding material from potential markers, and complicate the task of sorting orthology from paralogy.

Thus, most available studies focus on model taxa only, are generally limited to the comparison of datasets of either different quantity or different nature of data, but rarely both, using criteria related to tree robustness or time and money needed, but rarely both. To provide arguments to choose one strategy over another in phylogenetic reconstruction in non-model taxa, we here provide a comparison of several sequencing and tree-reconstruction strategies in terms of robustness of resulting trees, and time and money needed to produce and analyze the datasets. To do so, we used the family Turridae (Conoidea, Gastropoda), a group of marine molluscs, as an example. The Turridae constitute a good example of non-model animal taxon because of the lack of genomic resources (e.g. no assembled and annotated genome, less than ten nuclear markers represented in public databases, no karyotypes available), even in closely related groups. The closest reference would be the recently published genome of *Pomacea canaliculata* (Liu et al., 2018) and its divergence time with the family Turridae is estimated at 283 Ma (Zapata et al., 2014). The family comprises 216 species (WoRMS, checked on May 2019) but this number is largely underestimated (Puillandre et al., 2012). Most of the molecular phylogenies published for this group used the same classical mitochondrial (*cox1*, 12S and 16S rRNA) and/or nuclear (28S rRNA) markers (Heralde et al., 2007, 2010; Olivera et al., 2008; Puillandre et al., 2012; Fedosov et al., 2011; Todd and Rawlings, 2014; Puillandre et al., 2017). Two venom-gland transcriptomes are published (Gonzales and Saloma, 2014) that were not used primarily for phylogenetic purposes but for toxin research, because the Turridae are venomous and constitute a group of interest for bioactive compound discovery (Puillandre and Holford, 2010). Finally, only one phylogenomic (RAD-seq) study (Abdelkrim et al., 2018a) was published for species-delimitation purposes on eight species in the *Xenuroturrus/lotyrris* complex.

We analyzed 32 transcriptomes (29 Turridae + 3 outgroups),



corresponding to 18 species, from different tissues (venom gland, salivary gland or foot), from which we extracted *in silico* five datasets: (i) the barcode fragment (658 bp) of the *cox1* gene, the most frequently sequenced marker in gastropod systematics; (ii) a multilocus dataset that is typically produced using Sanger sequencing to conduct phylogenetic studies in gastropods, corresponding to fragments of the mitochondrial *cox1*, 16S rRNA and 12S rRNA genes, and the nuclear 28S rRNA, 18S rRNA and histone *h3* genes (e.g. Fedosov et al., 2018; Johnson et al., 2010 ); (iii) complete mitochondrial genomes (e.g. Uribe et al., 2018 ); (iv) a sequence capture approach, targeting a selection of nuclear loci (e.g. Abdelkrim et al., 2018b ); (v) an RNA-seq dataset (e.g. Cunha and Giribet, 2019 ). Because this dataset includes only a limited number of Turridae lineages, the goal is not to resolve the Turridae phylogeny, but to compare usefulness of those five datasets for phylogenetics. We empirically evaluated the capacity of each dataset to resolve relationships among the 32 samples, within which divergence ranges between 0 (intra-specimen divergence) and 79.4 Ma (estimated age of origin of the family Turridae; Abdelkrim et al., 2018b ). We also evaluated the time necessary for sample preparation, sequencing and data analysis, along with the monetary costs of each step to estimate the overall cost of producing each dataset.

## 2. Material and methods

### 2.1. Sampling

Twenty-eight specimens, representing six genera of Turridae and related outgroups (Conidae and Mitridae) were collected during several field expeditions organized by the Muséum national d'Histoire naturelle (MNHN; "KAVIENG" in Papua New Guinea, "KANACONO" in South New Caledonia), by joined Russian-Vietnamese Tropical Center (Vietnam), and by the University of Utah in collaboration with the University of the Philippines (Philippines). Specimens were photographed and the shells were broken to access the animal. For twenty-seven specimens, only one tissue type was sampled (venom gland, salivary gland or foot) depending on the project they were associated with; for one specimen, both venom gland and salivary gland tissue were sampled, resulting in a total of 29 tissues ( Supplementary Table 1 ). Remains of vouchers, when available, were kept and are deposited in the MNHN collections.

In addition, we used publicly available transcriptomes from three species: *Unedogemmula bisaya* Gemmula speciosa (Turridae) and *Terebra subulata* from a closely related family Terebridae (NCBI Sequence Read Archive (SRA) accession no.'s SRR1574923, SRR1574907 and SRR2060989, respectively; Gonzales and Saloma, 2014; Gorson et al., 2015 ).

### 2.2. RNA extraction, library preparation and sequencing

RNA was extracted using a Trizol protocol or the Qiagen RNeasy Micro kit, following the manufacturer's recommendations. Bioanalyzer traces were used to assess total RNA quality and determine suitability for sequencing. The cDNA libraries were prepared and sequenced either at the New York Genome Center or at the Evolutionary Genetics Lab at UC Berkeley ( Supplementary Table ). In New York, libraries were prepared using the automated polyA RNAseq library prep protocol and sequenced with Illumina HiSeq 4000 with 150-bp paired-end reads. In Berkeley, the KAPA Stranded mRNA-Seq kit was used to synthesize cDNA, ligate adapters using TruSeq HT adapters and barcode samples. Samples were then sequenced with Illumina HiSeq 2000 or 4000 (see Supplementary Table 1 ) with 100-bp paired-end reads.

### 2.3. Transcriptome assembly and quality assessment

All the transcriptomes, including the ones downloaded from GenBank, were assembled following the same procedure. Trimmomatic

v.0.36 ( Bolger et al., 2014 ) was used to remove adapters 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). Reads were merged using FLASH v1.2.8 ( Magoč and Salzberg, 2011 ) with a min. overlap parameter of 5, a max. overlap parameter of 100 and a mismatch ratio of 0.05. FastQC ( Andrews, 2010 ) was used for raw reads quality control. Transcripts were assembled using Trinity v2.4 with default parameters ( Grabherr et al., 2011 ). Cap3 ( Huang and Madan, 1999 ) with default parameters and cd-hit v4.6 (percent identity = 99%; Li and Godzik, 2006 ) were finally applied to reduce redundancy in the assemblies.

BBMap ( Bushnell, 2014 ) was used to generate basic assembly statistics and BUSCO ( Simão et al., 2015 ) to evaluate transcriptome completeness. Finally, bowtie2 v2.2.6 ( Langmead and Salzberg, 2012 ) and samtools v1.3 ( Li et al., 2009 ) were used to evaluate read representation in each assembled transcriptome, as recommended in the Trinity manual.

### 2.4. Transcriptome orthology inference

Two approaches were used to assess orthology among transcripts, from here onwards referred to as "reference-based" approach and "graph-based" (without a reference genome) approach ( Fig. 2 ).

For the reference-based approach, the *Pomacea canaliculata* genome (ASM307304v1; Liu et al., 2018 ) was used as a reference. Following the pipeline described in Phuong and Mahardika (2018) and Phuong et al. (2019), blastx was used to associate transcripts to peptide sequences of *P. canaliculata* and tblastn to associate peptides of *P. canaliculata* to transcripts from the BLAST + v2.2.31 suite ( Altschul et al., 1990 ) with an e-value threshold of  $1e10^{-10}$  and a word size value of 11. For each sample, bowtie2 v2.3.4.1 was used with the very sensitive-local alignment option and not allowing for discordant pair mapping (unexpected paired read orientation during mapping) to map reads to the selected transcripts from the reciprocal blast step. Duplicates were marked using picard-tools v2.0.1 ( <http://broadinstitute.github.io/picard> ) using default parameters. All positions with a coverage < 5X were masked and the entire sequence was removed if > 30% of the sequence was masked. To fix assembly errors, single nucleotide polymorphisms (SNPs) were called using samtools v1.3 (default parameters) and bcftools v1.3 ( Li et al., 2009 ) using the call command. Transcripts for each locus were aligned as nucleotides using MAFFT v7.222 ( Katoh et al., 2005 ) option -auto. To limit misalignments and paralogs inclusion, uncorrected pairwise distances were calculated at each locus for all possible pairwise comparisons and sequences were removed if the uncorrected pairwise distance was greater than the 90th percentile (threshold was set empirically) of pairwise distances across all loci for that pair of species.

For the graph-based approach, we used UPHO (Unrooted Phylogenetic Orthology; Ballesteros and Hormiga, 2016 ), a method that uses the topology of individual gene trees to identify clades corresponding to orthologous groups. Following the workflow established by the authors, all transcripts in open reading frame (ORF) were extracted from the transcriptome assemblies with custom Python scripts, and all ORFs that were less than 100 amino-acid long were eliminated. An all-versus-all blastp search was then performed, using a relaxed expectation value threshold of  $e = 1 \times 10^{-5}$ .

To reduce missing data, only the clusters that contained the maximum number of samples (32) were selected. The gene-family amino-acid sequence clusters were aligned and cleaned using mafft (option '-auto'), trimAL (option '-gappyout') and Al2phylo (-m 32 -t 300 -p 0.80). After alignments, the sequences were converted from amino acids back to nucleotides to increase the number of informative sites and improve the phylogenetic pipeline accuracy. Gene-family trees (GFTs) were estimated using IQ-tree ( Nguyen et al., 2014 ). The best

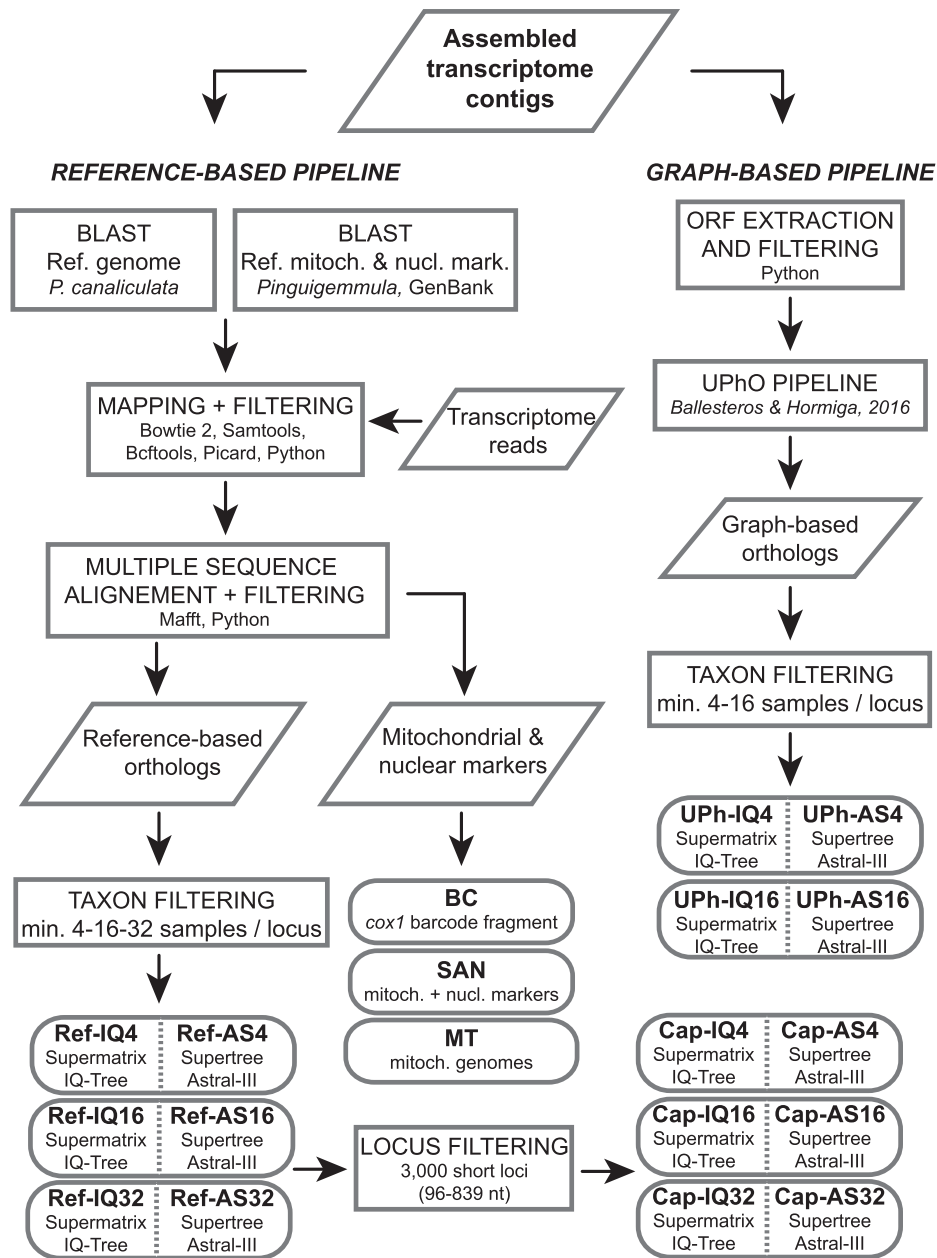


Fig. 2. Flowchart summarizing the in-silico approach used to generate all the datasets. Data are framed by parallelograms, tasks by rectangles and final datasets used for downstream analyses by rounded rectangles.

substitution model for each GFT was estimated with ModelFinder (Kalyanamoothy et al., 2017) following the BIC criterion. Subsequently, 1000 ultrafast bootstraps (UFBoot) (Hoang et al., 2017) were performed on each GFT to obtain branch support. The branches representing putative orthogroups were finally extracted with UPHo (-m 4 -S 0.80). The orthogroup alignments obtained were cleaned and analyzed using MAFFT, trimAL, Al2phylo and IQ-tree with the same parameters as above (except for the -m parameter in Al2phylo, set to 4).

### 2.5. Transcriptome-based phylogeny

Ten datasets were generated. For the reference-based approach three subsets were defined with a minimum of 4, 16 and 32 samples/locus. These subsets were analyzed using a supermatrix - concatenated alignment of all the loci - and a supertree approach, resulting in six datasets referred to as follows: Ref-IQ4, Ref-IQ16, Ref-IQ32, Ref-AS4, Ref-AS16 and Ref-AS32 (IQ referring to IQ-tree and AS to ASTRAL - see

below). Similar subsets were constructed for the graph-based approach with 16 and 4 samples/locus (the 32 sample/locus dataset was not analyzed here because only one locus was retrieved). The resulting four datasets are referred to as follows: Uph-IQ4, Uph-IQ16, Uph-AS4 and Uph-AS16.

Best substitution models were estimated for each partition (locus) in each concatenated dataset with ModelFinder following the BIC criterion. Supermatrix trees were reconstructed using IQ-tree and 1000 UFBoot were performed on each dataset. An individual tree for each locus was also generated with IQ-tree, using the associated best substitution model for datasets Ref-AS4, Ref-AS16, Ref-AS32, Uph-AS4 and Uph-AS16. The supertree approach implemented in the program ASTRAL-III (Zhang, Sayyari and Mirarab, 2017) was then applied to combine the single-locus trees into a single supertree for each of these datasets.

## 2.6. Sequence capture

We used the Ref-AS4 dataset and selected the 3000 shortest loci (ranging from 96 to 839 bp) to simulate a sequence capture datasets ( [Bi et al., 2012](#); [Jiang et al., 2019](#); [Abdelkrim et al., 2018b](#) ). Three subsets were generated, with a minimum of 4, 16 and 32 samples/locus for which both supermatrix and supertree approaches were applied, as explained above. These datasets will be referred as follow: Cap-IQ4, Cap-IQ16, Cap-IQ32, Cap-AS4, Cap-AS16 and Cap-AS32.

## 2.7. Mitogenomes and nuclear markers

The *Pinguicula vulgaris* (Turridae) mitogenome (MH308408.1; [Uribe et al., 2018](#) ) was used as a reference to extract partial (up to 20% missing data) to complete mitogenomes (including tRNAs) from the transcriptomes and create the dataset "MT." Several sequences of 28S rRNA, 18S rRNA and histone 3 ( h3) of Turridae from GenBank were used as references to extract the corresponding loci from the 32 transcriptomes by BLAST. Along with the mitochondrial *cox1*, 12S and 16S fragments, they constitute the Sanger multilocus dataset "SAN." Finally, the *cox1* alone constitutes the Sanger barcode dataset "BC." The same protocol as for the reference-based approach was applied for mapping, filtering and alignment. For the MT, SAN and BC datasets, each codon position of the protein coding genes was treated as an independent partition, as well as each non-protein coding gene. The best substitution model was estimated for each partition in each concatenated dataset with ModelFinder following the BIC criterion and 1000 UFBoot were performed on each dataset to obtain branch support for the trees reconstructed with IQ-tree.

## 2.8. Tree topology evaluation

The Turridae trees published so far suffer from both incomplete sampling and lack of resolution (e.g. [Heralde et al., 2007](#); [Puillandre et al., 2012](#) ). Thus, these published trees can hardly be used as a reference tree for the Turridae. Consequently, two approaches were used to evaluate tree topology decisiveness and informativeness.

For the matrix and supermatrix datasets (BC, SAN, MT, Cap-IQ, Ref-IQ, Uph-IQ), the log-likelihood of multiple constrained tree searches for each dataset was compared and the results were statistically tested with IQ-TREE using the [Shimodaira and Hasegawa \(1999\)](#) (SH) test. The trees were constrained respectively following all the different topologies retrieved with the different datasets, except for the intra-specific and outgroup nodes, resulting in a total of eight unique constrained topologies (the same topologies found for Cap-IQ32 and Cap-IQ16, Cap-IQ4 and Ref-IQ32, Ref-IQ16 and Ref-IQ4).

For the supertree datasets (Cap-AS, Ref-AS and Uph-AS), tree metrics were used to evaluate loci quality. The normalized quartet distance of each locus was calculated using TreeCmp ( [Bogdanowicz et al., 2012](#) ) with reference to the corresponding supertree with collapsed intraspecific nodes. Additionally, the quartet distance metric score distribution of BUSCO (single-copy + fragmented) loci trees versus all other single-locus trees for Ref-AS16 and Ref-AS32 were compared to evaluate the quality of the reference-based approach. The quartet score (proportion of quartets satisfying the supertree) was also used to evaluate the overall support of supertree analysis using ASTRAL-III's log.

## 2.9. Data, time and money evaluation

Data – The AMAS python program ( [Borowiec, 2016](#) ) was used to calculate alignment statistics for each dataset, including the number of loci, the alignment length (in the case of ASTRAL-III, the median length of all loci), the total number of matrix cells and undetermined cells (to evaluate missing data) and the proportion of variable and parsimony-informative sites.

Time and money – Comparisons of costs (time and money) were

measured respectively in number of days and euros but did not take into account specimen collection and salary costs, both varying too much depending respectively on the taxon, the country where research is carried out, or the academic level of the person employed (e.g. graduate or engineer). Costs were evaluated by the Service de Systématique Moléculaire (SSM) platform at the MNHN (UMS 2700). The time estimates were based on a realistic best-case scenario, meaning that each step of lab preparation and data analysis are supposed to work on the first try with the methods used at the SSM.

## 3. Results

### 3.1. Transcriptome sequencing, assembly and quality assessment

The total number of raw reads used for transcriptome assembly ranged from 42,770,212 to 138,181,918 and the number of assembled contigs ranged from 46,027 to 283,318. The mean value of N50 is 539. At least 80% of nput reads mapped back to the transcriptome assemblies. The mean BUSCO completeness value is 49.1%, ranging from 36% to 83.7% ( [Supplementary Table 1](#) ). Pearson's r showed a strong correlation between assembly size and BUSCO completeness (  $\rho = 0.78$ ,  $p\text{-value} = 1.54E-07$ ) but no correlation between the number of raw reads and BUSCO completeness (  $\rho = -0.01$ ,  $p\text{-value} = 0.98$ ) ( [Supplementary Table 2](#) ). Transcriptomes produced from foot tissue (*Gemmula* sp. and *M. mitra*) showed a greater BUSCO completeness than transcriptomes produced from venom or salivary glands, suggesting transcript abundance variation among tissues and/or over-representation of some transcripts in glands (e.g. highly expressed toxins – [Duterte et al., 2014](#) ). However, more transcriptomes assembled based on different tissues from the same specimen are needed to properly test this hypothesis.

### 3.2. Phylogenetic results

The monophyly of the ingroup Turridae is always confirmed, except with two datasets (Uph-AS16 and Uph-AS4), where the outgroup *Terebra* is found in the ingroup ( [Supplementary Fig. 1](#) ). The genera *Gemmula* and *Turris* are systematically retrieved polyphyletic ( [Fig. 3](#) ), as shown in previous studies ( [Puillandre et al., 2012](#); [Fedosov et al., 2011](#) ). The species represented by several specimens ( *X. legitima*, *I. cingulifera*, *I. musivum* and *I. olangoensis*) are always recovered as monophyletic groups except for one dataset (Uph-IQ4), in which a specimen of *I. cingulifera* is placed as a sister group of the other members of *Iotyrris*. Apart from the Uph-IQ4 dataset, the relationships inferred among *X. legitima* and all three *Iotyrris* species are always identical. The long branches *Turris* and *Lophiotoma* are found as sister groups only in the "Ref" and "Cap" datasets. Finally, the relationships among *Gemmula* sp., *T. nadaensis*, *Unedogemmula* – the earliest offshoots in the ingroup – and the rest of the Turridae appear to be the most problematic ( [Fig. 3](#) ). The phylogenetic results are globally congruent with previous studies (e.g. [Puillandre et al., 2012](#) ), despite the heterogeneity in the number of species per lineage and several missing lineages. Overall, the graph-based approach (UPhO) shows very low taxon occupancy (see also [Fernández et al., 2018](#) ) and fewer parsimony-informative sites, and hence results in shortest branches and incongruent results with the reference-based approach. An extreme case is the specimen of *I. cingulifera* not retrieved within the *I. cingulifera* species node in the Uph-IQ4 dataset. This specimen's transcriptome shows poorer results in terms of assembly size (38,931,364 bp, compared to the 56,711,565 mean) and BUSCO completeness (23.8% of complete single loci). Nevertheless, the reference-based reconstructions do not suffer from this low-quality transcriptome.

Except for the BC and SAN datasets, support for specific to supra-specific nodes ranged between 75% and 100% ( [Table 1](#) ), and showed no correlation with the dataset size. Interestingly, in the mitogenome dataset (MT), bootstrap supports were similar or superior to those of

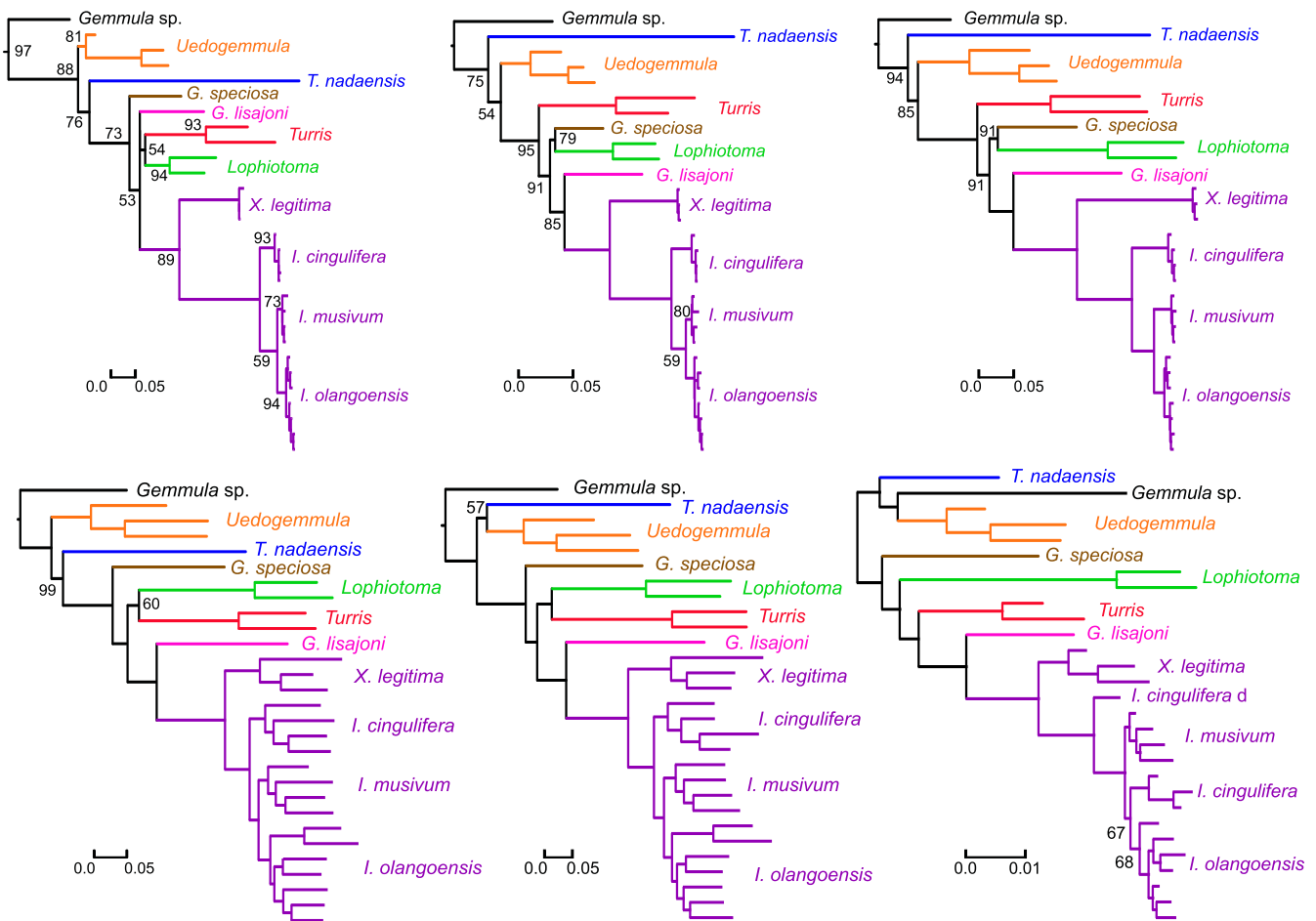


Fig. 3. From top to bottom and left to right: phylogenetic trees corresponding to the BC, SAN, MT, Cap-IQ16, Ref-IQ16 and UPh-IQ4 datasets. Outgroups are not shown. Bootstrap values for the fully supported and intraspecies nodes are not shown. Colors represent genera or genus-level groups. Scale: average frequency of substitution per site.

larger datasets, but those values were negatively affected by the removal of some regions such as tRNAs ( Supplementary Fig. 1 ).

### 3.3. Topology evaluation

Except for UPh-IQ4, all the datasets had at least one alternative constrained topology credible under the SH test ( Table 2 ). The credible sets of trees for the smallest datasets (BC, SAN and MT) contained more constrained trees than the credible sets of trees for the larger datasets (Cap, Ref & UPh).

Not a single-locus tree with 16 or more terminal entities fully matches its corresponding supertree ( Fig. 4 ). This is also true for the UPhO-AS16 single-locus tree distribution ( Supplementary Fig. 2 ). The student's t test results for quartet distance metric score distribution of BUSCO (single-copy + fragmented) loci trees versus all other loci trees for Ref-AS16 showed a significant difference between the two distributions ( $p$ -value <  $2.2e-16$ ; Fig. 4 b). The quartet score decreases when reducing taxon occupancy: for Ref-AS32, Ref-AS16 and Ref-AS4 the normalized quartet scores were respectively 0.730, 0.709 and 0.707 (Supplementary Table 3 ).

### 3.4. Data, time and money

All markers traditionally obtained by Sanger sequencing were extracted from the transcriptomes except for h3, lacking in 25 of the transcriptomes. The largest dataset (DS5aIQ4) is a concatenated alignment of 14,586,607 bp (71.7% of missing data), corresponding to

9232 loci (DS5aAS4), of which all other datasets were constructed, except for the graph-based approach ones. The graph-based approach generated too few loci with no missing data (32 terminals/locus), therefore only four datasets were retained ( Table 1 ). As shown on the Fig. 5 , the reference-based and graph-based approach used respectively 285,660 and 35,595 transcripts for each pipeline, but only 19,008 (6.8%) of the total transcripts are in common between the two pipelines.

Unsurprisingly, the larger datasets are also more costly ( Table 1 ), ranging from an estimated 226€ for the CO1 dataset to 8828€ for transcriptomes, for the production of a 32 terminal entity phylogeny (as in this study). But, while the 'Sanger' datasets (BC, SAN) costs increase proportionally with the number of specimens and number of loci targeted, the mitochondrial and sequence capture datasets costs will dramatically reduce when pooling a lot of specimens. This is particularly the case for the sequence capture dataset, especially when considering the price of custom baits. By pooling 100 post-capture libraries on a single sequencing lane (instead of the 32 in this study), the cost per specimen goes down from 196€ to 81€ (273€ to 105€ if including the transcriptome sequencing and the design of the probes). Finally, the transcriptomes dataset is the only HTS dataset not following the rule of decreasing costs when pooling more specimens, simply because there is a limit on the number of transcriptomes that can be sequenced on a single lane.

**Table 1**  
Description of the datasets analyzed. Me = median locus length. For the robustness evaluation, only the nodes between the ingroup to the species nodes were taken into account. More details on time and money evaluation is available in [Supplementary Table 4](#).

Data		Time (days)					Money (euros)			Robustness						
Dataset type	Orthology assessment	Phylogenetic method	Dataset name	No. of loci	Alignment length	Missing data	Variable sites	Parsimony informative sites	Lab work	Data analysis	Cost per specimen	Cost per base (per specimen)	Cost per variable site (per specimen)	% nodes > 80% BS or 95 PP*	% nodes = 1*	
Sanger - DNA barcoding	gene	ML (IQ-TREE)	BC	1	658	76 (0.4%)	258 (39.2%)	165 (29%)	2	-	7	0.011	0.027	61.1	16.6	
	multilocus	ML (IQ-TREE)	SAN	6	4,787	12,820 (8.4%)	889 (18.6%)	565 (11.8%)	2	-	40	0.01	0.045	77.7	55.5	
	Mitogenome	ML (IQ-TREE)	MT	1	14,927	27,562 (5.8%)	6,491 (43.5%)	4,922 (33%)	5	1	54	0.0036	0.0085	100	77.7	
Sequence capture	Genome reference	ML (IQ-TREE)	Cap-IQ32	274	136,799	249,086 (5.7%)	46,491 (34%)	28,083 (20.5%)	10	6 to 10	196	0.001432759	0.00445	100	83.3	
		Supertree (ASTRAL-III)	Cap-AS2	274	Me = 498	Me = 402.5 (2.6%)	Me = 165.5 (32.6%)	Me = 95.5 (18.6%)						77.7		
		ML (IQ-TREE)	Cap-IQ16	1373	743,778	8,009,019 (33.6%)	266,325 (35.8%)	148,171 (19.9%)				0.000263519	0.00078	100	88.8	
		Supertree (ASTRAL-III)	Cap-AS16	1373	Me = 548	Me = 1,901 (15.2%)	Me = 182 (34%)	Me = 92 (17.5%)						94.4		
		ML (IQ-TREE)	Cap-IQ4	3000	1,623,052	31,758,137 (61.1%)	499,798 (30.8%)	218,629 (13.5%)				0.00012076	0.00041	94.4	94.4	
		Supertree (ASTRAL-III)	Cap-AS4	2999	Me = 555	Me = 1,491 (26.4%)	Me = 154 (29.9%)	Me = 56 (11.8%)						94.4		
	Transcriptomes	Genome reference	ML (IQ-TREE)	Ref-IQ32	473	480,293	2,533,447 (16.5%)	158,798 (33.1%)	91,619 (19.1%)	8	20 to 40	275	0.000572567	0.00071	94.4	94.4
			Supertree (ASTRAL-III)	Ref-AS32	473	Me = 698	Me = 1,046 (4.2%)	Me = 239 (31.8%)	Me = 139 (17.8%)						88.8	
			ML (IQ-TREE)	Ref-IQ16	4663	8,187,363	153,998,814 (58.8%)	2,450,395 (29.9%)	1,147,534 (14%)				0.00003359	0.000046	94.4	94.4
			Supertree (ASTRAL-III)	Ref-AS16	4663	Me = 1,276	Me = 9,438 (34.6%)	Me = 409 (31.2%)	Me = 183 (14.8%)						94.4	
		ML (IQ-TREE)	Ref-IQ4	9232	14,586,607	334,525,406 (71.7%)	3,832,278 (14.3%)	1,465,372 (10%)				0.00001885	0.000029	94.4	94.4	
		Supertree (ASTRAL-III)	Ref-AS4	9232	Me = 1,173	Me = 5,877.5 (42.9%)	Me = 314 (27.7%)	Me = 100 (9.6%)						94.4		
UPhO			ML (IQ-TREE)	Uph-IQ16	347	245,095	2,812,587 (35.9%)	43,022 (17.6%)	20,211 (8.2%)	30 to 50			0.001122014	0.00026	88.8	88.8
			Supertree (ASTRAL-III)	Uph-AS16	345	Me = 618	Me = 0 (0%)	Me = 88 (14.3%)	Me = 41 (6.7%)						NA (84.2)	
			ML (IQ-TREE)	Uph-IQ4	7313	6,681,038	170,796,960 (79.9%)	1,165,551 (17.4%)	368,737 (5.5%)				0.000041161	0.000097	88.8	88.8
			Supertree (ASTRAL-III)	Uph-AS4	7058	Me = 645	Me = 2 (0%)	Me = 82 (11.6%)	Me = 16 (2.2%)						NA (73.7)	



Table 2

Summary table of the Shimodaira-Hasegawa tests for each dataset constrained with each topology, with 1000 resamplings using the RELL method. The topologies are on the top (as column headers) and the datasets on the side (as row headers). "+": the corresponding topology is not rejected; "-": vice-versa.

Dataset / Topology	BC	SAN	MT	Cap-IQ32	Cap-IQ16	Cap-IQ4	Ref-IQ32	Ref-IQ16	Ref-IQ4	Uph-IQ16	Uph-IQ4
BC	-	+	+	+	+	+	+	+	+	+	-
SAN	+	-	+	+	+	+	+	+	+	+	-
MT	+	+	-	+	+	+	+	+	+	+	-
Cap-IQ32	-	-	-	-	+	+	+	+	+	+	-
Cap-IQ16	-	-	-	+	-	+	+	+	+	+	-
Cap-IQ4	-	-	-	+	+	-	+	+	+	+	-
Ref-IQ32	-	-	-	+	+	+	-	+	+	+	-
Ref-IQ16	-	-	-	+	+	+	+	-	+	+	-
Ref-IQ4	-	-	-	+	+	+	+	+	-	+	-
Uph-IQ16	-	+	+	+	+	+	+	+	+	-	-
Uph-IQ4	-	-	-	-	-	-	-	-	-	-	-

4. Discussion

4.1. Comparison of the five sequencing strategies

In the present study, we compare datasets that are representative of the outputs of the pipelines used in most empirical phylogenetic studies in non-model animal taxa, and evaluate them in terms of costs (money and time) and robustness of the resulting tree. It should be noted that the conclusions on the cost evaluations rely on the assumption that the overall costs and timeframes of analyzed methodologies will be similar in other labs. Furthermore, another cost, the environmental cost (the impact of each pipeline on the environment), was not calculated due to the number and complexity of parameters to take in account. However, library preparations and the use of data centers ( Jones, 2018 ) would surely result in a substantial environmental cost for HTS-based trees. If this cost is currently rarely considered, in the future scientists might be encouraged to lower their ecological footprint.

Our results show that traditional Sanger sequencing of one to six loci will retrieve trees with robust nodes for more than half of the clades, quickly and at very affordable costs. Indeed, the *cox1* barcode tree alone retrieved both monophyletic species and most nodes well supported. This particular result might partially explain why, despite 15 years of HTS development and democratization, the vast majority of

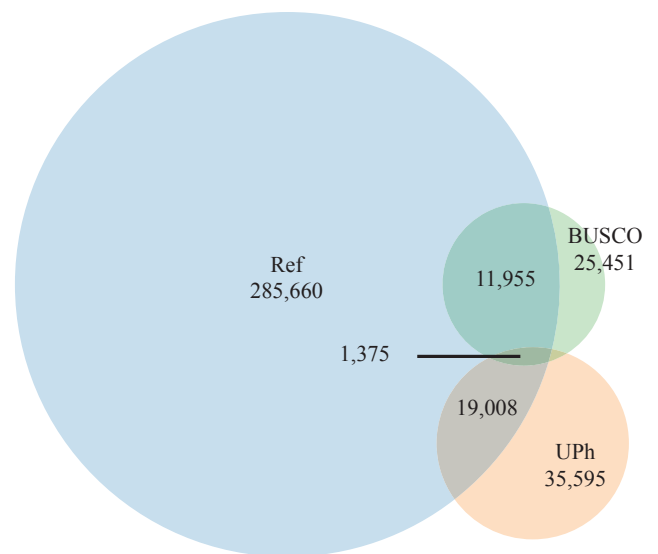


Fig. 5. Number of unique and shared transcripts for the Ref, UPH and BUSCO sets of loci recovered after the first blast step of each pipeline. Total number of transcripts for all transcriptomes is 3,634,333 ( supplementary Table 1 ).

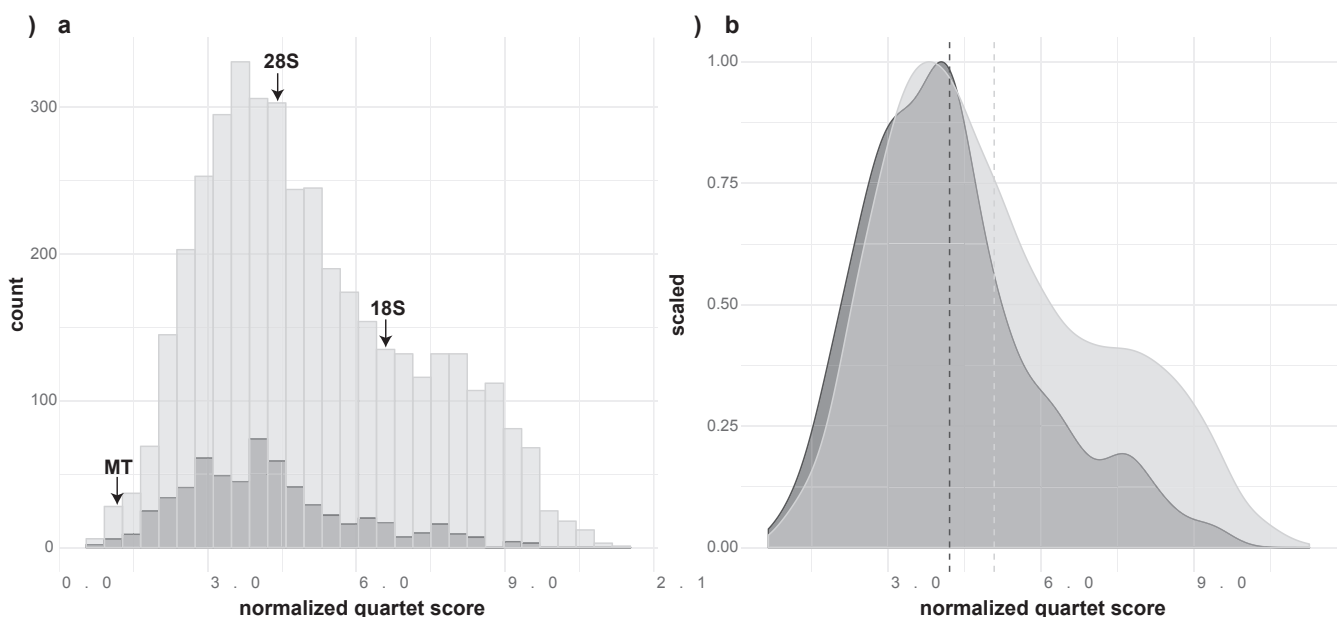


Fig. 4. Distribution of quartet distance of single-locus trees of the Ref-AS16 dataset against the Ref-AS16 supertree, for the BUSCO loci (dark grey) and the other loci (light grey). (a) Total number of counts, with indication of some specific loci (e.g. 28S) distance to supertree (arrows). (b) Scaled density plot, with dotted lines representing the mean values.

articles is still presenting trees produced with such datasets ( Fig. 1 ). Surely, the “Sanger era” has not yet arrived to its end, and many more phylogenies with such datasets will be published in the years to come.

Nevertheless, some nodes remain unsupported, in particular the deeper nodes. We found that the best compromise for retrieving a fully resolved and highly supported tree is the mitogenome dataset, for which all nodes have > 80% bootstrap and the costs are less than half the price of a sequence capture. However, previous studies have already shown that mitogenomic trees are subject to artifacts, such as long-branch attraction generated because of the high rates of mutation of the mitochondrial genome, especially in the third codon positions (Bergsten, 2005, Arabi et al., 2010 ). Moreover, a mitogenome can be considered as a single locus and thus cannot be subjected to congruence tests. The use of nuclear HTS data becomes even more indispensable when investigating biological processes such as introgression (e.g. Eaton et al., 2015; Zhang et al., 2016 ), where analysis of unlinked markers is necessary.

The sequence capture and RNA-seq datasets (based on a reference genome) yielded similar results in terms of phylogenetic reconstruction accuracy, number of credible sets of trees passing the SH test and single-loci tree metrics distribution. However, the costs of sequence capture are by far more affordable than costs of producing and analyzing transcriptomes. Furthermore, RNA-seq requires high-quality, fresh RNA samples, not often available for a representative set of taxa. These considerations led us to the conclusion that sequence capture might be the best method to produce a complete, high resolution tree for a non-model taxon, with a cost per specimen estimated at 80-100€ (if at least 100 specimens are sequenced on one lane) and a processing time of a few weeks to a few months ( Supplementary Table 4 ). Nevertheless, transcriptomic data remains necessary to identify suitable markers that will be targeted by sequence-capture, especially when there is no available genome. Furthermore, transcriptomic data might be more suitable for backbone phylogenetic trees, including very deep relationships (i.e. several hundreds of millions of years; Cunha and Giribet, 2019; Kocot et al., 2011 ). But very deep relationships also imply that it will be harder to distinguish orthology from paralogy. In summary, the Sanger approach still remains well suited to resolve phylogenetic relationships at a low price (both time and money), and can provide a preliminary outline of the taxon diversity, useful to select a subset of taxa that can be analyzed with a more costly approach. However, some gene markers might not be as useful as thought, depending on the taxon (e.g. 18S, see Fig. 4 ), and 12S and 16S will generally only corroborate the *cox1* results. We thus recommend starting with generating a vast DNA barcoding dataset but from there going directly to sequence capture (if there is a strong need to resolve the remaining challenging nodes). Mitogenomes indeed provide the best compromise between tree quality and costs, but are subject to potential biases. Finally, RNA-seq appears only appropriate for constructing phylogenies in the case of very deep relationships or simply to identify suitable markers for sequence capture.

Another class of HTS datasets that has not been explored is the reduced-representation approaches such as RAD-seq (e.g. Baird et al., 2008). RAD-seq has already been established as a suitable tool for phylogenetic inference (e.g., Cariou et al., 2013; Cruaud et al., 2014 ). In a recent *in silico* study (Collins and Hbrek, 2018 ), the authors even found that RAD and sequence capture datasets gave highly congruent results. However, RAD-seq datasets are reduced-representation of genomes, and extracting an *in silico* RAD-seq dataset from our transcriptomes may have produced biased results, not equivalent to other RAD-seq datasets. Nonetheless, it could be argued that sequence capture methods are more promising for phylogenetic studies, because markers are not anonymous, and their sets can be tailored with more versatility according to the needs, samples with fragmented DNA can be sequenced more efficiently, information content per locus is higher (allowing the use of supertree approaches) and larger evolutionary time scales are covered (Harvey et al., 2016 ).

#### 4.2. A note on topology accuracy assessment

As shown in Table 1, the majority of the concatenated datasets show > 80% or even 100% bootstrap values for all nodes – the same applies for ASTRAL support values – even though the amount of data can vary by a factor of 100 between datasets. Despite high node support, several topologies are in conflict, especially for the earliest relationships of the Turridae ( *Gemmula* sp., *Unedogemmula* and *T. na-daensis*). Even if the true tree is unknown, we know that, at best, only one of these topologies is correct. It has already been showed that the bootstrap support value can rapidly saturate when increasing the number of sites (especially invariant ones), proportion of missing data, or both (Simmons and Freudenstein, 2011 ). Furthermore, when using supermatrix approaches, log-likelihood ratio tests have been used to statistically test if a given dataset can accommodate several topologies (e.g. McFadden et al., 2006 ). In our case, all datasets (except the particular case of Uph-AS4) tolerated at least one, but not all, different (constrained) topologies, suggesting that the unconstrained topology is equal to or only slightly better than alternative one(s). The high-bootstrap values and non-conclusive log-likelihood ratio tests for each phylogenomic datasets calls for alternative methods to measure tree robustness.

The normalized quartet score ( Bayzid et al., 2014 ) is the proportion of quartets from the input single-locus trees that agree with the resulting supertree. We used it to measure the performance of datasets with low taxon occupancy (e.g. Ref-AS4) when considering a supertree approach. Our results show that the normalized quartet scores for Cap-AS, Ref-AS and Uph-AS datasets are systematically lower with low taxon occupancy. Such results would imply that, as for supermatrix (Philippe et al., 2017 ), datasets with low taxon occupancy should be avoided (but see e.g. Kallal et al., 2018 ). Graphical representations of single-loci tree distribution, sometimes referred to tree space visualization in its extended version ( Hillis et al., 2005 ) show promising results for understanding inconsistency among the datasets. The distribution of single-loci tree distance to a reference tree ( Fig. 4 a) has already been used to compare the quality of different datasets (Simmons, 2017 ), but also within-dataset informativeness (e.g. intron vs. exon; Chen et al., 2017 ). In the case of non-model taxa, such distribution patterns can be used to compare loci with high reliability of orthologous relationships (e.g. BUSCO single-copies) versus shallow orthologous loci (e.g. from a reference-based or a graph-based approaches) and thus evaluate the quality of a pipeline ( Fig. 4 b). In our case, we show that a simple blast and downstream filtering approach against a reference genome, even a very distant one, gives satisfactory results, although not sufficient to obtain orthologous loci of similar confidence to BUSCO single-copy loci.

#### 4.3. Improving sequence capture: challenges and perspectives

An important challenge of HTS in phylogenetic reconstruction is to a priori identify loci that better reflect evolutionary relationships among taxa. Our reference-based and graph-based approaches implemented herein correspond to the two alternative strategies, widely used to infer orthologous loci from *de novo* assembled transcriptomes (as reviewed in Laumer, 2018 ). In our case, the graph-based approach with UphO yielded poor results in comparison to the reference-based approach, but additional empirical and *in silico* generated datasets need to be analyzed to properly compare them. The UphO approach was especially sensitive to missing data (specimen of *I. cingulifera* not found with other *I. cingulifera* specimens in DS Uph-IQ4) and the tree reconstruction method (*Terebra* found in the ingroup for DS Uph-AS16 and Uph-AS4). One of the reasons that *Terebra* was found in the ingroup for the Uph-AS16 and Uph-AS4 datasets could be that the orthologs found with the graph-based approach were generally poorly informative (~7% parsimony-informative sites on average), thus resulting in poorly resolved single-locus trees. Conversely, the reference-based approach showed satisfying



results, both in terms of pipeline celerity (avoiding “all-vs-all” blast use), tree robustness and congruency between subsamples. Furthermore, it retrieved far more loci than the BUSCO database. However, single-loci tree evaluation ( Fig. 4 ) showed that the loci retrieved with our reference-based approach are not all informative and/or accurate, and the loci selection could be improved. The use of other alignment statistics, such as the proportion of parsimony-informative sites, could allow for a more precise a priori selection of loci (e.g. HaMStR; Ebersberger et al., 2009 ). Nonetheless, in our dataset, and quite paradoxically, there is a slight negative correlation between the number of parsimony-informative sites in the single-locus alignments and the single-locus-tree distances to the supertree (e. g. for Ref-AS32  $\rho = 0.40$ ,  $p\text{-value} = 9.93\text{E}-20$  – Supplementary Table 5 ), suggesting that most of the phylogenetic signal retrieved in the single-locus trees would not be conveyed by the parsimony-informative sites. Approaches to finding true, orthologous, informative loci still needs development, especially when no close reference genome is available. This relies on finding a better combination of filtering thresholds, alignments statistics and tree metrics to reduce the costs and increase the tree robustness, generating a solid framework to test evolutionary hypotheses.

Finally, one particular advantage of the Sanger approach to reconstructing phylogenies is its routine application. A phylogenetic dataset can be completed regularly, by adding additional sequences on a day-to-day basis, with little doubt on the loci sequenced (but see Mutanen et al., 2016 ). This is less true for HTS based approaches, which usually provide a large amount of data requiring significant investment and staff trained in bioinformatics to eventually combine several datasets, produced in several batches and/or by different research teams. To combine the advantage of both approaches, i.e. a small set of well identified loci that can easily be incremented and a larger, more informative dataset, we propose the following strategy: together with the loci that will be identified as targets in the exon capture approach, the mitochondrial and nuclear loci traditionally used in Sanger sequencing (typically, the *cox1*, 16S, 12S, 28S, 18S and *h3* for the mollusks), and even full mitogenomes, could also be captured (e.g. Espeland et al., 2018 with the *cox1* only). Hence, the backbone phylogeny obtained with a sequence capture dataset can further be completed with additional nuclear core markers or mitochondrial genomes, using a multi-level dataset approach.

#### Funding

This work was supported by the CONOTAX project funded by the French Agence Nationale pour la Recherche (ANR-13-JSV7-0013-01). Contribution of AF was supported by the Russian Science Foundation (grant RSF 16-14-10118, PI Y. Kantor).

#### Acknowledgments

Material was collected during several expeditions: the Kavieng Lagoon Biodiversity Survey in Papua New Guinea (June 2014, MoU with the University of Papua New Guinea, PIs: Philippe Bouchet, Jeff Kinch), as part of the Our Planet Reviewed expeditions organized jointly by Muséum national d’Histoire naturelle (MNHN), Pro-Natura International (PNI) and Institut de Recherche pour le Développement (IRD), with support from Papua New Guinea’s National Fisheries Authority, the Total Foundation, the Laboratoire d’Excellence Diversités Biologiques et Culturelles (LabEx BCDiv, ANR-10-LABX-0003-BCDiv), the Programme Investissement d’Avenir (ANR-11-IDEX-0004-02), the Fonds Pacifique, and CNRS’ Institut Ecologie et Environnement (INEE); the KANACONO expedition in New Caledonia (August 2016, convention MNHN-Provence Sud, APA\_NCP5\_2016\_012; PI N. Puillandre and S. Samadi), as part of the Our Planet Reviewed expeditions and the Tropical Deep-Sea Benthos program (expedition.mnhn.fr), with support from the LabEx BCDiv; the Nha-Trang expedition in Vietnam, supported by the Russian–Vietnamese Tropical

Center, with support from the staff of the Tropical Center for assistance in organization of the field sampling and loan of some laboratory equipment; and a collection trip supported by the ‘Conus-Turrid’ project (principal investigator B. M. Olivera, University of Utah, USA). These expeditions were operated under the regulations then in force in the country in question and satisfy the conditions set by the Nagoya Protocol for access to genetic resources. This project was partly supported by the Service de Systématique Moléculaire (UMS 2700 CNRS-MNHN). The authors also thank Laetitia Aznar-Cormano, Juliette Gorson and Mandë Holford for their help in the lab work, Yuri Kantor for his help in Turridae dissections, Barbara Buge for her help in curating the specimens, Mark Phuong, Lou Mary and Jérémie Bardin for their help with bioinformatics scripts.

#### Declaration of Competing Interest

The authors declare no competing interests that could inappropriately influence (bias) their work.

#### Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2019.106660>.

#### References

- Abdelkrim, J., Aznar-Cormano, L., Fedosov, A.E., Kantor, Y.I., Lozouet, P., Phuong, M.A., Zaharias, P., Puillandre, N., 2018a. Exon-capture-based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Mol. Biol. Evol.* 35 (10), 2355–2374 .
- Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., Zaharias, P., Puillandre, N., 2018b. Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Mol. Ecol.* 27 (22), 4591–4611 .
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410 .
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data.
- Arabi, J., Cruaud, C., Couloux, A., Hassanin, A., 2010. Studying sources of incongruence in arthropod molecular phylogenies: sea spiders (Pycnogonida) as a case study. *C.R. Biol.* 333 (5), 438–453 .
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3 (10), e3376 .
- Ballesteros, J.A., Hormiga, G., 2016. A new orthology assessment method for phylogenomic data: unrooted phylogenetic orthology. *Mol. Biol. Evol.* 33 (8), 2117–2134 .
- Bayzid, M.S., Hunt, T., Warnow, T., 2014. Disk covering methods improve phylogenomic analyses. *BMC Genomics* 15 (6), 57 .
- Bergsten, J., 2005. A review of long-branch attraction. *Cladistics* 21 (2), 163–193 .
- Bi, K., Vanderpool, D., Singhal, S., Linderroth, T., Moritz, C., Good, J.M., 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics* 13 (1), 403 .
- Bogdanowicz, D., Giaro, K., Wróbel, B., 2012. TreeCmp: Comparison of trees in polynomial time. *Evol. Bioinform.* 8, EBO-59657 .
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15), 2114–2120 .
- Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4, e1660 .
- Bushnell, B., 2014. BBMap: a fast, accurate, splice-aware aligner (No. LBNL-7065E). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).
- 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 .
- Chen, M.Y., Liang, D., Zhang, P., 2017. Phylogenomic resolution of the phylogeny of Laurasiatherian mammals: Exploring phylogenetic signals within coding and non-coding sequences. *Genome Biol. Evol.* 9 (8), 1998–2012 .
- Collins, R.A., Hrbek, T., 2018. An in silico comparison of protocols for dated phylogenomics. *Syst. Biol.* 67 (4), 633–650 .
- Cunha, T.J., Giribet, G., 2019. A congruent topology for deep gastropod relationships. *Proc. Roy. Soc. B* 286 (1898), 20182776 .
- Cruaud, A., Gautier, M., Galan, M., Foucaud, J., Sauné, L., Genson, G., Dubois, E., Deuve, T., Rasplus, J.Y., 2014. Empirical assessment of RAD sequencing for interspecific phylogeny. *Mol. Biol. Evol.* 31 (5), 1272–1274 .
- Dutertre, S., Jin, A.H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nat. Commun.* 5, 3521 .
- Eaton, D.A., Hipp, A.L., González-Rodríguez, A., Cavender-Bares, J., 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69 (10), 2587–2601 .
- Ebersberger, I., Strauss, S., von Haeseler, A., 2009. HaMStR: profile hidden markov model based search for orthologs in ESTs. *BMC Evol. Biol.* 9 (1), 157 .

- Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? *Evol. Int. J. Organic Evol.* 63 (1), 1–19 .
- Edwards, S.V., 2016. Phylogenomic subsampling: a brief review. *Zool. Scr.* 45, 63–74 .
- Espeland, M., Breinholt, J., Willmott, K.R., Warren, A.D., Vila, R., Toussaint, E.F., Maunsell, S.C., Aduse-Poku, K., Talavera, G., Eastwood, R., Jarzyna, M.A., Guralnick, R., Lohman, D.J., Pierce, N.E., Kawahara, A.Y., 2018. A comprehensive and dated phylogenomic analysis of butterflies. *Curr. Biol.* 28 (5), 770–778 .
- Faircloth, 2013. <http://s3.ultraconserved.org/talks/faircloth-evolution-2013.pdf>
- Fedosov, A., Watkins, M., Heralde III, F.M., Corneli, P.S., Concepcion, G.P., Olivera, B.M., 2011. Phylogeny of the genus *Turris*: Correlating molecular data with radular anatomy and shell morphology. *Mol. Phylogenet. Evol.* 59 (2), 263–270 .
- Fedosov, A., Puillandre, N., Herrmann, M., Kantor, Y., Oliverio, M., Dgebuadze, P., Modica, M.V., Bouchet, P., 2018. The collapse of *Mitra*: molecular systematics and morphology of the Mitridae (Gastropoda: Neogastropoda). *Zool. J. Linn. Soc.* 183 (2), 253–337 .
- Fernández, R., et al., 2018. Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. *Curr. Biol.* .
- Gonzales, D.T.T., Saloma, C.P., 2014. A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. *Toxicon* 92, 66–74 .
- Gorson, J., Ramrattan, G., Verdes, A., Wright, E.M., Kantor, Y., Rajaram Srinivasan, R., Musunuri, R., Packer, D., Albano, G., Holford, M., 2015. Molecular diversity and gene evolution of the venom arsenal of terebridae predatory marine snails. *Genome Biol. Evol.* 7 (6), 1761–1778 .
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Palma, F.D., Birren, W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652 .
- Haddad, et al., 2018. *Syst. Entomol.* <https://doi.org/10.1111/syen.12257> .
- Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., Brumfield, R.T., 2016. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Syst. Biol.* 65 (5), 910–924 .
- Heather, J.M., Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics* 107 (1), 1–8 .
- Heralde III, F.M., Watkins, M., Ownby, J.P., Bandyopadhyay, P.K., Santos, A.D., Concepcion, G.P., Olivera, B.M., 2007. Molecular phylogeny of some Indo-Pacific genera in the subfamily Turridae H. Adams and A. Adams, 1853 (1838)(Gastropoda: Neogastropoda). *Nautilus* 121 (3), 131–138 .
- Heralde, F.M., Kantor, Y., Astilla, M.A.Q., et al., 2010. The Indo-Pacific *Gemmula* species in the subfamily Turridae: aspects of field distribution, molecular phylogeny, radular anatomy and feeding ecology. *Philippine Sci. Lett.* 3, 21–34 .
- Hillis, D.M., Heath, T.A., John, K.S., 2005. Analysis and visualization of tree space. *Syst. Biol.* 54 (3), 471–482 .
- Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2017. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35 (2), 518–522 .
- Huang, X., Madan, A., 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9 (9), 868–877 .
- Jiang, J., Yuan, H., Zheng, X., Wang, Q., Kuang, T., Li, J., Liu, J., Song, S., Wang, W., Cheng, F., Li, H., Huang, J., Li, C., 2019. Gene markers for exon capture and phylogenomics in ray-finned fishes. *Ecol. Evol.* 9 (7), 3973–3983 .
- Johnson, S.B., Warén, A., Lee, R.W., Kano, Y., Kaim, A., Davis, A., Strong, E.E., Vrijenhoek, R.C., 2010. *Rubyspira*, new genus and two new species of bone-eating deep-sea snails with ancient habits. *Biol. Bull.* 219 (2), 166–177 .
- Jones, N., 2018. How to stop data centres from gobbling up the world's electricity. *Nature* 561 (7722), 163 .
- Kallal, R.J., et al., 2018. A phylotranscriptomic backbone of the orb-weaving spider family Araneidae (Arachnida, Araneae) supported by multiple methodological approaches. *Mol. Phylogenet. Evol.* .
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., von Haeseler, A., Jermini, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14 (6), 587 .
- 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, K., et al., 2011. Phylogenomics reveals deep molluscan relationships. *Nature Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods* 9 (4), 357 .
- Laumer, C.E., 2018. Inferring ancient relationships with genomic data: a commentary on current practices. *Integr. Comp. Biol.* 58 (4), 623–639 .
- Leaché, A.D., Chavez, A.S., Jones, L.N., Grummer, J.A., Gottscho, A.D., Linkem, C.W., 2015a. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biol. Evol.* 7 (3), 706–719 .
- Leaché, A.D., Banbury, B.L., Felsenstein, J., De Oca, A.N.M., Stamatakis, A., 2015b. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst. Biol.* 64 (6), 1032–1047 .
- Leaché, A.D., Oaks, J.R., 2017. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annu. Rev. Ecol. Syst.* 48, 69–84 .
- Lee, K.M., Kivelä, S.M., Ivanov, V., Hausmann, A., Kaila, L., Wahlberg, N., Mutanen, M., 2018. Information dropout patterns in restriction site associated DNA phylogenomics and a comparison with multilocus sanger data in a species-rich moth genus. *Syst. Biol.* 67 (6), 925–939 .
- Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61 (5), 727–744 .
- Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. *Annu. Rev. Ecol. Syst.* 44, 99–121 .
- Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22 (13), 1658–1659 .
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/mapformat and SAMtools. *Bioinformatics* 25, 2078–2079 .
- Liu, C., Zhang, Y., Ren, Y., Wang, H., Li, S., Jiang, F., Yin, L., Qiao, X., Zhang, G., Qian, W., Liu, B., Fan, W., 2018. The genome of the golden apple snail *Pomacea canaliculata* provides insight into stress tolerance and invasive adaptation. *GigaScience* 7 (9), giv101 .
- Magoč, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* .
- Mardis, E.R., 2008. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* 9, 387–402 .
- McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Res.* 22 (4), 746–754 .
- McFadden, C.S., France, S.C., Sánchez, J.A., Alderslade, P., 2006. A molecular phylogenetic analysis of the Octocorallia (Cnidaria: Anthozoa) based on mitochondrial protein-coding sequences. *Mol. Phylogenet. Evol.* 41 (3), 513–527 .
- McKain, M.R., Johnson, M.G., Uribe-Convers, S., Eaton, D., Yang, Y., 2018. Practical considerations for plant phylogenomics. *Appl. Plant Sci.* 6 (3), e1038 .
- Moreau, C.S., Wray, B.D., 2017. An empirical test of reduced-representation genomics to infer species-level phylogenies for two ant groups. *Insect Systemat. Diversity* 1 (2) .
- Mutanen, M., Kivelä, S.M., Vos, R.A., Doorenaere, C., Ratnasingham, S., Hausmann, A., Huemer, P., Dincă, V., Van Nieuwerkerken, E.J., Lopez-Vaamonde, C., Vila, R., Aarvik, L., Decaens, T., Efetov, K.A., Hebert, P.D.N., Johnsen, A., Karsholt, O., Penttinen, M., Rougerie, R., Segerer, A., Tarmann, G., Zahiri, R., Godfrey, H.C.J., 2016. Species-level para- and polyphyly in DNA barcode gene trees: Strong operational bias in European lepidoptera. *Syst. Biol.* 65, 1024–1040 .
- Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32 (1), 268–274 .
- Olivera, B.M., Hillyard, D.R., Watkins, M., 2008. A new species of *Gemmula*, Weinkauff 1875; Evidence of two clades of Philippine species in the genus *Gemmula*. *Philipp Sci. Lett.* 11, 11–15 .
- Philippe, H., Vienne, D.M.D., Ranwez, V., Roure, B., Baurain, D., Delsuc, F., 2017. Pitfalls in supermatrix phylogenomics. *Eur. J. Taxonomy* 283, 1–25 .
- Puong, M.A., Mahardika, G.N., 2018. Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. *Mol. Biol. Evol.* 35 (5), 1210–1224 .
- Puong, M.A., Alfaro, M.E., Mahardika, G.N., Marwoto, R.M., Prabowo, R.E., von Rintelen, T., Vogt, P.W.H., Hendricks, J.R., Puillandre, N., 2019. Lack of signal for the impact of conotoxin gene diversity on speciation rates in cone snails. *Syst. Biol.* .
- Puillandre, N., Modica, M.V., Zhang, Y., Sirovich, L., Boisselier, M.C., Craud, C., Holford, M., Samadi, S., 2012. Large-scale species delimitation method for hyperdiverse groups. *Mol. Ecol.* 21 (11), 2671–2691 .
- Puillandre, N., Holford, M., 2010. The Terebridae and terebratoxins: Combining phylogeny and anatomy for concerted discovery of bioactive compounds. *BMC Chem. Biol.* 10 (1), 7 .
- Puillandre, N., Fedosov, A.E., Zaharias, P., Aznar-Cormano, L., Kantor, Y.I., 2017. A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess. *Zool. J. Linn. Soc.* 181 (2), 243–271 .
- Ruane, S., Raxworthy, C.J., Lemmon, A.R., Lemmon, E.M., Burbrink, F.T., 2015. Comparing species tree estimation with large anchored phylogenomic and small Sanger-sequenced molecular datasets: an empirical study on Malagasy pseudoxiphophine snakes. *BMC Evol. Biol.* 15 (1), 221 .
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16 (8), 1114 .
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31 (19), 3210–3212 .
- Simmons, M.P., Freudenstein, J.V., 2011. Spurious 99% bootstrap and jackknife support for unsupported clades. *Mol. Phylogenet. Evol.* 61 (1), 177–191 .
- Simmons, M.P., 2017. Relative benefits of amino-acid, codon, degeneracy, DNA, and purine-pyrimidine character coding for phylogenetic analyses of exons. *J. Systemat. Evol.* 55 (2), 85–109 .
- Todd, J.A., Rawlings, T.A., 2014. A review of the Polystira clade—the Neotropic's largest marine gastropod radiation (Neogastropoda: Conoidea: Turridae sensu stricto). *Zootaxa* 3884 (5), 445–491 .
- Uribe, J.E., Zardoya, R., Puillandre, N., 2018. Phylogenetic relationships of the conoidean snails (Gastropoda: Caenogastropoda) based on mitochondrial genomes. *Mol. Phylogenet. Evol.* 127, 898–906 .
- Washburn, J.D., Schnable, J.C., Conant, G.C., Brutnell, T.P., Shao, Y., Zhang, Y., Ludwig, M., Davide, G., Pires, J.C., 2017. Genome-guided phylo-transcriptomic methods and the nuclear phylogenetic tree of the paniceae grasses. *Sci. Rep.* 7 (1), 13528 .
- Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C., Jörger, K.M., Schrödl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc. Roy. Soc. B: Biol. Sci.* 281 (1794), 20141739 .
- Zhang, W., Dasmahapatra, K.K., Mallet, J., Moreira, G.R., Kronforst, M.R., 2016. Genome-wide introgression among distantly related Heliconius butterfly species. *Genome Biol.* 17 (1), 25 .
- Zhang, C., Sanyari, E., Mirarab, S., 2017. ASTRAL-III: Increased scalability and impacts of contracting low support branches. In: RECOMB International Workshop on Comparative Genomics. Springer, Cham, pp. 53–75 .

## 2. A la recherche de nouveaux marqueurs de type « barcode »

Ce travail a été réalisé en collaboration avec une étudiante en post-doc dans notre équipe, Giulia Fassio, avec laquelle nous avons travaillé pendant 4 mois (Janvier-Avril 2018) sur l'identification de marqueurs de type « barcode » universels qui soient séquençables en approche Sanger. Je me suis occupé de la partie bioinformatique tandis qu'elle s'est occupée du dessin de sondes et de la partie expérimentale en laboratoire.

Au préalable, nous disposions de 10 couples de primers dessinés à l'aide de 19 transcriptomes du complexe *Iotyrris/Xenuroturris* (Tbl. 1) par un collègue, Mark Phuong, qui avait suivi l'approche suivante. L'orthologie des loci a été évaluée en suivant le pipeline « reference-based » détaillé dans l'article de la première partie de ce chapitre, mais en utilisant comme génome de référence *Lottia gigantea* (Simakov *et al.* 2013) et en annotant les exons en utilisant les limites d'exons du génome de *L. gigantea* sur la base de données EnsemblMetazoa. Les exons ont ensuite été filtrés pour qu'il y ait 10 à 16 spécimens par alignement, et pour que la distance moyenne pour chaque paire d'individus de soit supérieure

Tableau 6 : Les 10 couples d'amorces dessinés par Mark Phuong et testés sur les Conoidea

N° primer	Id protéine <i>Lottia</i>	N° exon	Séquence	% GC	Tm	longueur du primer	longueur du locus
primer1	222603	exon4	GATTGGGAGAACAAGCAGGA	50	60,2	20	351
primer1	222603	exon4	GTGGGACTCATCCTTCTCCA	55	60,05	20	351
primer2	183748	exon3	CTGTGCTGTGGTCCAGAAGA	55	60,02	20	313
primer2	183748	exon3	CTTCTAACGTCGCTGCCATT	50	60,41	20	313
primer3	196232	exon0	AGTGGCGCATCTTCTTCTGT	50	60,02	20	330
primer3	196232	exon0	GAGAGTCGATGATGGGTGCT	55	60,23	20	330
primer4	213398	exon3	CCTTTGGGATCCGTGGTG	61,1	62,33	18	321
primer4	213398	exon3	AGTCTTCGTCGGTTTGATGG	50	60,11	20	321
primer5	224000	exon1	ACCAGTGTGGGTTTTGGCTA	50	60,41	20	350
primer5	224000	exon1	TTGAGCTTGGGCAAGTAAAGA	42,9	60	21	350
primer6	233882	exon2	CTGTTTCGCCACCTTGAAAT	45	60,11	20	384
primer6	233882	exon2	GGGTTCAGCTGACACAGTT	55	60,16	20	384
primer7	239271	exon5	GAAATCAAGAACGGCTACGC	50	59,85	20	246
primer7	239271	exon5	CACGGTCTGCTTCATGTCAC	55	60,32	20	246
primer8	119163	exon0	CAAGAAGGGGCATCTGGAC	57,9	60,61	19	255
primer8	119163	exon0	AAGTCGTATTTCCCCCGAAC	50	60,19	20	255
primer9	198678	exon1	ACCGTCTCTTCGCTACCTC	60	60,79	20	443
primer9	198678	exon1	AGCCAGCAATCGTTTCAGTT	45	59,88	20	443
primer10	225029	exon7	GCTGTCAGCCTTGCCATAG	57,9	59,56	19	324
primer10	225029	exon7	AGAAGCGCCATGGAGGAG	61,1	61,48	18	324

à 1%. Les séquences ont également été filtrées pour être de longueur supérieure à 350 paires de bases (pb). Les primers ont ensuite été dessinés en utilisant Primer 3 (Untergasser *et al.*, 2012).

Seuls cinq couples de primers (P1, P2, P6, P7 et P8) ont passé l'étape de PCR et de séquençage avec succès. Parmi ces cinq couples, seul P7 (locus 239271\_exon5) a permis de retrouver toutes les hypothèses d'espèces CO1 du complexe *Iotyrris/Xenuroturris*, les autres n'étant pas assez variables. Mais lorsque nous avons utilisé le couple de primers P7 sur d'autres genres de Turridae (*Lophiotoma*, *Unedogemmula*, *Gemmuloborsonia*) et chez un genre de Conidae (*Profundiconus*), le locus 239271\_exon5 ne s'est pas révélé assez variable pour retrouver les hypothèses d'espèce CO1 déjà établies dans ces genres. De plus nous avons constaté chez *Lophiotoma* la présence de double pics dans les chromatogrammes des séquences, indiquant une possible hétérozygotie pour ce locus. Enfin, de manière générale, les exons retrouvés à partir des différents couples d'amorce fournissent un signal phylogénétique très faible. Dans l'idéal, un marqueur servant à la délimitation d'espèces serait également utilisable pour reconstruire des phylogénies, comme c'est déjà le cas de la plupart des marqueurs Sanger servant à la délimitation d'espèces. Une approche différente d'identification de marqueurs a donc été mise en place, combinant analyse phylogénétique et évaluation des distances intra-/inter-spécifiques.

L'approche que j'ai développée est similaire en tous points à celle de Mark Phuong en ce qui concerne la délimitation et l'annotation des exons, mais diffère dans la sélection des exons à tester expérimentalement. Cette fois, 58 transcriptomes de Neogastropoda ont été utilisés, pour la plupart non-publiés. Pour certains genres comme *Iotyrris/Xenuroturris*, *Turricula* et *Profundiconus* nous disposons de plusieurs transcriptomes par espèce. L'approche suit le déroulement suivant :

- (i) Concaténation de tous les exons annotés (voir détails dans Article 2) et construction d'un arbre phylogénétique (maximum de vraisemblance, mêmes paramètres que pour les arbres Ref-IQ ou Cap-IQ de l'article 2) de référence
- (ii) Sélection d'exons de taille supérieure à 300 pb, de sorte à maximiser le nombre de sites informatifs par locus. L'idéal serait un marqueur entre 500 et 700 pb (comme le barcode CO1), relativement facile à séquencer en approche Sanger.
- (iii) Construction d'arbres d'exons (Maximum de Vraisemblance, mêmes paramètres que pour les arbres Ref-AS ou Cap-AS de l'article 2)

(iv) Utilisation d'une métrique de distance entre arbres (la distance Robinson-Foulds normalisée ; Robinson & Foulds, 1981) pour sélectionner les 100 arbres d'exons les plus proches de l'arbre concaténé de référence, et extraction des alignements correspondants.

(v) Retour aux alignements afin de vérifier qu'il n'y ait pas d'erreurs (ex : contaminations). Evaluation à l'œil et/ou utilisation de méthodes automatisées de délimitation d'espèces (ABDG, PTP) afin de sélectionner un lot d'exons les plus adéquats pour des tests de séquençage en laboratoire. Ainsi, les métriques d'arbres permettent de s'assurer de la qualité du signal phylogénétique de l'exon, tandis que les méthodes automatisées de délimitation d'espèce permettent de s'assurer de l'efficacité de l'exon pour délimiter les espèces.

Au total, 14 exons d'intérêt ont été identifiés, pour lesquels 37 primers ont été dessinés (Tbl. 2). En outre, nous voulions tester le séquençage à l'aveugle de quelques introns, car les régions non-codantes du génome sont généralement plus variables que les régions codantes. L'approche « exon-primed, intron crossing » (EPIC ; Palumbi & Baker, 1994) a été utilisée, permettant de capturer des introns à partir de primers dessinés sur les exons flanquants. Un total de 18 primers permettant de capturer cinq introns ont été dessinés. Enfin, nous avons également testé l'intron 9 du gène gamma-glutamyl carboxylase (*ggcx*) déjà identifié et testé avec succès chez des Conidae (Kraus *et al.*, 2011 ; Kraus *et al.*, 2012). Pour cela, huit primers ont été dessinés.

La grande majorité des loci testés n'ont pas marché en PCR, soit parce que les primers n'ont pas fonctionné, soit car les PCRs ont produits des bandes multiples. Le seul locus ayant fonctionné est le locus 229923, composé de deux exons et d'un intron. Deux couples de primers, fonctionnant a priori sur tous les Conoidea, permettent de séquencer deux fragments de ce locus, respectivement l'exon 1 et l'intron 1-2 + une partie de l'exon 2. La combinaison de ces deux fragments en une séquence pour évaluer l'efficacité de ce gène pour délimiter les espèces n'a pas été véritablement testée, pour des raisons de temps et d'argent. Cependant, les quelques résultats obtenus sur l'exon 1 nous indiquent que le locus 229923 ne serait pas suffisamment variable pour pouvoir être utilisé comme barcode. De plus, des doubles pics dans les chromatogrammes ont été relevés chez certains spécimens de *Gemmuloborsonia*. En revanche, les séquences de *Gemmuloborsonia* et d'*Unedogemmula* indiquent une cinquantaine de nucléotides de différence (~9% de divergence). Ce locus pourrait donc s'avérer être un bon marqueur pour reconstruire des phylogénies, même si cette proposition reste à tester sur différents groupes de Neogastropoda.

Tableau 2 : Tableau récapitulatif des amorces testées sur les Conoidea avec Giulia Fassio.

Identifiant primer	N° protéine <i>Lottia</i>	Locus	groupe visé	primer (5'->3')	% GC	Tm	longueur du primer	longueur du locus
ExonX_F	224100	Exon 1	Conoidea	CATCACCCCTCAACACCATCAC	52,38	58,91	21	815
ExonX_R	224100	Exon 1	Turridae	GAAGAGGTCTGAACAATGCC	52,38	58,91	21	815
Exon1_F	229923	Exon 1	Turridae	CAGGGCATGGGTGAGGAATC	60	60,47	20	508
Exon1_R	229923	Exon 1	Turridae	GGGGCCTTCATCTTCTTGGA	55	59,37	20	508
Exon2_F	229923	Exon 2	Conoidea	GGCAACCGTTTCCGCTAC	61,11	59,13	18	254
Exon2_longR	229923	Exon 2	Conoidea	GGCCTGCCCTGGAAGTAG	66,67	59,41	18	254
Exon2_shortR	229923	Exon 2	Turridae	CCTGGAAGTTGGTCATGGTG	55	58,46	20	200
Exon4_longF	239271	Exon 4	Clade A	ATCGTGGCTGCTGGTACTGG	60	62,24	20	388
Exon4_shortF	239271	Exon 4	Conoidea	GTCAACAAGATGGACAGCAC	50	57,02	20	277
Exon4_R	239271	Exon 4	Conoidea	CTCCGATCTTGTAGACATCCTG	50	57,76	22	277
Exon7_F	23927	Exon 7	Clade A	GTCATCATCCTGAACCACCC	55	57,96	20	313
Exon7_R	23927	Exon 7	Turridae	CTTCTGAGCGGCCTTGGT	61,11	59,65	18	313
224100_EX_1f	224100	Exon 1	Conoidea	CCCATCACCCCTCAACACCAT	55	59,7	20	744
224100_EX_2r	224100	Exon 1	Turridae	TCCGGAGATTCCCACAGCTA	55	60	20	
224100_EX_d1r	224100	Exon 1	Conoidea	GCCRGTTCTGYCCCACTTGCAT	57,1	62,9	21	699
229923_E1_1f	229923	Exon 1	Conoidea	CCCCACCAACCTCATCATC	60	60,1	20	553
229923_E1_2f	229923	Exon 1	Conoidea	CCCACCAACCTCATCATCAACTA	47,8	60,1	23	551
229923_E1_d1r	229923	Exon 1	Conoidea	GGGGADGGVACRTACTGCAC	60	60,7	20	
183748_E1_1f_I	183748	Exon 1	Conoidea	TGCCAAGATGGACAAGAGCA	50	59,6	20	633
183748_E1_1r_I	183748	Exon 1	Conoidea	ACCAGGGAGTGTGCTAGACA	55	60,2	20	
183748_E1_2f_II	183748	Exon 1	Conoidea	ACAACGAGTGGGTGGTGAAG	55	60,2	20	
183748_E1_2r_II	183748	Exon 1	Conoidea	GGGTGACAGGGTCAAAGTT	55	59,9	20	841
183748_E1_3f_II	183748	Exon 1	Turridae	CCGGATGGGCTTCATACACA	55	59,8	20	746
183748_E1_3r_II	183748	Exon 1	Turridae	AGAAAGCATGCTCTGGGTGT	50	59,6	20	
209765_E1_1f	209765	Exon 1	Turridae	AACAGGCAGTTGTTACCCCG	55	60,5	20	
209765_E1_1r	209765	Exon 1	Turridae	ATGACGTCAGACAACGAGCG	55	60,7	20	480
209713_E3_1f	209713	Exon 3	Turridae	CTCCCAGTTCTCCAAGCTGG	60	60	20	
209713_E3_1r	209713	Exon 3	Turridae	GACTCCATGGGCGTGATCT	57,9	59,2	19	269
209713_E3_2r	209713	Exon 3	Turridae	GACACCACCTCTCCTTCAC	60	59,7	20	218
239532_E1_2f_II	239532	Exon 1	Turridae	AGCAAGAAAGGTCTGCGTCT	50	59,6	20	
239532_E1_2r_II	239532	Exon 1	Turridae	GTTCATGGGCTGTTCCCTCCT	55	59,7	20	816



239532_E1_1f_I	239532	Exon 1	Turridae	AAGACCCTCAAGCCTTCAGC	55	60	20	
239532_E1_1r_I	239532	Exon 1	Turridae	AGACGCAGACCTTTCTTGCT	50	59,6	20	786
239532_E3_1f	239532	Exon 3	Turridae	GGGCAGGGAAAGGAGGAGAT	60	61	20	
239532_E3_1r	239532	Exon 3	Turridae	CCTTGCTTTTGGCACAGGA	52,6	58,6	19	511
239532_E3_2f	239532	Exon 3	Turridae	GAGCTCCTGCCTTCCCAAAA	55	60,3	20	
239532_E3_2r	239532	Exon 3	Turridae	TGCACAGCACCTCAATCTCA	50	59,6	20	684
229923_E1_e1f	229923	Exon 1-2	Conoidea	GCCAAGATCATCCGGGACAA	55	60,1	20	
229923_E1_e2f	229923	Exon 1-2	Turridae	CAAACGCATCAAAGTGGCCA	50	60	20	
229923_E2_e1sr	229923	Exon 1-2	Conoidea	GTTGCCTTGGTTACGCATGG	55	60,1	20	543
229923_E2_e1lr	229923	Exon 1-2	Conoidea	CATTCTGGGCCTCCTGGAAG	60	60,1	20	787
229923_E2_e2r	229923	Exon 1-2	Conoidea	CTGCCCTGGAAGTAGAAGCC	60	60,1	20	705
209713_E2_e1f	209713	Exon 2-3	Conoidea	ACAAGAAGCCGTTGATGGAGA	47,6	59,7	21	
209713_E2_e2f	209713	Exon 2-3	Conoidea	GCTGCAGGCCATCAAGAAAG	55	59,8	20	
209713_E3_e1r	209713	Exon 2-3	Conoidea	GTACTTGCTGACCACGTGGG	60	60,9	20	415
209713_E3_e2r	209713	Exon 2-3	Conoidea	CTTGCGCACTCGTTGAAG	57,9	59,8	19	369
239532_E1_e1f	239532	Exon 1-2	Conoidea	CTCAAACAGGTGGCCAGTCA	55	60,2	20	
239532_E2_e1r	239532	Exon 1-2	Clade A	TTGGTCTCCACCTCTTTGC	52,6	57,3	20	121+33
239532_E1_e2f	239532	Exon 1-2	Conoidea	AGGAGGAACAGCCCATGAAC	55	59,7	20	182+33
239532_E2_e3f	239532	Exon 2-3	Conoidea	CCCCATGGCAGCAAATGTA	50	58,8	20	
239532_E3_e3r	239532	Exon 2-3	Turridae	TTTTGGGAAGGCAGGAGCTC	55	60,3	20	477+78
239532_E2_e5f	239532	Exon 2-3	Conoidea	TACCCCCAGTACTGTCAACA	50	57,6	20	
239532_E3_e5r	239532	Exon 2-3	Conoidea	TGGCCAGAGACAGGTTGTTG	55	60,2	20	208+78
239532_E3_e4f	239532	Exon 3-4	Turridae	TGAGATTGAGGTGCTGTGCA	50	59,6	20	
239532_E4_e4r	239532	Exon 3-4	Conoidea	CCACAGGTGATGTTCCGCAT	55	60,7	20	345+90
E9_1f	182855	Exon 9-10	Turridae	CAACAGCTGGTCASAGGGTC		55,9	20	
435+deg	182855	Exon 9-10	Turridae	GGMTAYTCSTGGGAYATGATG		56,3	21	
E9_2f	182855	Exon 9-10	Turridae	CACYCAGCACACMAGGATCT		55,9	20	
458+deg	182855	Exon 9-10	Turridae	GACACTGGAGARCAAGGGTA		53,8	20	
E10_3R	182855	Exon 9-10	Turridae	TCGWYGTGTRYGGTCCATGC		60,2	20	
E10_4R	182855	Exon 9-10	Turridae	GYAGACGATCAGCAATGCAG		53,8	20	
E10_5R	182855	Exon 9-10	Turridae	TCCACATTGTGAATYCCRTGC		54,4	20	
E10_6R	182855	Exon 9-10	Turridae	CTGGAARCGWTBRTTGAGGGAT		56,7	21	



Pour l'intron 9 du gène *ggcx* (protéine 182855 chez *Lottia*), le dessin de nouvelles sondes a permis de séquencer ce locus chez des spécimens de *Gemmulobrsonia*. Toutefois il s'est avéré que dans la plupart des chromatogrammes des doubles pics étaient présents, pouvant indiquer une nouvelle fois de l'hétérozygotie ou bien de l'existence de paralogues.

La recherche, mais surtout le test expérimental de marqueurs nouveaux est une entreprise dispendieuse, chronophage et souvent peu productive, et s'est donc arrêtée à ce niveau. Un des principaux obstacles est à chercher à la source des approches présentées ; l'approche « reference-based », sur un génome éloigné de plusieurs centaines de millions d'années, ne permet que de trouver des loci orthologues très conservés à l'échelle des gastéropodes. En travaillant avec ce type de loci, les chances de trouver un marqueur suffisamment variable sont donc faibles. En revanche, cette approche peut permettre d'identifier des séquences informatives phylogénétiquement et potentiellement conservées à l'échelle des gastéropodes, comme c'est le cas du locus 229923.

En attendant un génome assemblé et annoté de Neogastropoda, une approche alternative pourrait se trouver dans la combinaison des approches PCR et des technologies de séquençage HTS. Des projets émergents, comme le projet CODEX mené en ce moment à l'UMS 2700 du MNHN, proposent de séquencer à moindre coût des séquences courtes, en utilisant une approche de double-tagging des échantillons, et en utilisant la technologie de séquençage Illumina. Les loci d'intérêt sont amplifiés en fragments suffisamment courts pour pouvoir être séquencés en Illumina (ex : 2 fragments pour le CO1). Afin de pouvoir séquencer un très grand nombre de produits, chaque fragment est indexé avec deux tags. Ainsi 100 tags fournissent 10000 combinaisons possibles, permettant de séquencer 10000 produits de PCR en même temps. Au-delà de l'aspect financièrement intéressant, cette méthodologie propose d'autres avantages. Nous l'avons vu plus haut, les problèmes de bandes multiples ou de doubles-pics sur les chromatogrammes, dus à des locus hétérozygotes ou à la présence de paralogues, ne peuvent être résolus par l'approche de séquençage Sanger. En revanche, en utilisant les reads de séquençage des techniques HTS, toutes les copies de locus (bandes multiples, hétérozygotie) peuvent être indépendamment assemblées.

L'avenir du barcode sera donc avec les technologies de séquençage haut-débit, en attendant que le prix du séquençage de génome devienne assez intéressant pour se passer de PCR.

### **3. Conclusion**

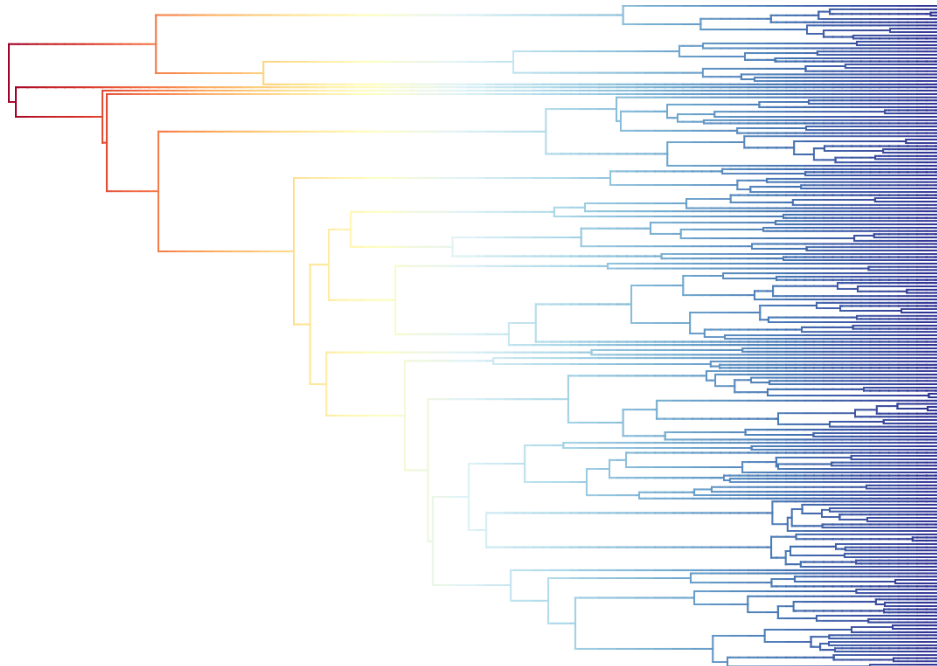
L'utilisation expérimentale de transcriptomes pour la Systématique présente des intérêts variés en délimitation d'espèces et en phylogénie. Elle permet à la fois d'évaluer des pipelines bioinformatiques pré-existants, d'en créer de nouveaux, de s'assurer de la validité de nos stratégies d'échantillonnage de loci, et de rechercher des marqueurs alternatifs aux quelques rares existant. L'augmentation quasi-exponentielle du nombre de transcriptomes produits et la facilité à les manipuler devrait encourager les biologistes à produire plus d'études expérimentales permettant de mieux comprendre les résultats empiriques de certaines méthodes, comme les alignements multiples de séquences, les reconstructions phylogénétiques, les mesures de soutiens, le traitement du polymorphisme nucléotidique, etc.

Plus important, en tout cas dans le cadre de cette thèse, les transcriptomes doivent servir de terrain expérimental pour mieux encadrer les approches dites « génome réduit », utiles à la fois en délimitation d'espèce et en phylogénie. La recommandation proposée à la fin de l'article de la partie 1 de ce Chapitre – que les génomes d'organites et les marqueurs classiques, notamment ribosomiques, devraient systématiquement être inclus dans les approches de type capture – n'est que très peu suivie dans la littérature, y compris dans cette thèse (voir Chapitre 3). Une des raisons est peut-être la surreprésentation des ADN ribosomiques, mitochondriaux ou chloroplastiques dans les extraits d'ADN totaux, pouvant alors affecter négativement le taux de capture de l'ADN génomique. Cela dit, une solution simple consisterait à diminuer le nombre de sondes permettant de capturer ces loci par rapport au nombre de sondes total (ex : Espaland *et al.*, 2018). La construction de jeux « génomes réduits » efficaces, c'est à dire maximisant le nombre de loci informatifs et minimisant les coûts, passe aujourd'hui par l'exploration et l'examen des transcriptomes.

A l'issue de ces deux premiers chapitres, nous disposons maintenant d'une partition d'espèces nous permettant de sélectionner les spécimens qui seront inclus dans la phylogénie, et nous avons validé la méthodologie de « sequence capture » pour reconstruire cette phylogénie.



**CHAPITRE 3**  
**PHYLOGENIE ET DIVERSIFICATION DES**  
**TURRIDAE**



Ce chapitre constitue à la fois le cœur et l'aboutissement de cette thèse. Son objectif est triple ; (i) construire une phylogénie robuste et fiable des Turridae afin de pouvoir (ii) établir une classification stable du groupe et (iii) tester des hypothèses relatives à la diversification des Turridae.

Les deux chapitres précédents ont permis de délimiter les unités terminales pouvant être utilisées pour la phylogénie et de valider la méthodologie utilisée pour construire cette phylogénie. L'historique de la phylogénie/classification des Turridae a été décrit en détail dans les parties 1 et 2 de l'introduction, et ne sera que brièvement rappelé dans la partie Introduction de ce chapitre.

Parce qu'il est prévu que les résultats de chapitre soient publiés sous la forme d'articles scientifiques, et parce que les deux rapporteurs de cette thèse sont anglophones, il a été décidé que ce chapitre soit rédigé en anglais, sans pour autant qu'il constitue un article pouvant se lire indépendamment du reste de la thèse. L'intérêt principal de ce format est de faciliter une critique non-seulement de fond mais aussi de forme, ainsi que de faciliter le formatage de ce chapitre pour publication après la thèse.

## **1. Introduction**

Understanding patterns of diversification is the starting point for better understanding past and current biodiversity dynamics as well as empirically exploring complex evolutionary concepts such as “adaptation”, “evolvability” or “evolutionary success”. However, reconstructing such diversity patterns remains challenging for most taxa, because it needs several prerequisites such as (i) an as complete as possible sampling or robust species hypothesis (see the case of *Lautoconus* in Phuong *et al.*, 2019), (ii) a robust (but see Thomas *et al.*, 2013), dated (although it is not a strict requirement, cf. Morlon, 2014) species-level (but see Alfaro *et al.*, 2009) phylogeny, and (iii) documented traits that may correspond to key innovations. Moreover, and this concerns all the taxa, another challenge is the development of reliable methods that can estimate rates of diversification or test correlations between traits and diversification rates. The democratization of molecular methods, especially HTS, during the past 10 years (see also article 2 - Chapter 2) allowed the standardization of the two first points. Moreover, the development of imagery and microscopy techniques are currently accelerating the acquisition and visualization of morpho-anatomical data (Ziegler & Menze, 2018), with innovative results for malacological research (Ziegler *et al.*, 2018). Finally the development of user-friendly programs for inferring diversification rates (e.g. Alfaro *et al.*, 2009; FitzJohn, 2012; Rabosky, 2014; Morlon *et al.*, 2016) have greatly simplified the use of complex statistical methods for the layperson. These progresses firstly concerned already well-studied groups, and molecular-based diversification studies have traditionally been limited to groups such as vertebrates (e.g. Jetz *et al.*, 2012; Shi & Rabosky, 2015; Malmstrøm *et al.*, 2016; Feng *et al.*, 2017; Arbour & Santana, 2017; Stepan & Schenk, 2017), angiosperms (e.g. Weber & Agrawal, 2014; Puttick *et al.*, 2015; Lagomarsino *et al.*, 2016; Landis *et al.*, 2017; Igea *et al.*, 2017; Rose *et al.*, 2018) or terrestrial arthropods (e.g. Espeland *et al.*, 2015; Kozak *et al.*, 2015; Legendre & Condamine, 2018; Zhang *et al.*, 2018; Economo *et al.*, 2018; Fernandez *et al.*, 2018). In contrast, fewer studies have been conducted in other groups, in particular for marine invertebrates (but see Davis *et al.*, 2016; Li *et al.*, 2016; Pérez-Escobar *et al.*, 2017; Cunha *et al.*, 2017; Davis *et al.*, 2018 ; O’Hara *et al.*, 2019), for which sampling often constitutes an additional obstacle to fulfill the prerequisites stated above.

Marine gastropods, along with other mollusks and a few other taxa (e.g. Brachiopoda, Cnidaria), constitute an exceptional model for studying diversification dynamics, because of their abundance and diversity in the fossil record. For this reason, the diversification history

of marine gastropod have long been studied (e.g. Sohl, 1969; Kohn, 1990), usually through diversity curves or kite diagrams generally inferred from fossil species or genera richness collected level by level in the sedimentary record. Malacologists are sitting on a goldmine, and combining data (mostly paleontological) from previous studies with modern approaches is one of their biggest challenges for the next century. Very few contemporary studies, however, have combined the three prerequisites presented above. Several reasons (unstable taxonomy, phenotypic plasticity, etc.) have been detailed in the Introduction and Chapter 1 of this manuscript, well exemplified by the Turridae. Thus, when looking at the pre-HTS gastropod literature, only a few studies on gastropod diversification based on molecular data are available (e.g. Cunha *et al.*, 2005). While the combined development of integrative taxonomy and HTS methods should have led to an increase of gastropod diversification studies, the literature appears to be poor on that matter. In fact, a quick bibliometric survey of articles in the past 15 years citing either LASER (Rabosky, 2006), GEIGER (Harmon *et al.*, 2008), MEDUSA (Alfaro *et al.*, 2009), DIVERSITREE (FitzJohn, 2012), BAMM (Rabosky, 2014) or RPANDA (Morlon *et al.*, 2016) - six of the most cited tools for studying diversification in a phylogenetic context - shows that only a handful of gastropod studies used these mainstream tools (Williams & Duda, 2008; Williams *et al.*, 2013; Postaire *et al.*, 2014; Krug *et al.*, 2015; Cunha *et al.*, 2014; Cunha *et al.*, 2017; Abdelkrim *et al.*, 2018a; Phuong *et al.*, 2019; Modica, Gorson *et al.*, 2019). Among these nine articles, five publications concern conoideans, of which two were the only ones based on HTS data. In five studies, diversification was studied in relation with geological/geographical scenarios (Williams & Duda, 2008; Williams *et al.*, 2013; Postaire *et al.*, 2014; Cunha *et al.*, 2014; Cunha *et al.*, 2017), while Krug *et al.* (2015) presented results of correlation between diversification dynamics and larval development types. For the three Conoidea studies published in the past two years, in addition to diversification shifts identification, diversification was systematically studied in relation with trait correlation (i.e. potential key innovations). It is important to stress out that important sampling effort was conducted to document many different types of extrinsic and intrinsic traits (especially in Phuong *et al.*, 2019; Gorson, Modica *et al.*, 2019). Despite these efforts, no particular trait was found to be correlated with diversification rates, except maybe a promising yet inconclusive result in the whole Conoidea family, for which the families with higher radula diversity showed increases in diversification rates (Abdelkrim *et al.*, 2018a). This lack of results includes traits related to the (supposedly) key innovation of the group – the venom gland.



As all Conoidea families, the Turridae have suffered from taxonomic instability (see the Introduction). However, 15 years of exhaustive molecular sampling and integrative taxonomy practice (see Chapter 1) are slowly leading towards a comprehensive set of species hypotheses. Using this stabilized, yet not fully complete, taxonomic context, we here investigate the phylogeny of this group using a sequence capture technique previously successfully applied to other conoideans (Abdelkrim *et al.*, 2018a; Phuong *et al.*, 2019). Diversification rates and possible rate shifts were analyzed using several, complementary methods. Diversification dynamics in comparison to two other Conoidea families, the Conidae and the Terebridae, are discussed. We hypothesized that the Turridae, as being part of the same hyperdiverse group as the Conidae and Terebridae, should follow the same diversification dynamics. In addition to the venom apparatus, shell size and shape are influenced by ecological and evolutionary history, and trait diversity patterns associated with diversification shifts might reveal unsuspected innovations. Similarly, the protoconch type is highly suspected of having an impact on diversification processes (Krug *et al.* 2015), but this character is always treated as a two-state character, and its continuous nature has thus been neglected (Krug *et al.*, 2015; Modica, Gorson *et al.*, 2019). Another trait of importance would be the bathymetric distribution of species. Deep environments might shape evolutionary constraints, and be linked to niche colonization events (Modica, Gorson *et al.*, 2019). Such environmental adaptations might also shape the nucleotide composition of genomes (Foerstner *et al.*, 2005) and this composition could be detected using the genome-reduced approaches applied here. To test all these hypotheses, several continuous traits – some of which never treated as such - were documented, among which shell and protoconch biometrics, shell shape vectors, bathymetry and GC content of the transcriptomes. Correlations between diversification rates and traits and among traits were tested in order to better understanding the Turridae evolutionary history.

## **2. Material and Methods**

### **2.1 Bait Design, Sampling, Library Preparation, Hybridization and Sequencing**

A target sequencing approach was used to recover exons for phylogenetic inference. First, a set of 27 transcriptomes (including three outgroups) was processed following the pipeline from article 2 (Chapter 2) and using *Lottia gigantea* as a reference-genome. Exon-boundaries were denoted by comparing the transcriptome sequences to the *L. gigantea*

genome. Only exons > 180 bp were retained. Ancestral sequences were generated for each exon alignment using FastML v3.1 (Ashkenazy *et al.*, 2012). Then, markers with a GC content <30% or >70% were excluded because extreme GC contents can reduce capture efficiency (Bi *et al.* 2012). Loci with repeats identified through the RepeatMasker v4.0.6 web server (Smit *et al.* 2015) were also removed. Then, we performed a self-blast with the target sequences via blastn v2.2.31 (evalue = 1e-10) and removed loci that did not blast to themselves with sequence identity >90%. The sequences were sent to Mycroarray (Ann Arbor, MI) to synthesize a custom MYbaits-1 kit, which allowed for 19,966 bait sequences. The bait length was 120 bp with ~54bp spacing (~2.2x tiling density). Overall, the bait design targeted about 1,425,000 bp.

Specimens were selected to represent all supported lineages in the CO1 tree presented in Chapter 1. Particular attention was paid to sampling *Gemmula* specimens, because of its long-time suspected paraphyly. Most specimens in the final exon-capture dataset come from the MNHN collection, except for *G. hindsiana* (Museum of Comparative Zoology of Harvard), *Polystira picta* and *G. deshayesii* (Natural History Museum in London) (see Elec. Supp. Mat.). In total, 172 samples were selected, divided in three batches of respectively 60, 60 and 52 specimens each. DNA was extracted using the E.Z.N.A.® Mollusc DNA Kit (Omega Bio-tek, Norcross, GA, USA), following the manufacturer's recommendations.

Library preparation followed Meyer and Kircher (2010) with some minor modifications (see Abdelkrim *et al.*, 2018a). 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 7min (30s ON/30s 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 of H3 fragment 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. For our last batch (batch 3), most indexing PCR product started amplification after 20 cycles. Despite this anomaly, and after multiple tries, products were processed anyway for the next steps. After quantification of the indexing PCR, 120 ng of each library were pooled by groups of 10 samples, except for the batch 3 for which we pooled 13 samples per pool.

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 Technologies were used. Capture was conducted for 24h at 60°C on a BIO-RAD C1000 touch thermal cycler. Post-capture libraries were cleaned-up following MyBaits protocol and quantified using Qubit. Each library was amplified through a PCR reaction of 10 cycles. At this stage, capture success was assessed using two positive and two negative controls. These 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 observing a gain after capture in the number of cycles needed for the PCR to reach the threshold cycle 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 and characterized using an Agilent 2100 Bioanalyzer. All batches were sequenced on a lane of Illumina HiSeq 4000 paired-end (100 bp reads) (QB3, Berkeley, USA).

## 2.2 Data filtering and 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 (Magoc 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. The 29 Turridae transcriptomes used in Chapter 2 were added to the exon-capture bioinformatic pipeline, as well as the transcriptomes of three outgroups (different than those used for bait design) *Mitra mitra* (Neogastropoda, Mitridae), *Conus tessulatus* (Conidae, unpublished) and *Terebra subulata* (Terebridae, SRR2059224). The exon-capture raw reads from the five samples of Turridae from a previous exon-capture study (Abdelkrim *et al.*, 2018a) were added. We used blastn v2.2.31 (evalue = 1e-10, word size = 11) to associate contigs with the targeted loci. Then, we used EXONERATE v2.4.0 under the est2genome model to redefine our target sequences because many of the original predicted exons were actually composed of several smaller exon. To fix misassemblies and estimate average heterozygosity, we mapped reads using bowtie2 v2.2.7 (Langmead and Salzberg, 2012) with the very sensitive local and no

discordant options, marked duplicates using picard-tools v2.1.1 (<http://broadinstitute.github.io/picard>), 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  $<4X$  coverage, and 4) removed sequences if there was not a minimum of  $4X$  coverage across 70% of a particular sequence. Alignments for each locus were generated using mafft v7.222 (Katoh *et al.* 2005). Finally, custom python scripts were used on aligned loci for additional filtering. Alignments with less than or equal to 40% missing data were discarded. Sequences within alignments with less than 50 nucleotides were removed. Fragments at both beginning and ends of alignments were trimmed if they were represented by less than 50% of the tips in the alignments.

### 2.3 Phylogenomic analysis

Two sets of loci were used for phylogenetic inference, depending on taxon occupancy, i.e. the proportion of missing tips in a given loci (Sayyari *et al.*, 2017): DS1 with loci containing at least 4 tips/locus and DS2 with loci containing at least 50 tips/locus. Each dataset is then concatenated in a single matrix on which a maximum likelihood analysis was performed using IQ-tree (Nguyen *et al.*, 2015), using a GTR+F+R6 model. We then applied 1,000 ultrafast bootstraps (UFBoot) (Hoang *et al.* 2017) on each data set to obtain branch support. Individual trees for each locus were also generated for each dataset using IQ-tree and the associated best substitution model found with ModelFinder (Kalyaanamoorthy *et al.* 2017) following the BIC criterion. Then, the sets of trees for both datasets were used to infer a phylogeny using the coalescent-based approach implemented in ASTRAL-III (Zhang *et al.*, 2017). These four analyses will be referred as DS1-AS, DS1-IQ, DS2-AS, DS2-IQ. To account for extreme evolutionary rates, we also build a third dataset. Following Cunha & Giribet (2019), the 10% slowest and the 20% fastest evolving loci of the DS2 dataset were identified using trimAl (Capella-Gutierrez *et al.*, 2009) and removed. For this dataset (DS3), we inferred two trees using a concatenated (DS3-IQ) and a summary approach (DS3-AS) using the same parameters as for DS2IQ. The Matching Cluster (MC) metric as implemented in the TreeCmp software (Bogadnowiz *et al.*, 2012) was used to compare the six trees.

A dated tree was inferred on DSIQ50 (without the outgroups), referred herein as DS2-MB, using MrBayes 3.2.6 (Ronquist *et al.*, 2012), with three calibration points. The root

calibration points is actually a “secondary calibration” derived from the Conoidea tree of Abdelkrim *et al.* (2018a). A second calibration point for *Polystira* was identified from Todd & Rawlings (2014), who identified *Pleuroliria subsimilis* Casey, 1904 (Red Bluff Formation, Early Oligocene of Red Bluff, Mississippi) as possessing the characteristics of the clade including all recent *Polystira* species. Finally, the third calibration point corresponding to the “core-clade” of Turridae of the Indo-Pacific was defined with the help of Pierre Lozouet (MNHN) who identified *Gemmula nilssoni* (Deshayes, 1865) (Lower Eocene) as the oldest representative of this clade based on shape of the shell, presence of gemmules and protoconch characters. In addition, the nodes in DS1-IQ and DS2-IQ trees corresponding to the divergence of *G. hindsiana* and *P. chilensis*, two specimens with a lot of missing data, were constrained in order to accelerate Markov chain convergence. Two parallel analyses were run, each consisting of four Markov chains of 100,000,000 generations with a sampling frequency of one tree every 10,000 generations. The number of swaps was set 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. Convergence of the analysis was evaluated using Tracer 1.4.1 (Rambaut *et al.*, 2014), and the runs were stopped when all ESS values were >200. A consensus tree was then calculated after omitting the first 10% trees as burn-in.

In addition to the exon-capture dated tree, we reconstructed a more complete, CO1-based, dated tree using what could be referred as a “Russian Doll” approach consisting in three steps. First, the “genus-level” (see Tbl. 4) nodes of the DS2-MB tree were collapsed. Then, all remaining nodes were used as constraints and fixed calibration points for the CO1 dataset. Finally, we ran MrBayes with the same parameters as described above (except for the calibration points). This dataset will be referred herein as DSCO1.

#### 2.4 Trait data acquisition

So far, trait acquisition has been conducted only for species sampling of the exon-capture datasets. When possible, five adults and two juveniles per MOTU were selected.

Shell and protoconch sizes were measured and converted to log, and shell shapes were acquired following the protocols detailed in the article 1 (Chapter 1). We made sure that

shells, especially adults, were not broken, truncated or aberrant. Protoconchs were measured in standard position and the number of protoconch whorls counted according to Bouchet & Kantor (2004). Shells were measured for adult specimens only, in standard position. The variation of shell shapes was analyzed through their outlines, reconstructed from photographs using a Photoshop CC 2019 script and the package Momocs (Bonhomme *et al.* 2014) from the R software. We used Elliptical Fourier analysis (Giardina and Kuhl, 1977), which presents many advantages over alternative Fourier analyses (Bonhomme *et al.* 2014), to decompose the closed outlines of the shells into periodic functions. The number of harmonics necessary to describe the shape was selected to reach 99.9% of the cumulative harmonic power, ensuring that the majority of the shell outline variation was considered. Shapes were also checked by eye to ensure that enough complexity was captured. The obtained coefficients were normalized to allow the superposition of the first ellipse and make the shapes comparable. We checked that all the reconstructed outlines were appropriately aligned. A principal component analysis (PCA) was performed on the harmonics coefficients in order to capture an optimal shape variation with a minimal number of principal components. A total of four principal components were retained for diversification analysis. The median for each MOTU was calculated to obtain only one value per MOTU.

Radulae diversity is already documented for some groups that have already been revised or partially revised such as *Iotyrris/Xenuroturris*, *Gemmuloborsonia*, *Turris/Anulaturris*, *Lophiotoma* and *Cryptogemma*. More radulae material was also available from taxonomically broader studies (Kantor, 2006) or from unpublished material (Kantor, pers. com.). For the remaining lineages, we sampled at least two radula from two different species per lineage. Radulae were prepared using standard protocols (Kantor & Puillandre, 2012) and examined by scanning electron microscopes JEOL JSM 840A and Tescan VEGA II LSU at the MNHN.

Bathymetric distribution for each MOTU was extracted from the TDSB stations data retrieve through the «BasExp» application from MNHN. Non-MNHN species depths were retrieved from the literature.

Finally, GC content for each specimen on the DS2-IQ dataset using a perl script from J. Caballero found on the Biostar forum (<https://www.biostars.org/p/353323/>).

## 2.5 Diversification rates

Diversification rates were inferred from the DS2-MB and DSCO1 trees using Bayesian Analysis of Macroevolutionary Mixtures (BAMM) v.2.5.0 (Rabosky *et al.* 2013; Rabosky, 2014; Shi & Rabosky, 2015). Additionally, dated trees of the Conidae (Phuong *et al.*, 2019) and Terebridae (Modica, Gorson *et al.*, 2019) were also analyzed using BAMM. The exhaustiveness of the sampling of the Conidae, and in particular within the *Lautoconus* clade, where species oversplitting has been recently revealed (Abalde *et al.*, 2017), has been discussed in Phuong *et al.* (2019). However, instead of randomly removing half of the species as in Phuong *et al.* (2019), we removed one species from each of the eight pairs of sister species present in this group, to stay consistent with the species delimitation approach developed in Chapter 1. For each analysis, 20,000,000 generations of reversible jump Markov Chain Monte Carlo were run, with a burn in set at 10%. Priors were chosen using the `setBAMMpriors` command in the R package `BAMMtools` (Rabosky *et al.* 2014). Following Modica, Gorson *et al.* (2019), we also tested values ranging from 0.1 to 10 for the prior probability of rate shift, which has been shown to affect BAMM results (Moore *et al.* 2016; Rabosky *et al.* 2017), and selected the best final analysis according to the runs with the highest ESS values. To account for incomplete taxon sampling in the DS2-MB dataset, we used the CO1-based MOTU sampling to estimate total species diversity for each “genus-level” node. For the DSCO1 dataset, multiple runs with global sampling fraction ranging from 0.9 to 0.1 were performed. For the Conidae and the Terebridae trees, the sampling fractions from the original publications were kept.

The RPANDA method (Morlon *et al.*, 2016) was used as an alternative approach to estimate diversification rates. The same protocol and script used in Modica, Gorson *et al.* (2019) was applied to our DS2-MB and DSCO1 trees, testing six nested diversification models and allowing scenarios where extinction is not constant, contrary to BAMM (see Modica, Gorson *et al.* (2019). The best fitting model was chosen according to the resulting corrected Akaike information criterion (AICc).

## 2.6 Trait-dependent diversification and correlations among traits

The QuaSSE (Quantitative State Speciation and Extinction) algorithm (FitzJohn, 2010) was used as implemented in `diversitree` to test the relationship between diversification rates and the various traits obtained (except for radula for which not enough samples were available). Diversification is expressed as  $r = \lambda - \mu$ , where  $\lambda$  is the rate of speciation and



mu the rate of extinction (FitzJohn, 2010), both measured in lineages/Ma. Linear, sigmoidal and “hump” models (FitzJohn, 2010) were fitted against the DS2-MB tree, and then contrasted against a null model of constant diversification. When trait values were missing from some tips, the tips were discarded from the tree. Models were then compared using AIC scores.

It has been shown that model-based inferences of trait-dependent speciation methods, such as QuaSSE, can be very sensitive to type I error, i.e. too liberal in rejecting the null model (Rabosky & Goldberg, 2015). To account for that, we applied the *Es-sim* test (Harvey & Rabosky, 2018), which uses tree simulations to build a null distribution of trait-speciation associations for testing. *Es-sim* was run on the same tree (DS2-MB) and traits tested with QuaSSE, and results compared against a null distribution of 1,000,000 simulated trees.

We used comparative methods implemented in the R package “caper” (Orme *et al.*, 2013), with Pagel’s lambda optimized with maximum likelihood, to fit our data on models : number of protoconch whorls vs shell size (Modica, Gorson *et al.*, 2019), bathymetry vs shell size, bathymetry vs shape (principal components), bathymetry vs number of protoconch whorls and bathymetry vs GC content.

### **3. Results**

#### **3.1 Targeted sequencing data**

On average, 5,473,290 reads per specimen were sequenced, with a high heterogeneity (from 2764 to 40,868,381 reads). Moreover, some specimens did not pass the cleaning and filtering pipeline, despite having a proper number reads, sometimes due to a high proportion of duplicate reads. In particular, only five samples from batch3 (out of 52) passed the exon-cleaning pipeline. Our final sampling for the exon-capture phylogeny includes 110 Turridae terminals. For these, we recovered 4,180 targeted loci, with an average of 49 species per loci. Despite the unexpected poor sequencing quality, especially for the batch 3, almost all targeted lineages were retrieved. Overall, *Gemmula* sp. 27 and *G. leucotropis* are the only problematic (i.e. representing a deep lineage which placement is not supported or in a spurious phylogenetic position in the CO1 tree) lineage identified in the CO1 tree that are not represented in the exon-capture dataset. We also note the poor results on the important

specimen *Gemmula hindsiana* (type species of the polyphyletic *Gemmula*), for which only 14 loci were recovered.

### 3.2 Phylogeny

Statistics for the five phylogenetic analyses are summarized hereafter (Table 1):

Table 7: Summary statistics for DS1-3 analysis. Tax. occ. – mean number of taxa / locus; % supp. nodes all – mean percentage of nodes with 80% bootstrap or more; % supp. nodes – percentage of nodes with 80% bootstrap or more, or 95 posterior probability or more for summary trees; Locus dist. – mean single-locus tree distance against respective species tree; Dist. to root – mean distance of each tip to the root

Dataset	Nb of loci	Tax. occ.	Supp. nodes all	Align. Length (+ missing)	% supp. nodes	Mean support	Locus dist.	Dist. to root
DS1-IQ	4,178	49	28.8	634,645 (58%)	6	97.3	0.59	0.22
DS1-AS				-	32	88.5	0.59	4.83
DS2-IQ	2,001	73	30.0	314,854 (37%)	6	97.1	0.61	0.21
DS2-AS				-	30	91.4	0.61	4.92
DS3-IQ	1,380	73	29.8	214,812 (37%)	11	95.0	0.59	0.20
DS3-AS				-	32	88.5	0.59	4.87

Overall, the relationships remained very stable between all datasets. Most differences either concern a few short branches (e.g. the sister group of the “G. sp 03/G. sp 29” clade), or a few samples with a lot of missing data (e.g. *Gemmula hindsiana*). Despite having a more homogeneous composition of loci than for DS1 or DS2 the DS3-IQ dataset showed an overall smaller bootstrap support. The biggest difference between the DS3-IQ dataset and the DS1-IQ and DS2-IQ datasets is the key position of *Gemmula hindsiana*, however always very poorly supported. The overall support in the summary trees is less important than for the concatenated trees (Tbl. 1), especially for the short branches (see also Fig. 1). As already shown in previous empirical studies (e.g. Abdelkrim *et al.*, 2018b – see Appendix 3) and discussed in the literature (e.g. Siu-Ting *et al.*, 2019), the amount of missing data in concatenated datasets seems to increase support values.

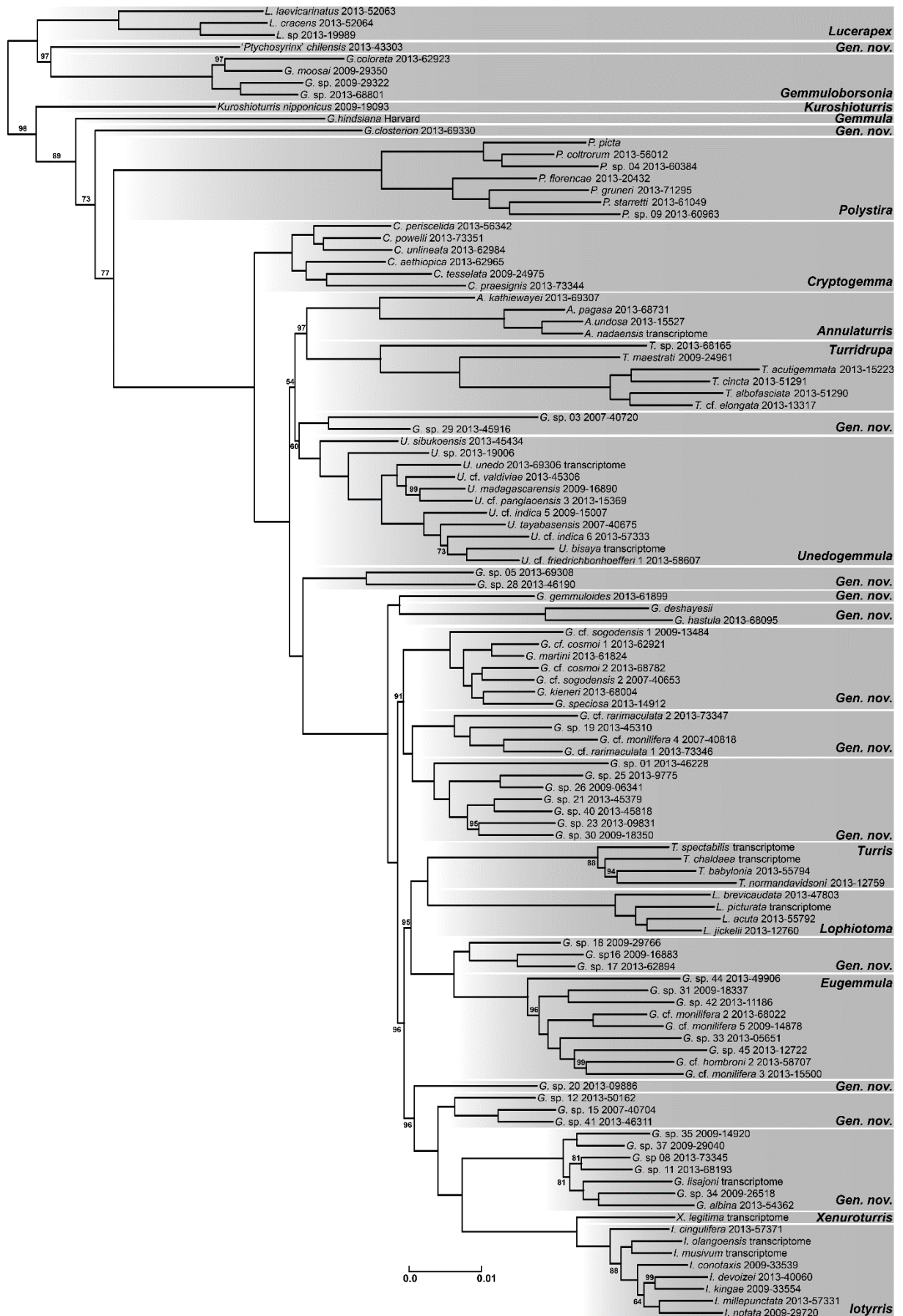


Figure 8: ML tree from DS2-IQ. Fully supported nodes not shown.

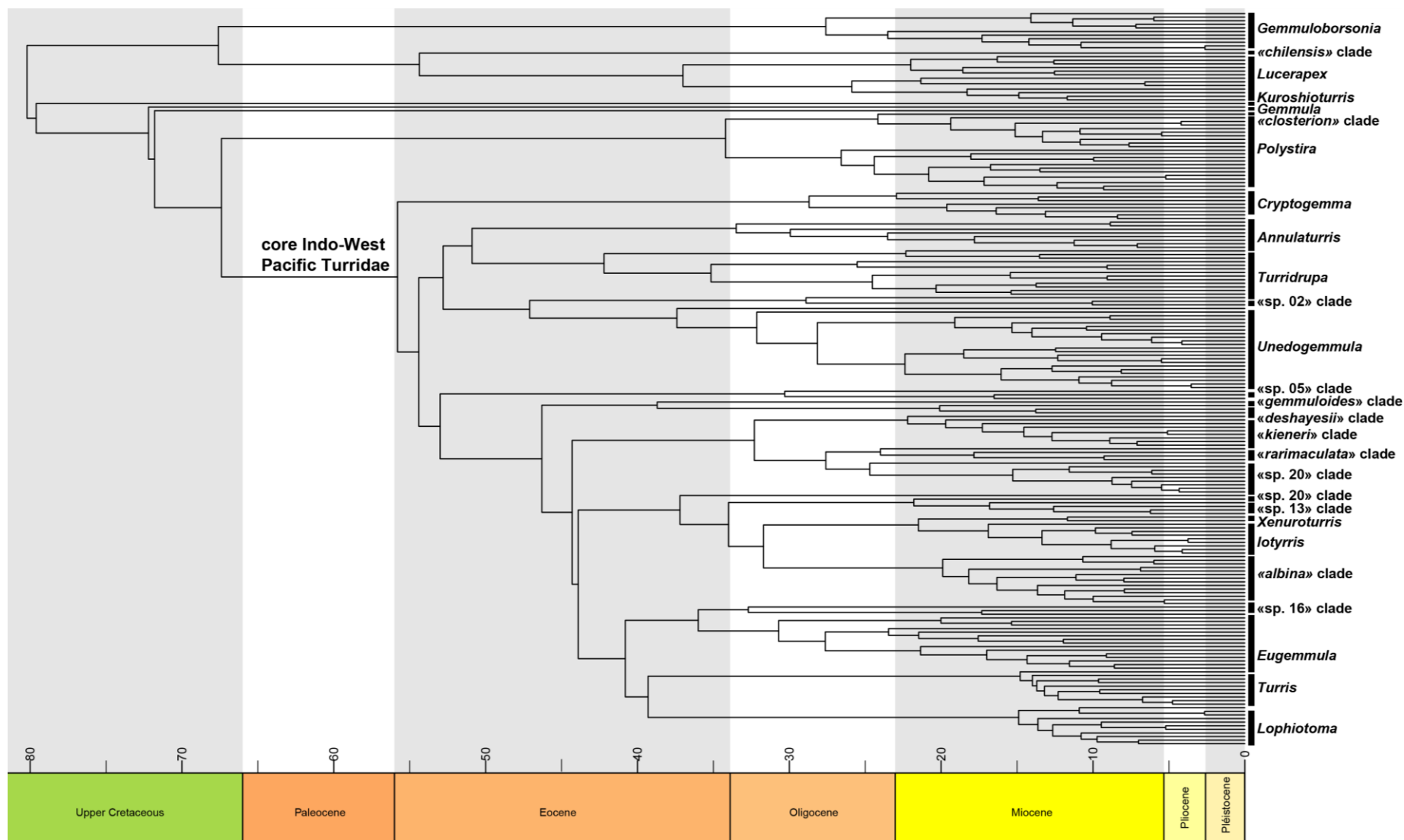


Figure 2: Dated tree obtained with DSCO1 dataset.

Unfortunately, at this stage of the study, the DS2-MB analysis did not reached enough generations to achieve proper convergence ( $ESS > 200$ ) on all parameters, currently being at 4,500,000 generations. Nevertheless, the log likelihood value of both chains has converged (-2,389,080) after ~2,000,000 generations, so we only sampled trees after this threshold. DivTime (dos Reis & Yang, 2019) and PATHd8 (Britton *et al.*, 2006) analyses were also run on the DS2-IQ tree to corroborate the DS2-MB tree. The internal calibrated nodes (*Polystira* and the “core-clade” of Turridae from Indo-Pacific) were at the lower boundaries of the confidence interval, indicating a possible older age for these nodes. The divergence of the “core-clade” of the Turridae from Indo-Pacific from its sister group, *Polystira*, could have occurred before the Cretaceous–Paleogene (K–Pg) extinction event. The most recent event of diversification is the divergence between *Unedogemmula cf. panglaoensis* and *U. cf. madagascarensis*, estimated at ~5.6 Ma. As expected because of its more complete sampling, the DSCO1 dated tree shows many more recent divergences (Fig. 2; Elec. Supp. Mat.). In most cases, a great heterogeneity is found between the divergence time estimates of sister species between the DS2-MB and DSCO1 datasets, e.g. *Gemmula.sp25/G.sp26* (15.8 vs. 5.2 My), *G. martini/G. cf. cosmoi* 1 (6.2 vs 1.4 My), *G/ sp16/G. sp 17* (25.3 vs. 16.6) (Fig. 2; Elec. Supp. Mat.).

### 3.3 Diversification rates

BAMM results on DS2-MB and DSCO1 trees showed consistent and comparable results, indicating no rate shift in a particular clade, but rather a general, positive, low and decreasing diversification rate for the whole Turridae (Fig. 3). The extinction rate was found to be almost null, so that the diversification rate can be considered synonym to the diversification rate.

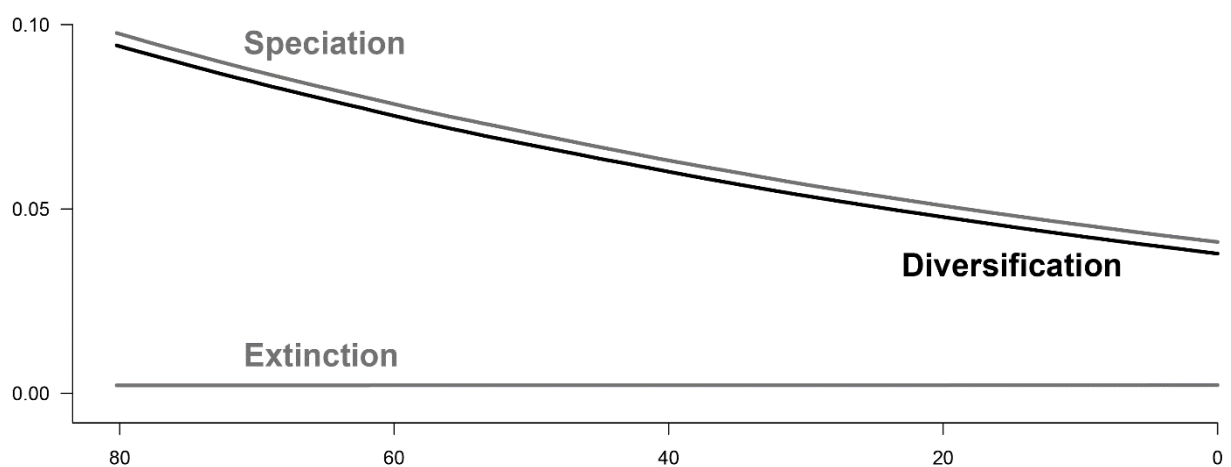


Figure 3: BAMM results on DS2-MB dataset. Rate values are indicated in the vertical axis, time before present is indicated in the horizontal axis.

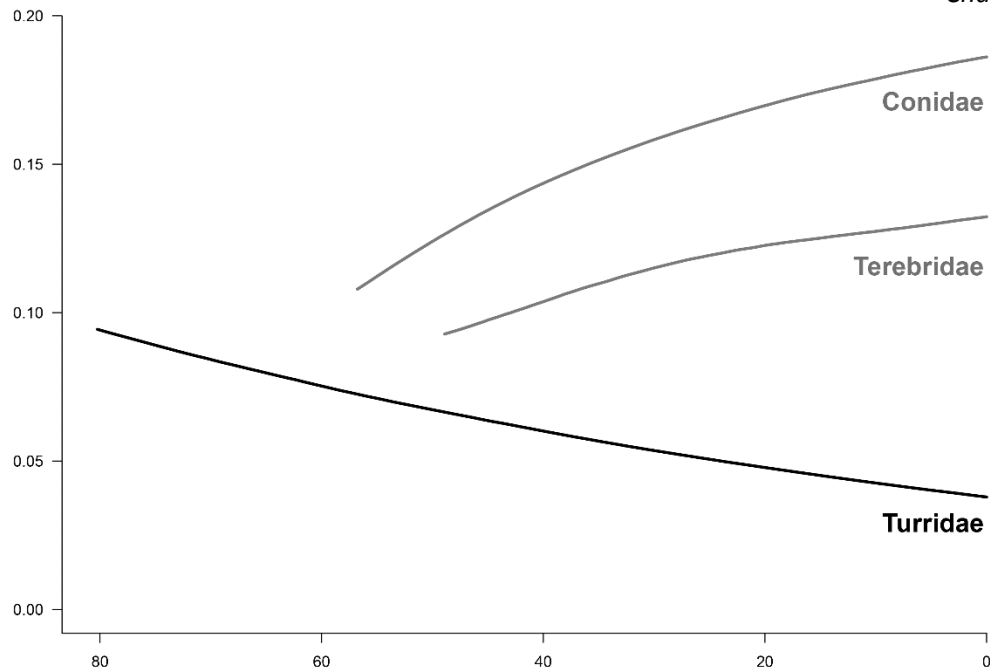


Figure 4: Bamm net diversification inferences for Conidae, Terebridae and Turridae. Rate values are indicated in the vertical axis, time before present is indicated in the horizontal axis.

BAMM results on Conidae and Terebridae are consistent with original results (Phuong *et al.*, 2019; Modica, Gorson *et al.*, 2019), indicating an increasing global diversification rate, greatly contrasting with the Turridae diversification curve (Fig. 4). The simulation of artificially increasing random sampling fractions on the DSCO1 dataset shows that BAMM would detect an increasing diversification in Turridae only if we were missing more than 50% of lineages (Fig. 5). The lambda and mu estimated values of BAMM on the DSCO1 showed very similar results, with a speciation rate of  $\sim 0.05$  at present and an extinction rate near zero. However, RPANDA yielded very contrasting results. The scenario selected with the highest AIC score is a time-dependent model, where both speciation and extinction are variable, each following a linear rate (Morlon *et al.*, 2011; Condamine *et al.*, 2019). The extinction rate was evaluated at  $\sim 0.03$  lineages/Ma while the speciation rate was evaluated at  $0.005$  lineages/Ma, resulting in a global negative diversification rate.

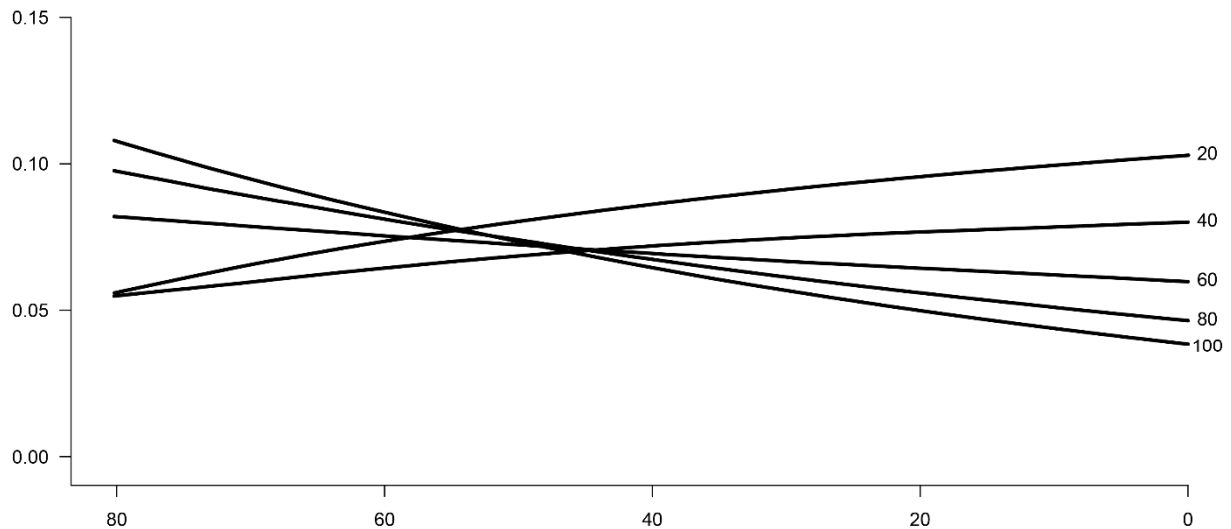


Figure 5: BMM net diversification rates for DSCO1 dataset, depending on the value fraction of sampling (numbers on the right, in %). Rate values are indicated in the vertical axis, time before present is indicated in the horizontal axis.

### 3.4 Trait-dependent diversification and trait correlations

QuaSSE analysis resulted in many statistically significant cases where birth and death depended on the character state, instead of rejecting the null model, with different best-fitted models (Tbl. 2). Most traits had speciation rates as linear functions, either positive (e.g. length and diameter of the shell, height and diameter of the protoconch) or negative (e.g. bathymetry, principal components 1 – corresponding mainly to the presence/absence of a siphonal canal). Other traits had speciation rates following a positive sigmoid function (e.g. number of protoconch whorls, aperture size). Overall, the only traits for which the null model of diversification was never rejected are the principal component 2 (corresponding mainly to the length/width ratio of the shell; Fig. 6) and the GC content.

In contrast, *Es-sim* yielded more conservative results, with significant positive correlations between size and speciation rate ( $\rho = 0.39$ ,  $p\text{-value} = 0.042^*$ ) and the number of protoconch whorls and speciation ( $\rho = 0.40$ ,  $p\text{-value} = 0.040^*$ ). Besides, *Es-sim* detected a slight negative correlation between bathymetry and speciation rates ( $\rho = -0.36$ ,  $p\text{-value} = 0.067\cdot$ ) (Tbl. 2).



Table 2: QuaSSE and ES-sim results on trait-dependent diversification. Only traits with significant correlation are shown. Length 1 corresponds to the total length of a shell. Length 2 corresponds to the length measure without siphonal canal.

Trait	QuaSSE					ES-sim	
	Best fitting model	Df	Pr(> Chi )	lambda	mu	rho	p-value
Length 1	linear	5	1.106e-05 ***	0.19	0.0004	0.39	0.042 *
Length 2	linear	5	6.311e-06 ***	0.28	0.0005	0.39	0.045 *
Width	linear	5	2.365e-05 ***	0.23	0.0005	0.36	0.066 .
PW	sigmoid	7	9.986e-06 ***	-	0.00004	0.40	0.040 *
PC1	linear	5	0.02448 *	-1.16	0.00004	-0.36	0.067 .
Prof	hump	7	3.628e-08 ***	-	0.019	-0.18	0.405

Finally, no significant relationship was found between bathymetry and most body size traits, bathymetry and the number of protoconch whorls and bathymetry and GC content. However, we found that the shell width, protoconch width and protoconch height variables explained the distribution of the bathymetric variable, showing significant relationships. Finally, a negative link was found between the bathymetry and the PC2.

## **4. Discussion**

### **4.1 Capture results**

The target sequencing results showed low performance in comparison with previous comparable studies (Abdelkrim *et al.*, 2018a; Phuong & Mahardika, 2018; Phuong *et al.*, 2019), in particular for the final number of taxa recovered at the end of the exon cleaning and filtering pipeline. This is mostly due to the poor results of the batch 3, which could be explained by the indexing PCR step and the high number of indexing PCR cycles that was needed for most samples. Many duplicates were found for most samples, possibly indicating that only a few targets were over-amplified during the indexing PCR step. A more positive result with the batch 3 is that the number of pooled samples before the capture step (13 in the batch 3 instead of 10 in the batches 1 & 2) did not affect the total number of reads recovered per specimen after sequencing. These results suggest that, in the future, more samples could be pooled together at the capture step, decreasing the cost per specimen.

#### 4.2. Locus filtering

A part of the exons with high evolutionary rates ( $> 10\%$  of overall divergence) were found to be exons with homology issues (i.e. heterozygosity and/or a mix of paralogous sequences) (results not shown). This implies that these non-orthologous exons still passed the pipelines, despite all the filtering steps. There is no doubt that the pipeline can be improved and that alternative methods to exon filtering based on similarity must be developed. A recent method, *Clan\_check* (Siu-Ting *et al.*, 2019) has been proposed to improve the process of filtering. It is a “graph-based” method, that supposes that incontestable “clans” (the unrooted equivalent of a clade; Wilkinson *et al.*, 2007), i.e. Mammalia or Aves if the ingroup is Vertebrata, should be systematically found in single-locus trees. By attributing a score for each clan found in each single-locus tree, one could choose to only keep the loci with a score above an arbitrary threshold. This method could be helpful in detecting locus alignments that include paralogues, for example if the same set of non-related taxa is often found together in single-locus trees\*. However, one limit of this method is that clans could be difficult to define in groups for which little phylogenetic investigation has been conducted, i.e. the Turridae or some other Conoidea family. Moreover, it implies that the loci have enough phylogenetic signal to recover the clan, which is not the case for most small loci (see article 2 - Chapter 2). In the future, a combination of such “graph-based” methods along with more traditional alignment metrics would surely help improving not only the detection of non-homologous or non-informative loci, but also cleaning and correction of those, instead of just removing them (as we did for DS3).

#### 4.3 Phylogenetic uncertainties

Phylogenomics has led to a renewal of phylogenetics, not only because of the unprecedented amount of data used in such analysis, but also, and above all, in the creation of many alternative methods of phylogenetic reconstruction (Edwards, 2009; Leaché & Oak, 2017). Past and current practice shows that two main families of methods are often used and compared: concatenated datasets and summary approaches (most of the time graph-based). The Matching Cluster metric showed that trees obtained with the same methods are more similar (Tbl. 3), despite being composed of very different sets of loci. This suggests in our

---

\*As an anecdote, a very similar method was proposed independently during intensive discussions with Alexander Fedosov and Nicolas Puillandre in spring 2019, precisely during the publication of Siu-Ting *et al.*'s article. A python code of the method has even been written by A. Fedosov and it is currently implemented in a phylogenomic pipeline still unpublished.

particular case that the impact of the method is greater than the impact of the dataset. There is a great amount of literature on the comparison of summary vs. concatenated methods to infer phylogenetic trees with the same datasets (Gatesy *et al.*, 2002; Bininda-Emonds, 2004), but one might ask: are the final trees obtained with each family of method really comparable? Obviously, the use of different entities (e.g. nucleotides vs. quartet graphs), combined with different criteria for calculating the tree (e.g. use of models, statistical consistent coalescent-based criterion, etc.) renders the different methods difficult to compare – some would say incommensurable. In addition to the incommensurability of the methods, their specificity could lead to the use of adapted datasets, e.g. minimizing taxon sampling for summary analysis (e.g. Streicher *et al.*, 2015), for each method, furthering apart the possible comparison between trees. Indeed, a few genes can have a great influence on phylogenies, even at genomic scales (Shen *et al.*, 2017). So, if the methods and the datasets are not comparable, can the final trees be? What happens if two different methods both show supported topologies, but that the clades are different? If all methods show the same topology, does that really mean that they are congruent? In fact, this debate exists since long time, and many discussions (e.g. Nelson & Platnick, 1981; Giribet *et al.*, 2002; Zaragüeta & Pécaud, 2016) have concluded on the importance of having a philosophical rationale when it comes to choosing methods. Regardless of the epistemological significance of those debates, the current, widespread pluralistic practice of phylogenetics (including in this PhD) shows us that philosophical rationale seems to be of little matter for most phylogeneticists. Once again, as shown in Chapter 1 about the species concept, there seems to be a gap between the philosophico-theoretical paradigm, the methodology and the reality of the scientific practice. In their lack of philosophical consistency, phylogeneticists will often rely on the experience acquired through the exploration of empirical datasets or external arguments from the literature to decide on which topology they will continue their investigation.

*Table 3:* Distance matrix between DS1-3 datasets final trees. The distance between each tree was measured using the Matching-Cluster metric.

	<b>DS1-AS</b>	<b>DS1-IQ</b>	<b>DS2-AS</b>	<b>DS2-IQ</b>	<b>DS3-AS</b>	<b>DS3-IQ</b>
<b>DS1-AS</b>	0	277	0	191	93	173
<b>DS1-IQ</b>		0	191	176	212	144
<b>DS2-AS</b>			0	277	93	173
<b>DS2-IQ</b>				0	204	146
<b>DS3-AS</b>					0	146
<b>DS3-IQ</b>						0

In the case of Turridae, we mostly relied on the concatenated datasets for discussing classification. Summary methods and statistical binning (Mirarab *et al.*, 2014) can give erroneous or problematic results when single-locus tree are short and poorly informative (Streicher *et al.*, 2018) and it has already been the case for a similar set of exons (Abdelkrim *et al.*, 2018a). Furthermore, we focused on the DS2-IQ for the diversification studies presented in this chapter, but further inquiry will require an improvement of the locus-cleaning pipeline and might lead to a newly designed dataset.

#### 4.4 Towards a new classification of the Turridae

We were able to sequence one or several samples for 15 out of the 16 extant and valid genera listed in Worms (2019). All phylogenies retrieved previous assessment of species to genera as in Worms in most cases: the monophyly of *Gemmuloborsonia*, *Lucerapex*, *Polystira*, *Cryptogemma*, *Unedogemmula*, *Lophiotoma Iotyrris/Xenuroturris* is confirmed. The two clades of *Turris* (*Turris* and ‘*Annulaturris*’) found in Fedosov *et al.*, 2012 are retrieved, confirming the necessity of resurrecting the name *Annulaturris*. The genus *Decolidrillia* Habe & Ito, 1965 is not a Turridae, but more likely close to the Fusiturridae-Horaiclavidae-Clavatulidae (FHC) group found in Abdelkrim *et al.* (2018a) (results not shown). The species *chilensis* Berry, 1968, formerly attributed to the genus *Ptychosyrinx*, synonymized with *Cryptogemma*, (see article 1 - Chapter 1), should be included in a new, monospecific genus. The major outcome of this phylogeny is the confirmation of the paraphyly of the genus *Gemmula*, that we propose to split into 15 genera; for 13 of them we were not able to attribute any available genus-level name, and they should thus be described as new genera (Tbl. 4). For now, we favored a topology (DS2-IQ) where *G. hindsiana* – the type species of the genus – is monotypic, but a new sampling of loci should be conducted on specimens of this species to ensure that it is not a sister species of *G. closterion* (as with DS3-IQ and summary datasets). The splitting of *Gemmula* into 15 genera might also be a slight over-splitting, especially since half of these genera are composed of one to three species. Should the sister-relationships between some of these groups be confirmed, the number of new genera to describe could be lowered. In particular, based on the DS1/DS2/DS3-IQ phylogenies, the “G. sp. 16 clade” (plates 5) could be combined with *Eugemmula* (C6-R6 Plates in Introduction), and the “G. sp. 01 clade” clade with the “*rarimaculata* clade” (C5-R5 Plates), considering that they share similar shell and radula features, leading to the description of 11 new genera instead of 13.

Another difficulty is the attribution of available species names to the unnamed MOTUs, and eventually the description of new species: as illustrated for *Cryptogemma* (article 1, Chapter 1), *Xenuroturrus/Iotyrris* (Appendix 3) and *Lophiotoma* (Appendix 1), these steps are very time consuming and can be problematic in the case of unpreserved or lost name-bearing specimens. Surely, the work should continue gradually, as it has been done until now, focusing on a single genus at a time.

#### 4.5 A note on some traits

The extensive sampling of radulae in “*Gemmula*” lineages confirmed the low diversity of radula types within Turridae (see also Introduction; Abdelkrim *et al.*, 2018a). However, some easily detectable features, such as the central tooth, indicate a certain plasticity of this structure. Hypothesizing that the ancestral state is the presence of a strong central tooth (Kantor, 2006), this state appears to have been lost several times independently during the evolution of the Turridae. Little is known about the evolutionary processes that could lead to a loss of this state. In other families of Conoidea (e.g. Conidae, Terebridae), the reduction of the teeth are generally linked to the acquisition of another derived state, the hypodermic tooth. However, no hypodermic teeth are found in Turridae. Thus, we cannot exclude the possibility that this particular state follows a random pattern in evolution. The use of methods like *Essim*, which compare trait-speciation correlations against a null – Brownian-like – model, could help us answering this question. However, a much more exhaustive and homogeneous sampling of radulae across the Turridae tree would be necessary in order to increase statistical power.

Table 4: Newly proposed classification of Turridae

Genus (Worms 2019)	Genus (this manuscript)	type species	Nb species (Worms 2019)	Nb MOTUs (this manuscript)	Corresponding plates
<i>Turris</i> Batsch 1789	<i>Turris</i> Batsch 1789	<i>Turris babylonia</i> (Linné 1758)	32	10	7
	<i>Annulaturris</i> Powell 1966	<i>Annulaturris amicta</i> (E. A. Smith, 1877)		8	7
<i>Cryptogemma</i> Dall 1918	<i>Cryptogemma</i> Dall 1918	<i>Cryptogemma benthima</i> (Dall, 1908)	11	8	cf. Chapter 1
<i>Pinguiggemula</i> McNeil 1961		<i>Pinguiggemula okinavensis</i> McNeil 1961 †	1		
<i>Ptychosyrinx</i> Thiele 1925		<i>Ptychosyrinx bisinuata</i> (Martens, 1901)	5		
	"chilensis" clade	N.A.	1	1	1
<i>Decollidrillia</i> Habe & Ito, 1965	not a Turridae	<i>Decollidrillia nigra</i> Habe & Ito, 1965	1	N.A.	X
<i>Epidirella</i> Iredale, 1931	never sequenced	<i>Epidirella xanthophaes</i> (R. B. Watson, 1886)	1	N.A.	X
<i>Gemmuloborsonia</i> Shuto, 1989	<i>Gemmuloborsonia</i> Shuto, 1989	<i>Gemmuloborsonia fierstinei</i> Shuto, 1989 †	7	11	1
<i>Iotyrris</i> Medinskaya & Sysoev, 2001	<i>Iotyrris</i> Medinskaya & Sysoev, 2001	<i>Iotyrris marquesensis</i> Sysoev, 2002	8	9	cf. Appendix 3
<i>Kuroshioturris</i> Shuto, 1961	<i>Kuroshioturris</i> Shuto, 1961	<i>Gemmula</i> ( <i>Kuroshioturris</i> ) <i>hyugaensis</i> Shuto, 1961 †	5	1	1
<i>Lophiotoma</i> T. L. Casey, 1904	<i>Lophiotoma</i> T. L. Casey, 1904	<i>Lophiotoma acuta</i> (Perry, 1811)	24	11	cf. Appendix 1
<i>Lucerapex</i> Wenz, 1943	<i>Lucerapex</i> Wenz, 1943	<i>Lucerapex casearia</i> (Hedley & Petterd, 1906)	8	15	1
<i>Polystira</i> Woodring, 1928	<i>Polystira</i> Woodring, 1928	<i>Polystira albida</i> (G. Perry, 1811)	25	22	2
<i>Turridrupa</i> Hedley, 1922	<i>Turridrupa</i> Hedley, 1922	<i>Turridrupa acutigemmata</i> (E. A. Smith, 1877)	18	13	2
<i>Unedogemmula</i> MacNeil, 1961	<i>Unedogemmula</i> MacNeil, 1961	<i>Unedogemmula unedo</i> (Kiener, 1839)	3	23	3
<i>Xenuroturris</i> Iredale, 1929	<i>Xenuroturris</i> Iredale, 1929	<i>Xenuroturris legitima</i> Iredale, 1929	6	2	cf. Appendix 3
<i>Gemmula</i> Weinkauff, 1875	<i>Gemmula</i> Weinkauff, 1875	<i>Gemmula hindsiana</i> Berry, 1958	62	1	1
	<i>Eugemmula</i> Iredale 1931	<i>Eugemmula hawleyi</i> Iredale, 1931		15	6
	"closterion" clade	N.A.		2	1
	"sp. 02" clade	N.A.		3	4
	"sp. 05" clade	N.A.		3	4
	"gemmuloides" clade	N.A.		1	4
	"deshayesii" clade	N.A.		3	4
	"kieneri" clade	N.A.		8	4
	"rarimaculata" clade	N.A.		4	5
	"sp. 01" clade	N.A.		9	5
	"sp. 16" clade	N.A.		3	6
	"sp. 20" clade	N.A.		1	8
	"sp. 12" clade	N.A.		5	8
	"albina" clade	N.A.		14	8
	"sp. 27" clade	N.A.		1	5

The use of a semi-automatic pipeline of morphometric data acquisition allowed us to characterize shell shapes for 384 specimens (corresponding to 110 MOTUs) in a short amount of time. However, the data acquisition relied on a heterogeneous set of pictures taken at different times, by different people, meaning that the orientation of the shells was not necessarily homologous from one picture to another. Thus, the first two principal components captured easily interpretable variation in shapes (Fig. 6), but interpreting the variation of the

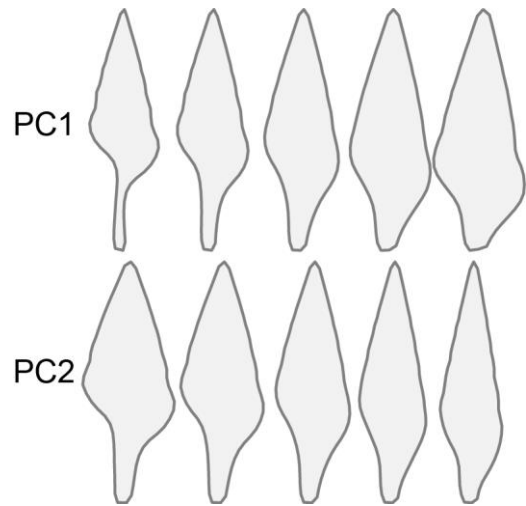


Figure 6: Shape variation in principal components 1&2

other principal components is more challenging. At this low level of variance explained by the principal components, the variability of a principal component could be as well explained by picture heterogeneity, and so we decided to discuss only the first two axes for our study.

#### 4.6 An unexpected low and decreasing rate of diversification

As a null hypothesis, we expected that the Turridae would follow about the same diversification patterns as the Conidae and Terebridae, considering that they are all part of the same hyperdiverse superfamily. However, our BAMM analyses of speciation and extinction estimations indicated an almost null extinction rate and a very low speciation rate ( $\sim 0.05$ ) at present (Fig. 3). Thus the BAMM diversification curve follows an exact inverse pattern as those found in Conidae and Terebridae (Fig. 4). BAMM analysis were run on alternative dated (using DivTime; Dos Reis & Yang, 2019 or PATHd8; Britton et al, 2006) or non-calibrated ultrametric trees (MrBayes) to confirm this decreasing rate tendency, producing similar results the for the DS2-MB and DSCO1 trees (results not shown). Finally, our simulation of increased missing data on BAMM analysis on the DSCO1 tree indicate that net diversity curve dynamics would invert if we were missing at least 50% of the lineages in Turridae, an improbable scenario given the diversity curves produced in Chapter 1. Nevertheless, while the speciation rate inferred with BAMM is realistic, a constant – nearly null – extinction rate makes little sense, especially because the Turridae have a rich fossil record (e.g. Todd & Jonhson, 2013) and evolved during 80 My. RPANDA allows for time-varying rates of extinction, and has been applied on a range of clade decline scenarios (Morlon *et al.*, 2011), so we expected more realistic values for  $\mu$ . Indeed, RPANDA found



positive extinction rates, even greater than speciation rates, resulting in a negative diversification rate (Fig. 7). This extreme case of clade decline is allowed by a few methods such as BAMM or RPANDA (Morlon *et al.*, 2011; Rabosky *et al.*, 2014), even in the case of extant-tips-only phylogenies. However, these results require further testing, and might necessitate the use of models that are more complex in order to account for diversification heterogeneity (Condamine, pers. com.).

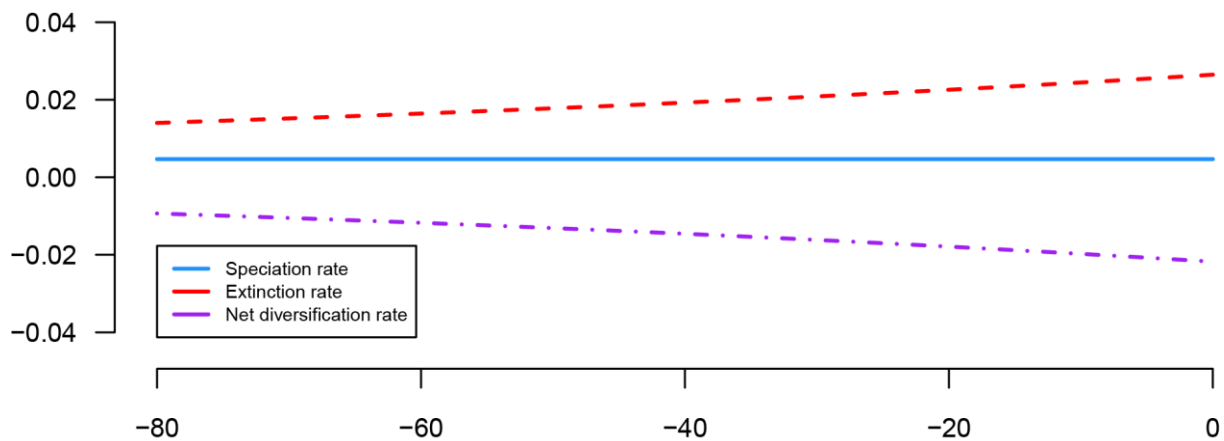


Figure 7: RPANDA results on diversification rates

Negative diversification rates has already been found in the literature but is a poorly discussed topic (see Burin *et al.*, 2019 for a bibliometric survey). Burin *et al.* (2019) produced one of the few – if not the only one – simulation studies comparing the ability of BAMM and RPANDA to detect declining diversification rates. Their results showed a tendency for RPANDA to favor increasing extinction and a tendency for BAMM to favor decreasing speciation when identifying decline in empirical trees. This is precisely what we found for both methods with our Turridae trees. However, their simulations also showed that both methods fail to characterize the temporal dynamics when extinction rate varies. Extinction rates are precisely what might be expected to vary, especially for old clades – such as the Turridae – that might have undergone several extinction crises. Burin *et al.* concluded “that current methods might have a difficult time on simultaneously inferring deep time dynamics and recent dynamics, especially for clades experiencing diversity decline”. We must therefore consider that both methods used in this chapter potentially provided erroneous results. We would tend to favor RPANDA results over BAMM results, because BAMM’s null and constant rate of extinction seems unrealistic, considering the richness of the Turridae fossil

record. Despite these inconclusive results, we will consider the scenario of a clade decline true for Turridae until proven otherwise.

#### 4.7 Traits affecting diversification

Despite the doubtful inferred diversification rates and the differences between the two methods on many key aspects, such as accounting for missing tips (*ES-sim* doesn't allow it), our trait-dependent speciation analyses resulted in many correlations rejecting the null model, with both QuaSSE and *Es-sim* (Tbl. 2). The results can be summarized in three main points: (i) overall shell size, especially shell length, is positively linked with speciation rate; (ii) the number of protoconch whorls is positively linked with speciation rate; (iii) bathymetry is negatively linked with speciation rates. An additional fourth result has been found only with QuaSSE: the principal component 1, corresponding more or less the presence/absence of a siphonal canal, is negatively linked to speciation rates, indicating a potential evolutionary advantage of possessing such a feature. Indeed, a recorded case of a *Turris* specimen feeding showed that the specimen was buried beneath the substrate but that the presence of a siphonal canal apparently allowed it to move its proboscis towards the polychaete prey and successfully envenomate the prey while still being buried (Olivera *et al.*, 2014). However, QuaSSE tends to report spuriously significant results when diversification decelerates (Machac, 2014), which would be precisely the case here. Since this result has not been confirmed by *Es-sim*, we will no further discuss it here.

Since all extinction rates found with QuaSSE were almost zero, we consider the speciation rate to be equivalent to the diversification rate thereafter. The three detected correlations indicate the possibility of an increase in speciation rates in taxa with either bigger shells, with more protoconch whorls, living in shallower habitats, or possibly any combination of these three parameters. However, we observe few examples of shallow-water Turridae species that are relatively big and with many protoconch whorls (but *Xenuroturris legitima* is a counter-example), suggesting that these tree states might not be linked together in the process of favoring speciation. Furthermore, our PGLS results indicate no link between these three variables, although testing properly this hypothesis would require model fitting between speciation rates and all these traits within in a multivariate analysis. However, PGLS showed links between bathymetry and protoconch width, shell width, and shell width/size ratio (PC2), suggesting that bathymetry might constraint shell shape.

The correlation between bathymetry and speciation has been discussed in the literature. Shallow water environments are expected to be more fragmented than deep-water environment, thus favoring speciation (Williams & Duda, 2008). However, the hypothesis that the colonization of deep-water environment favored the diversification of some taxa has also been proposed (e.g. Sysoev, 1991). Our results would tend to favor the first hypothesis, even if we report a number of shallow-water Turridae species with large distributions (Puillandre *et al.*, 2017; Abdelkrim *et al.*, 2018b – Appendix 1 and 3). Concerning the shell size, smaller sizes could be favored because of their ability to colonize microhabitats inaccessible for bigger specimens, such as rock crevices (Bouchet, pers. com.). But this hypothesis would apply to very small specimens (<10mm), while most of the Turridae adult species express a range of ~10-70 mm, a range of shell sizes in which the difference in terms of habitats accessibility might be negligible. Inversely, a bigger shell size could imply better protection against predators, or even unknown feeding strategies such as the *Turris* "hide-and-catch" strategy described above. Further investigation would require exploring other disciplines such as population genomics to test the fragmentation of different populations at different bathymetric levels, quantification of predatory marks on Turridae species of different shell sizes or even behavioral studies to compare to feeding strategies of smaller and bigger Turridae.

The remaining character – the number of protoconch whorls – shows a positive correlation between speciation and the number of protoconch whorls. Since the postulate is that a high number of protoconch whorls (i.e. multispiral type) allows for increased dispersal abilities, then it is expected to reduce speciation rate because it allows populations to maintain a high genetic connectivity over large distances, even in the presence of biogeographic barriers (Harvey & Rabosky, 2017). Thus we would expect, conversely, that a lower number of protoconch whorls would result in a higher speciation rate. Indeed, Krug *et al.* (2015) showed that lecithotrophic clades in Sacoglossa have an increased speciation rate but also an increased extinction rate, lowering the net diversification, while planktotrophic clades speciate less but also extinct much less, resulting in a higher net diversification rate than lecithotrophs. Our results contradict this hypothesis. We found that speciation rates vary as a sigmoid-shaped function of the number of protoconch whorls, suggesting that speciation rate is higher in lineages with (potentially) more dispersal abilities. One hypothesis could be that extreme capacity of dispersal could favor founding effect events. Indeed, most Turridae species could theoretically have an Indo-West Pacific distribution (see Chapter 1), but what

about the multiple islands from Central and East Pacific? Sampling effort has been limited and little is known on Turridae diversity in these regions. The reduction of genetic connectivity on remote Islands and/or divergent selection linked to the availability of new preys could favor speciation, as suggested for a cone snails species from the Marquesas Islands (Puillandre *et al.*, 2014).

The State Speciation and Extinction (SSE) family of methods, such as QuaSSE, as undergone a lot of criticisms. In addition to sensitivity to type I error already discussed above, the problem of model inadequacy has also been raised (Rabosky & Goldberg, 2015). Indeed, why should the null model (trait independent diversification on a constant rate tree) explain the data well? An even more serious problem is phylogenetic pseudo-replication (Maddison & FitzJohn, 2014): if a trait has evolved only once or just a few times, an observed effect on rates could be spurious. In Turridae, all the species with more than 3.5 protoconch whorls are found in the “core Indo-West Pacific” clade, suggesting a strong effect of non-replication for this character state. Alternative solutions have been developed to overcome some of these issues, such as the possibility to construct alternative null models (HiSSE; Beaulieu & O’Meara, 2016). But most of these methods have been designed for binary-state models. Therefore, it might be useful to go back to binary state approach of testing trait-dependent diversification, such as lecito-/planctotrophic states, shallow/non-shallow states and “big” and “small” shells (Modica, Gorson, *et al.*, 2019), even though it implies to not account for the continuous nature of these characters.

If so many traits (shell size, shallow bathymetry and increased number protoconch whorl) are positively linked to speciation, then why the global net diversification of Turridae is decreasing? There are many hypotheses that can explain diversification slowdowns (Moen & Morlon, 2014). One possible explanation is the niche saturation hypothesis: the Turridae would have conquered the niche they occupy today early during their history, followed by a quick saturation. This hypothesis would fit well the BAMM results, indicating a speciation slowdown during the history of Turridae. It would also fit the RPANDA results, if we combine them with the results on the number of protoconch whorls: the speciation rates have remained constant through time, but while niche occupation became more and more saturated, lectitotrophic species became less and less competitive, leading to an increase of extinctions and thus a decrease in diversification in the Turridae. These hypotheses are very preliminary and should be considered very cautiously, until further testing has been conducted.

#### 4.8 Future evolutionary research in Turridae

Although a member of the hyperdiverse Conoidea superfamily, the Turridae expressed a decrease in diversification rates, suggesting that the clade is in decline. Contrarily to previous similar studies on cone snail (Phuong *et al.*, 2019), we did not test the main hypothesis for diversification, that is, venom components and diversity would have a major impact on diversification. However, our analysis of the impact of some traits on diversification indicates a link between speciation and more traditionally-studied traits such as size, protoconch type, bathymetry, and possibly the presence of a siphonal canal.

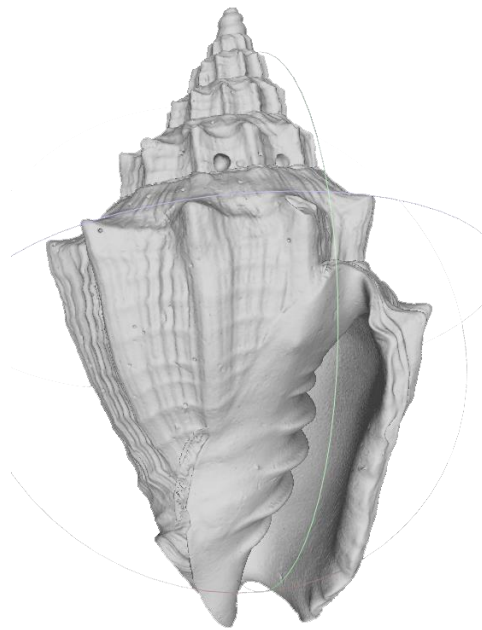
These results are preliminary and investigation must be continued. In particular, trait acquisition must be achieved for all samples from the DSCO1 tree, in order to double the sampling and increase statistical power. Since shell size and width appear to have a link with speciation, the acquisition of more complex shell traits could be done: integrating Raup parameters (Raup, 1966) could provide indirect information on ontogeny and shed new light on some possible links with speciation. This could be done semi-automatically by acquiring tridimensional shell shape (using e.g. micro-computed tomography). 3D shell morphometrics could also refine protoconch data acquisition, which also appears to be of great importance.

Past considerations on Turridae *s.l.* (see Introduction) already hypothesized that such groups possessing mostly plesiomorphic anatomical characters, especially radula, could be groups in decline. In the current phylogenetic framework of Conoidea, conducting diversification studies on a single family is not enough to answer this hypothesis. A diversification study encompassing all the Conoidea, using the most complete sampling possible, along with a clear characterization of plesiomorphic/apomorphic radula and anatomical states, should help to confirm – or not – if the Turridae, as other groups expressing solely pleomorphic states (e.g. Drillidae), are effectively in decline.

---

**All supplementary figures and tables cited in this chapter are available in electronic supplementary material.**

## **CONCLUSIONS & PERSPECTIVES**



© Atlas 3D – Sorbonne Universités

### A. Une stratégie opérationnelle de systématique exploratoire

Même s'ils ont été présentés à plusieurs reprises au cours de cette thèse comme un taxon non-modèle, les Turridae ont pourtant bien servi de modèle pour mettre à l'épreuve la stratégie de systématique proposée dans cette thèse (Fig. 1), stratégie préfigurée depuis 15 ans par l'équipe de recherche 3E mais adaptée et complétée au cours de cette thèse. A chaque étape, cette stratégie repose sur une approche semi-automatisée permettant un échantillonnage optimisé des individus, le séquençage ciblé de l'ADN et de l'ARN (~500 séquences CO1, plus de 4000 loci capturés pour plus de 110 spécimens et 21 transcriptomes obtenus pendant cette thèse), et la documentation de traits phénotypiques et environnementaux. Toutefois, aucune étape ne saurait se passer d'expertise et de connaissances préalables afin de les mener à bien :

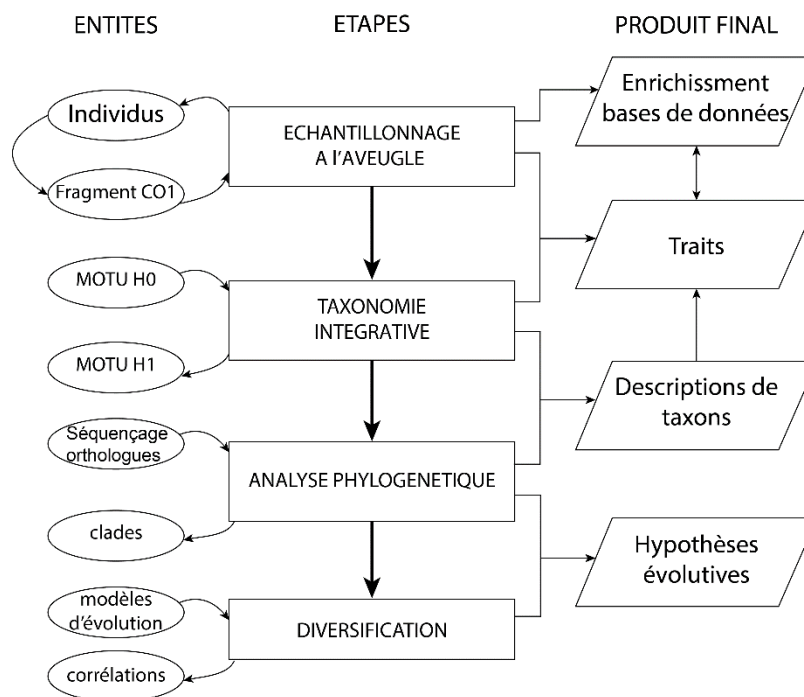


Figure 9: Stratégie générale de systématique utilisée pour les Turridae et proposée pour des taxons non-modèles

(i) Expertise de terrain. Plusieurs raisons peuvent expliquer qu'un taxon reste sous-étudié, comme le manque de spécialistes (par exemple, beaucoup de groupes de la méiofaune ; Bekkouche, 2016) ou une diversité trop importante à embrasser (par exemple certains groupes d'insectes). Une autre raison possible est la difficulté d'échantillonnage et la méconnaissance de l'environnement du groupe d'étude. Chez les Conoidea, la plupart des espèces sont petites (<5 cm) et vivent en milieu profond (> 300 m) (Bouchet *et al.*, 2009). La stratégie pratiquée depuis 43 ans dans le cadre du programme *Tropical Deep-Sea Benthos* n'est pas seulement



unique dans l'amplitude de son échantillonnage (7200 stations en 40 ans; Bary, 2018), mais également dans l'expérience et la connaissance de ce terrain méconnu qu'est le benthos profond. Nouveaux modèles de drague (« drague Warèn » en 1985), techniques de manipulation des chaluts adaptées aux monts-sous-marins, application d'un protocole opérationnel de préparation des spécimens à bord du navire, etc. La stratégie développée pour les Turridae et les données présentées dans cette thèse n'ont pu l'être que grâce à l'effort d'exploration des fonds-marins réalisé par mon équipe d'accueil.

(ii) Expertise de laboratoire. Bien entendu, qui dit moléculaire dit paillasse. Le développement du séquençage ADN – dont il ne serait pas exagéré de qualifier de « à croissance exponentielle » (Stein, 2010) – implique de s'adapter en permanence à de nouveaux protocoles de laboratoires. Même l'acquisition de petits fragments moléculaires universaux, comme le barcode CO1, et dont le séquençage par méthode Sanger est maîtrisé depuis 20 ans, est amenée à évoluer comme j'ai pu l'illustrer dans le chapitre 2 avec le projet CODEX. Malgré tout, le coût (en temps et en argent) des génomes (séquençage et assemblage) est encore trop élevé pour l'appliquer de manière exhaustive sur chaque individu. La communauté des systématiciens est donc en ce moment tournée vers des approches dites « génome réduit » (par exemple, séquençage RAD-tags – Annexe 3 - ou capture d'exon – Chapitre 3, Annexe 2) pour résoudre des problèmes de délimitation d'espèces ou de phylogénie. Toutefois, et ce malgré l'universalité de l'ADN, la spécificité de chaque groupe implique des protocoles adaptés (par exemple l'utilisation de kit spécialisés pour l'extraction d'ADN de mollusques) et pour lesquels nous disposons de peu de recul. Un exemple malencontreux est présenté rapidement dans le chapitre 3, où le troisième batch de capture d'exon sur les Turridae n'a pas été un succès à cause d'une seule étape du protocole de préparation de banques, pour laquelle il n'existe pas de contrôle qualité.

(iii) Expertise taxonomique. Quand bien même nous disposerions de génomes complets de tous les organismes, il est inconcevable d'envisager une taxonomie qui se prétende intégrative sans y intégrer des connaissances sur les traits morpho-anatomiques, fonctionnels, géographiques, écologiques, chimiques, etc. La taxonomie ne se limite plus à son étymologie (« règles de classification »), mais son objectif est bel et bien de produire une connaissance synthétique permettant d'identifier, reconnaître, caractériser, définir et décrire les taxons. En outre, la pratique de délimitation d'espèces implique de maîtriser un certain nombre de concepts, permettant d'éviter un certain nombre de pièges comme la confusion entre délimitation et description par exemple. Tout au long de ma thèse, j'ai pu collaborer avec des

malacologues experts du groupe des Conoidea (en particulier Yuri Kantor, Sasha Fedosov) afin d'apprendre les bases de l'observation, de la caractérisation et de la description des Turridae.

(iv) Expertise phylogénétique. Tout comme la taxonomie, construire une phylogénie implique de maîtriser un certain nombre de concepts anciens mais fondamentaux, comme l'homologie par exemple. De plus, le développement de méthodes toujours plus nombreuses de constructions d'arbres implique non seulement de les assimiler, mais aussi de justifier lesquelles choisir et utiliser. La tendance actuelle consiste à construire une série de jeux de données, puis d'utiliser quelques méthodes « en vogue », quelques méthodes plus « traditionnelles » et de comparer les topologies obtenues (ce que j'ai fait dans le chapitre 3). Certains auteurs n'hésitent pas à parler de congruence entre topologies, ou de soutien statistique d'un nœud s'il est retrouvé par plusieurs méthodes. Cette pratique pose toutefois problème si la comparaison entre ces différentes méthodes et topologies n'est pas véritablement cohérente et possible d'un point de vue théorique.

(v) Expertise bioinformatique et statistique. L'acquisition massive de données moléculaires et de traits phénotypiques et environnementaux pour une étude implique nécessairement l'utilisation d'outils informatiques et statistiques permettant de semi-automatiser le traitement des informations. L'utilisation de scripts personnalisés, où chaque étape est plus ou moins formalisée, permet d'éviter le problème de « boîte noire » que beaucoup d'outils informatiques constituent. En s'obligeant à comprendre ou à reconstruire soi-même les principales étapes d'un pipeline bioinformatique, le systématicien est forcé de prendre des décisions à chaque étape (par exemple, les valeurs d'un paramètre), et de se forger une opinion critique sur son propre jeu de données. De même, les méthodes modernes d'étude de diversification requièrent de construire des modèles (ou de poser des probabilités *a priori*) avant de tester les hypothèses, impliquant de posséder un certain bagage statistique. Un des regrets de ma thèse est d'avoir manqué de recul à l'égard des méthodes statistiques de diversification utilisées dans le chapitre 3. Pour contrebalancer cette impression, il est toutefois important de souligner que beaucoup de ces méthodes sont encore jeunes, et qu'elles n'ont pas encore été testées sur suffisamment de jeux de données (réels ou simulés) pour connaître les limites de chacune, et mieux interpréter les résultats parfois contradictoires qu'elles proposent (Burin *et al.*, 2019).

Ainsi, quel que soit la vitesse d'acquisition des données (moléculaires notamment), la systématique ne peut pas devenir une « turbo-systématique » comme certains le prétendent,

car cette discipline demande de s'approprier un contexte historique de l'étude d'un groupe (voir Introduction), une maîtrise de concepts théoriques et méthodologiques en perpétuels débats (voir chapitre 1-3), ainsi qu'une multitude de savoir-faire. De par les nombreux champs d'expertise à maîtriser, le systématicien se doit donc d'être un *jack of all trades* (« Touche-à-tout »). Mais en souhaitant s'approprier tous les domaines, il se condamne également à devenir – dans une certaine mesure – un *master of none* (« bon en rien »). Sans être polymathe, il doit être multipotentiel. En définitive, la spécialité du systématicien n'est-elle ni plus ni moins que le(s) groupe(s) qu'il étudie ?

Néanmoins, cette stratégie de systématique offre un cadre plus complet que les stratégies plus « traditionnelles ». Les hypothèses d'espèces sont construites sur la base d'un pipeline plus opérationnel et reproductible que l'utilisation de morpho-espèces. Combinées aux approches d'échantillonnage et de séquençage à l'aveugle, ces hypothèses d'espèces permettent des estimations d'abondance certainement plus pertinentes. Les approches « génomes-réduits » offrent la possibilité de véritablement mettre en congruence des milliers de loci, en comparaison des approches Sanger où l'on est bien souvent obligé de se contenter d'approches de concaténation. Tous ces avantages tendent vers un cadre systématique plus complet, précis et robuste, affectant directement les résultats des tests d'hypothèses de diversification.

## **B. Les Turridae ou comment être un clade en déclin et faire partie d'un groupe hyperdiversifié**

Les principaux résultats de ce travail peuvent être résumés en quelques points principaux :

(i) environ 204 hypothèses d'espèces ont été proposées chez les Turridae. Ces hypothèses n'ont pas toute la même robustesse – certains groupes (*Iotyrris*, *Lophiotoma*, *Cryptogemma*) ont été bien plus étudiés que d'autres. Ce chiffre confirme que la diversité spécifique des Turridae est bien moindre que dans d'autres familles de Conoidea.

(ii) Les approches de capture doivent être combinées avec des marqueurs classiques utilisés pour du séquençage Sanger, dans une optique finale de combiner des arbres à différentes échelles phylogénétiques. Une telle approche dite « poupées russes » a été expérimenté sur les Turridae en combinant les résultats de capture d'exon et un arbre CO1.

(iii) L'analyse phylogénétique confirme une grande diversité de clades chez le genre *Gemmula*, avec potentiellement 13 nouveaux genres à décrire. L'analyse datée confirme également un âge d'origine de cette famille bien plus ancien que d'autres familles de Conoidea comme les Conidae et les Terebridae. Cette différence d'âge pourrait potentiellement expliquer les différences constatées en termes de dynamiques de diversification, mais également compliquer une comparaison véritable avec ces deux autres familles.

(iv) les Turridae présentent un taux net global de diversification décroissant. Toutefois, il n'est pas clair si ce taux est décroissant mais positif et lié uniquement au taux de spéciation (résultats BAMM) ou bien négatif et lié uniquement au taux d'extinction (résultats RPANDA). Ces deux résultats sont également des illustrations des biais possibles de ces méthodes, et sont donc à considérer avec beaucoup de précaution.

(v) Il semble y avoir des liens positifs et indépendants entre le taux de diversification et d'une part la taille de la coquille et d'autre part le nombre de tours de la protoconque. Par ailleurs il semblerait qu'il y existe des liens négatifs plus faibles entre le taux de diversification et d'une part la bathymétrie et d'autre part la réduction du canal siphonal. Des hypothèses adaptatives ont été proposées pour chacune de ces corrélations. Toutefois, de même que pour les calculs du taux diversification, ces résultats sont à prendre avec énormément de précaution à cause des biais, ou tout simplement du manque de recul (personnel ou de la communauté) sur les méthodes utilisées.

(vi) Il semble y avoir un lien positif entre la largeur de la coquille (absolue et relative à la longueur) et la profondeur. Quelques hypothèses adaptatives peuvent être proposées, comme une meilleure résistance à la pression des profondeurs ou bien une aération plus importante de la cavité palléale, permettant de mieux s'oxygéner en profondeur (MacNeil, 1960). Toutefois ces hypothèses n'ont pas plus été explorées car elles nécessitent obligatoirement la mise en place de protocoles expérimentaux lourds, comme recréer des conditions de pression et température similaires à celles des grandes profondeurs, ou bien longs et fastidieux, comme des mesures précises de l'activité des organes respiratoires chez un grand nombre de spécimens.

Les résultats de ce travail, et en particulier le nombre d'espèces relativement limité par rapport aux autres familles de Conoidea et le taux de diversification décroissant, nous indique que les Turridae sont un clade en déclin, impliquant de ce fait qu'à une époque donnée, les

Turridae étaient plus diversifiés qu'aujourd'hui. La pratique traditionnelle de considérer les Conoidea ou les Turridae *s.l.* comme un groupe hyperdiversifié devient alors toute relative. Soit nous parlons des Conoidea comme groupe hyperdiversifié actuellement, et dans ce cas nous nous trompons et nous devons reconsidérer cette affirmation à une échelle phylogénétique plus petite (par exemple, les Conidae); soit nous parlons des Conoidea comme d'un groupe hyperdiversifié à une époque donnée, c'est-à-dire à l'époque où tout le groupe suivait un taux de diversification positif – si un telle période a réellement existé. Ces considérations peuvent paraître superflus, mais elles remettent en cause certains acquis. Par exemple, il est communément accepté que la glande à venin est une innovation adaptative des Conoidea, et que cette structure est, au moins en partie, responsable du succès évolutif du groupe. A la lumière des résultats de diversification sur les Turridae, cela n'a pu être vrai que jusqu'à une époque donnée. Les traits qui expliquent les milliers d'espèces actuelles de Conidae, Raphitomidae et Mangeliidae ne sont peut-être pas du tout liés à la glande à venin.

### **C. Perspectives**

La stratégie appliquée aux Turridae semble maintenant opérationnelle, et pourrait être appliquée à des familles plus diversifiées, comme les Raphitomidae ou les Mangeliidae, qui chacune inclurait environ 2000 espèces, ou bien des groupes de rang taxonomique supérieur comme les Conoidea et les Neogastéropodes. Cependant, plusieurs points peuvent être améliorés.

Si l'acquisition des données moléculaires de Conoidea est en plein essor depuis 15 ans, l'acquisition des traits morphologiques prend du retard. Pourtant les coquilles de gastéropodes ont cette spécificité unique d'enregistrer le développement de l'animal au cours de sa vie. De plus, une meilleure formalisation de ces traits permettrait des études combinant fossiles et actuels (Sanders *et al.*, 2019). Certes, chaque spécimen séquencé est associé à une photo dans la base de données du muséum. Mais beaucoup de ces photos ont été prises à titre illustratif, manquent souvent d'une échelle et possèdent différents niveaux de résolution. Dans une tentative de tester la possibilité d'exploiter ces photos pour les Turridae, j'ai collaboré avec Jérémie Bardin (IE, CR2P) afin d'acquérir des données morphométriques sur les contours de la coquille de façon quasi-automatisée. Des études utilisant les contours de la coquille pour des études de systématique existent déjà chez les Conoidea (Puillandre *et al.*, 2009; Puillandre *et al.*, 2010), mais le processus d'acquisition des données s'est fait en plaçant

manuellement des landmarks, un procédé long et fastidieux. L'approche utilisée dans l'article 1, ainsi que dans le chapitre 3, repose sur une acquisition automatique des contours (Bonhomme *et al.*, 2014), permettant d'analyser des centaines de photos en à peine quelques heures. Cette approche a fonctionné avec un certain succès, mais l'utilisation d'objets à deux dimensions pour caractériser des objets qui en ont trois a forcément ses limites, notamment en termes de quantité de variation capturée par l'approche. Pour ne plus négliger l'aspect tridimensionnel des coquilles, un projet possible serait l'acquisition de la forme totale de la coquille par microtomographie aux rayons X ( $\mu$ CT). L'acquisition est coûteuse : actuellement, une demi-journée d'utilisation du tomographe de la plateforme AST-RX du muséum coûte 125 €. Cependant, il y aurait moyen d'automatiser le pipeline acquisition et de faire passer plusieurs spécimens en même temps afin d'en réduire les coûts par spécimen. Prenons un exemple avec les Turridae. Leur taille adulte moyenne est d'environ 5 cm et, leur coquille étant longiforme, la largeur moyenne d'une coquille est d'1 cm. En construisant un cylindre de 5 cm de diamètre sur 5 cm de hauteur, on pourrait y insérer 19 tubes cylindriques (afin d'immobiliser correctement les spécimens) de 1 cm de diamètre sur 5 cm de hauteur. La résolution de chaque coquille, en voxel, serait de 0,035 mm, de quoi obtenir suffisamment de résolution par caractériser jusqu'à la protoconque même (Monet *et al.*, 2009). De plus, en appliquant une méthode de multiscan (Bellato, Germain, Wills, pers. com.), une colonne de quatre cylindres de 5x5 cm pourrait être scannée en 4-5h, et trois de ces colonnes pourraient être scannées à la journée. Au total, 228 spécimens pourraient être scannés en 1 journée, ce qui réduirait le coût à quasiment 1,1 € par spécimen, soit un coût presque 7 fois moins élevé qu'un séquençage CO1 actuel en Sanger. Bien entendu, le processus de traitement des images est également à prendre en compte. Cependant ce type d'analyses permet en une seule fois d'acquérir les données pour estimer une quantité de traits innombrable : taille, nombre de tours, protoconque, cavité interne, paramètres de Raup (Raup, 1966), quantification de l'ornementation, etc. Autant de traits qui pourraient être utilisés à toutes les étapes d'une étude de systématique : délimitation d'espèce, attribution de types porte-noms à des hypothèses d'espèces, données ontogénétiques (Monet *et al.*, 2009), attribution de fossiles à des groupes actuels pour la datation et la phylogénie, plasticité phénotypique, description (semi-automatisée) des espèces nouvelles, identification de spécimens inconnus, etc.

Pour faire le lien entre ces données morphologiques et moléculaires, l'assemblage et l'annotation complète de génomes est nécessaire. En effet, l'acquisition de génomes permettrait de débloquent un certain nombre de questions posées parfois depuis un demi-

siècle : quels sont les mécanismes d'expression de gènes codant pour certains trait-clefs comme par exemple le type de vie larvaire ? Quel est le degré de duplication des génomes dans les groupes majeurs de gastéropodes ? Comment s'organisent dans le génome les gènes codants pour les toxines ? ... et d'en amener bien d'autres. De plus, la comparaison de génomes à différentes échelles phylogénétiques permettrait d'identifier plus facilement les zones d'intérêt, et permettrait d'améliorer ainsi l'efficacité des approches « génomes réduit » pour la délimitation d'espèces ou pour la phylogénie. Etant donné la rapidité de développement des techniques de séquençage, l'acquisition à grande échelle de génomes de gastéropodes n'est qu'une question de quelques années, à condition que nous soyons capables de les assembler correctement.

D'autres limites à l'approche développée pendant cette thèse pourraient également être évoquées. Comme pour les données morphologiques, nous manquons d'information sur les habitats où vivent les Turridae, que ce soit des paramètres physico-chimiques classiques, comme la température, la pression, la salinité, etc. mais également une description plus fine de l'environnement immédiat dans lequel évoluent les spécimens. Grâce à l'utilisation de robots sous-marins, les missions récentes organisées par mon équipe ont permis d'obtenir des images des fonds sous-marins, mais elles restent peu informatives pour les Turridae, qui vivent probablement enfouis dans le sable pour la plupart. De la même manière, nous manquons d'information sur les proies des Turridae : les approches *in situ* sont difficiles à mettre en place, et si les approches indirectes (métabarcoding) sont prometteuses, elles doivent être complétées pour obtenir des résultats fiables et proposer des hypothèses robustes.

Plus généralement, l'accès au terrain, et donc aux spécimens et à la description de leur habitats, s'est grandement compliqué depuis quelques années, notamment avec la mise en place du Protocole de Nagoya. Les résultats de ma thèse, largement dépendante du large échantillonnage disponible, illustre parfaitement l'importance de combler nos lacunes en termes de connaissance sur la biodiversité, notamment marine. De la même façon, cette thèse illustre le besoin d'une approche pluridisciplinaire, et en particulier de l'importance des sciences « traditionnelles » en systématique (anatomie comparée, morphométrie, description d'espèces, etc.), qui souffrent également d'un contexte actuel plutôt défavorable, le nombre de « collections orphelines » dans les muséums ne cessant apparemment pas de croître.

Enfin, et en attendant les génomes complets, les approches « génomes réduits » – les différentes captures d'exons effectuées dans l'équipe – ont révélé la présence de nombreuses séquences non-homologues passant toutes les étapes de filtre de notre pipeline. La pratique



actuelle a tendance à supprimer les alignements dans leur ensemble, alors que cette opération n'est pas forcément nécessaire. Le développement conjoint de méthodes « tree-based » et de métriques d'alignement permettrait de s'assurer de garder un maximum de groupes de séquences orthologues à partir d'une sortie de séquençage. La question de l'homologie des séquences est fondamentale, et les alignements multiples de séquences sont la base de toutes les études moléculaires en systématique. C'est dans l'objectif d'en apprendre plus sur ces méthodes, et si possible de les développer, que je devrai effectuer un post-doc au sein du laboratoire de Tandy Warnow, à l'Université de l'Illinois Urbana-Champaign, à partir de février 2020.

**Bibliographie**

- Abalde, S., Tenorio, M. J., Afonso, C. M., Uribe, J. E., Echeverry, A. M., & Zardoya, R. (2017). Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitochondrial genomes. *BMC evolutionary biology*, 17(1), 231.
- Abdelkrim J., Aznar-Cormano L., Fedosov A.E., Kantor Y.I., Lozouet P., Phuong M.A., Zaharias P., Puillandre N. (2018a). Exon-capture based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Mol. Biol. Evol.* 35:2355–2374.
- Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., Zaharias, P., & Puillandre, N. (2018b). Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Molecular ecology*, 27(22), 4591-4611.
- Adams, H. & Adams, A. (1853–1858) *The Genera of Recent Mollusca Arranged According to Their Organization*. John van Voorst, London.
- Adams, A. & Reeve, L. (1850) *Mollusca. Part 2*. In Adams, A. (Ed.) *The Zoology of the Voyage of the H. M. S. Samarang; under the Command of Captain Sir Edward Belcher, C.B., F.R.A.S., F.G.S., during the Years 1843–1846*. Reeve & Benham, London, pp. 25–41, pls. 10–17.
- Aguilar, M. B., de la Rosa, R. A. C., Falcón, A., Olivera, B. M., & de la Cotera, E. P. H. (2009). Peptide pal9a from the venom of the turrid snail *Polystira albida* from the Gulf of Mexico: purification, characterization, and comparison with P-conotoxin-like (framework IX) conoidean peptides. *Peptides*, 30(3), 467-476.
- Alexander, S. A., de Bruin, A., & Kornet, D. J. (2015). An alternative construction of internodons: the emergence of a multi-level tree of life. *Bulletin of mathematical biology*, 77(1), 23-45.
- Alfaro, M. E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D. L. Carnevale, G. & Harmon, L. J. (2009). Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences*, 106(32), 13410-13414.
- Arbour, J. H., & Santana, S. E. (2017). A major shift in diversification rate helps explain macroevolutionary patterns in primate species diversity. *Evolution*, 71(6), 1600-1613.
- Arumugam, M., Giji, S., Tamilmozhi, S., Kumar, S., & Balasubramanian, T. (2013). Studies on biochemical and biological properties of turrids venom (*Turricula javana* and *Lophiotoma indica*).
- 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 research*, 40(W1), W580-W584.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A.V., Sirothkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A. & Pevzner, P. A. (2012). SPAdes: a new

- genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, 19(5), 455-477.
- Bary, S. (2018). Les représentations de la biodiversité dans les fonds marins : une approche épistémologique et scientifique. Thèse de doctorat de Biologie Interdisciplinaire, Université Paris Descartes 465p.
- Batsch AJGC (1789) Versuch einer Anleitung, zur Kenntniß und Geschichte der Thiere und Mineralien, für akademische Vorlesungen entworfen, und mit den nöthigsten Abbildungen versehen. Zweyter Theil. Besondre Geschichte der Insekten, Gewürme und Mineralien. Akademische Buchhandlung, Jena
- Baum, D. A., & Shaw, K. L. (1995). Genealogical perspectives on the species problem. *Experimental and molecular approaches to plant biosystematics*, 53(289-303), 123-124.
- Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic biology*, 65(4), 583-601.
- Bekkouche, N. T. (2018). Evolution and Structure of Neuromuscular Systems in Spiralian Meiofauna. PhD at University of Copenhagen, Copenhagen
- Berry, S.S. (1910). Publications Received: Report on a Collection of Shells from Peru, with a Summary of the littoral marine Mollusca of the Peruvian Zoological Province. By William Healey Dall: *Proc. U.S.N.M.* 37, P. 147-294, Pls. 20-28, 1909. *The Nautilus*, 23(10): 130-132.
- Berry, S.S. (1920). Turritidae vs. Turridae. *The Nautilus*, 33(4): 130-133.
- Berry, S.S. (1920). Turritidae vs. Turridae. *The Nautilus*, 34(2): 57-58.
- Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., & Good, J. M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC genomics*, 13(1), 403.
- Bininda-Emonds, O. R. (2004). Trees versus characters and the supertree/supermatrix" paradox". *Systematic Biology*, 53(2), 356-359.
- Bogdanowicz, D., Giaro, K., & Wróbel, B. (2012). TreeCmp: comparison of trees in polynomial time. *Evolutionary Bioinformatics*, 8, EBO-S9657.
- Bonhomme, V., Picq, S., Gaucherel, C., & Claude, J. (2014). Momocs: outline analysis using R. *Journal of Statistical Software*, 56(13), 1-24.
- Bouchet, P., & Waren, A. (1979). Planktotrophig larval development in deep-water gastropods. *Sarsia*, 64(1-2), 37-40.
- Bouchet, P., & Kantor, Y. I. (2004). New Caledonia: the major centre of biodiversity for volutomitrid molluscs (Mollusca: Neogastropoda: Volutomitridae). *Systematics and Biodiversity*, 1(4), 467-502.
- Bouchet, P., Lozouet, P., & Sysoev, A. (2009). An inordinate fondness for turrids. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(19-20), 1724-1731.

- Bouchet, P., Kantor, Y. I., Sysoev, A., & Puillandre, N. (2011). A new operational classification of the Conoidea (Gastropoda). *Journal of Molluscan Studies*, 77(3), 273-308.
- Boxshall, G., & Self, D. (2011). UK Taxonomy & Systematics Review–2010. Results of survey undertaken by the Review Team at the Natural History Museum serving as contractors to the Natural Environment Research Council (NERC).
- Britton, T., Anderson, C. L., Jaquet, D., Lundqvist, S., & Bremer, K. (2006). PATHd8—a new method for estimating divergence times in large phylogenetic trees without a molecular clock. Available from the authors ([www.math.su.se/PATHd8](http://www.math.su.se/PATHd8)).
- Burin, G., Alencar, L. R., Chang, J., Alfaro, M. E., & Quental, T. B. (2018). How well can we estimate diversity dynamics for clades in diversity decline?. *Systematic biology*, 68(1), 47-62.
- Candolle, A. P. (1865) *Géographie botanique raisonnée*, vol II. Masson, Paris
- Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972-1973.
- Casey, T. L. (1904). Notes on the Pleurotomidae with Description of Some New Genera and Species. *Transactions of the Academy of Science of St. Louis*, 14: 123–170.
- Cecca, F., & Zaragüeta, R. Z. B. (2015). *Paléobiogéographie*. EDP sciences.
- Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*, 265-270.
- Condamine, F. L., Rolland, J., & Morlon, H. (2019). Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecology letters*, 22(11), 1900-1912.
- Gotelli, N. J., & Chao, A. (2013). Measuring and estimating species richness, species diversity, and biotic similarity from sampling data.
- Chemnitz, J. H. (1780). *Neues systematisches Conchylien-Cabinet*. Vierter Band. Mit acht und dreyßig nach der Natur gemalten und durch lebendige Farben erleuchteten Kupfertafeln. - pp. [1-27], 1-344, [1], Tab. CXXII-CLIX [= 122-159]. Nürnberg. (Raspe)
- Cossmann, M. (1896). *Essais de paléoconchologie comparée*, 2. The author and Comptoir Géologique, Paris. 179 pp., 8 pls.
- Coyne, J. A., & Orr, H. A. *Speciation*. (2004). Sunderland, MA: Sinauer Associates, 545.
- Cruz, L. J., Gray, W. R., & Olivera, B. M. (1978). Purification and properties of a myotoxin from *Conus geographus* venom. *Archives of biochemistry and biophysics*, 190(2), 539-548.
- Cunha, R. L., Castilho, R., Rüber, L., & Zardoya, R. (2005). Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde Islands. *Systematic biology*, 54(4), 634-650.

- Cunha, R. L., Assis, J. M., Madeira, C., Seabra, R., Lima, F. P., Lopes, E. P., Williams, S.T. & Castilho, R. (2017). Drivers of Cape Verde archipelagic endemism in keyhole limpets. *Scientific Reports*, 7, 41817.
- Cunha, T. J., & Giribet, G. (2019). A congruent topology for deep gastropod relationships. *Proceedings of the Royal Society B*, 286(1898), 20182776.
- Dall, W.H. (1881) Reports on the results of dredging under the supervision of Alexander Agassiz, in the Gulf of Mexico and in the Caribbean Sea, 1877–79, by the U. S. Coast Survey Steamer “Blake”, Lieutenant Commander C. D. Sigsbee, U. S. N., and Commander J. R. Bartlett, U. S. N., commanding. XV. Preliminary report on the Mollusca. *Bulletin of the Museum of Comparative Zoölogy*, 9, 33–144.
- Dall, W. H. (1906). Early history of the generic name *Fusus*. *J. Conch.* 11(10): 289-297.
- Dall, W. H. (1908). *The Mollusca and the Brachiopoda (Vol. 37)*. Harvard College, Museum of Comparative Zoology.
- Dall, W.H. (1909). Contributions to the Tertiary paleontology of the Pacific Coast. I. The Miocene of Astoria and Coos Bay, Oregon. *United States Geological Survey Professional Paper*, 59, 1–278.
- Dall, W.H. (1910). Publication Received: Note on the Summary of the Mollusca of the Peruvian Province. *The Nautilus*, 23(11): 144.
- Dall, W.H. (1918b) Notes on the nomenclature of the mollusks of the family Turritidae. *Proceedings of the United States National Museum*, 54, 313–333.
- Dall, W.H. (1920). Turritidae vs. Turridae. *The Nautilus* 34(1): 27-28.
- Darwin, C. (1859). *On the origin of species*, 1859. Murray, London
- Dautzenberg, P. & Fischer, H. (1896) Dragages effectués par l’Hirondelle et par le Princess-Alice, 1888–1895. *Mémoires de la Société Zoologique de France*, 9, 395–498, pls. 15–22.
- Davis, K. E., Hill, J., Astrop, T. I., & Wills, M. A. (2016). Global cooling as a driver of diversification in a major marine clade. *Nature communications*, 7, 13003.
- Davis, K. E., De Grave, S., Delmer, C., & Wills, M. A. (2018). Freshwater transitions and symbioses shaped the evolution and extant diversity of caridean shrimps. *Communications biology*, 1(1), 16.
- Dayrat, B. (2005). Towards integrative taxonomy. *Biological journal of the Linnean society*, 85(3), 407-417.
- De Carvalho M.R. Ebach M.C. Williams D.M. Nihei S.S. Trefaut Rodrigues M. Grant T. Silveira L.F. Zaher H. Gill A.C. Schelly R.C. Sparks J.S. Bockmann F.A. Séret B. Ho H.C. Grande L. Rieppel O. Dubois A. Ohler A. Faivovich J. Assis L.C.S. Wheeler Q.D. Goldstein P.Z. De Almeida E.A.B. Valdecasas A.G. Nelson G. (2014) . Does counting species count as taxonomy? On misrepresenting systematics, yet again . *Cladistics* 30 : 322 – 329 .

- De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. *Endless forms: species and speciation*.
- De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, 56(6), 879-886.
- de Pinna, M. C. (1999). Species concepts and phylogenetics. *Reviews in Fish biology and Fisheries*, 9(4), 353-373.
- Der Sarkissian, C., Pichereau, V., Dupont, C., Ilsøe, P. C., Perrigault, M., Butler, P., ... & Orlando, L. (2017). Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past. *Molecular ecology resources*, 17(5), 835-853.
- Dézallier d'Argenville, A. J. (1742). *L'histoire naturelle éclaircie dans deux de ses parties principales, la lithologie et la conchyliologie, dont l'une traite des pierres et l'autre des coquillages, ouvrage dans lequel on trouve une nouvelle méthode & une notice critique des principaux auteurs qui ont écrit sur ces matières [sic]. Enrichi de figures dessinées d'après nature.* - pp. [1-8], 1-491, [1], Pl. 1-33. Paris. (De Bure).
- dos Reis, M., & Yang, Z. (2019). Bayesian molecular clock dating using genome-scale datasets. In *Evolutionary Genomics* (pp. 309-330). Humana, New York, NY.
- Dubois, A., & Bour, R. (2010). The distinction between family-series and class-series nomina in zoological nomenclature, with emphasis on the nomina created by Batsch (1788, 1789) and on the higher nomenclature of turtles. *Bonn Zoological Bulletin*, 57(2), 149-171.
- Duda, T.F., (2008). Differentiation of venoms of predatory marine gastropods: divergence of orthologous toxin genes of closely related *Conus* species with different dietary specializations. *Journal of Molecular Evolution* 67, 315-321
- Duruy, V. (1865). *Introduction Generale A L'Histoire de France*. Hachette
- Dutertre, S.; Jin, A.-H.; Vetter, I.; Hamilton, B.; Sunagar, K.; Lavergne, V.; Dutertre, V.; Fry, B.G.; Antunes, A.; Venter, D.J.; Alewood P.F., Lewis R. J. (2014) Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nat. Commun.*, 5, 3521
- Economio, E. P., Narula, N., Friedman, N. R., Weiser, M. D., & Guénard, B. (2018). Macroecology and macroevolution of the latitudinal diversity gradient in ants. *Nature communications*, 9(1), 1778.
- Espeland, M., Hall, J. P., DeVries, P. J., Lees, D. C., Cornwall, M., Hsu, Y. F., Wu, L.-W., Campbell, D. L., Talavera, G., Vila, R., Salzman, S., Ruehr, S., Lohman, D. J., Pierce, N. E. & Salzman, S. (2015). Ancient Neotropical origin and recent recolonisation: Phylogeny, biogeography and diversification of the Riodinidae (Lepidoptera: Papilionoidea). *Molecular phylogenetics and evolution*, 93, 296-306.
- Espeland M. , Breinholt J. , Wilmott K.R. , Warren A.D. , Vila R. , Toussaint E.F.A. , Maunsell S.A. , Aduse-Poku K. , Tatavera G. , Eastwood R. , Jarzyna M.A. , Guralnick R. , Lohman D.J. , Pierce N.E. , Kawahara A.V. (2018) . A comprehensive and dated phylogenomic analysis of butterflies. *Curr. Biol.* 28:1 –9.

- Faircloth, B. C., & Glenn, T. C. (2014). Protocol: preparation of an AMPure XP substitute (AKA Serapure). Web document doi, 10, J9MW2F26.
- Fassio, G., Modica, M. V., Mary, L., Zaharias, P., Fedosov, A. E., Gorson, J., Kantor, Y. I., Holford, M. & Puillandre, N. (2019). Venom Diversity and Evolution in the Most Divergent Cone Snail Genus *Profundiconus*. *Toxins*, 11(11), 623.
- Fedosov, A., Watkins, M., Heralde III, F. M., Corneli, P. S., Concepcion, G. P., & Olivera, B. M. (2011). Phylogeny of the genus *Turris*: Correlating molecular data with radular anatomy and shell morphology. *Molecular phylogenetics and evolution*, 59(2), 263-270.
- Fedosov, A. E., & Puillandre, N. (2012). Phylogeny and taxonomy of the *Kermia–Pseudodaphnella* (Mollusca: Gastropoda: Raphitomidae) genus complex: a remarkable radiation via diversification of larval development. *Systematics and Biodiversity*, 10(4), 447-477.
- Felsenstein, J., & Felsenstein, J. (2004). *Inferring phylogenies* (Vol. 2, p. 664). Sunderland, MA: Sinauer associates.
- Feng, Y. J., Blackburn, D. C., Liang, D., Hillis, D. M., Wake, D. B., Cannatella, D. C., & Zhang, P. (2017). Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. *Proceedings of the National Academy of Sciences*, 114(29), E5864-E5870.
- Fernández, R., Kallal, R. J., Dimitrov, D., Ballesteros, J. A., Arnedo, M. A., Giribet, G., & Hormiga, G. (2018). Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. *Current Biology*, 28(9), 1489-1497.
- FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, 3(6), 1084-1092.
- Fitzpatrick, B. M., Fordyce, J. A., & Gavrillets, S. (2008). What, if anything, is sympatric speciation?. *Journal of evolutionary biology*, 21(6), 1452-1459.
- Foerstner, K. U., Von Mering, C., Hooper, S. D., & Bork, P. (2005). Environments shape the nucleotide composition of genomes. *EMBO reports*, 6(12), 1208-1213.
- Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2016). Biodiversity and the species concept—lineages are not enough. *Systematic Biology*, 66(4), 644-656.
- Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in ecology & evolution*, 27(9), 480-488.
- García-López, A. L., Uria-Galicia, E. A., & Ortiz-Ordoñez, E. (2007). Estudio histológico de boca, esófago, estómago e intestino de *Polystira albida* (Gastropoda: Turridae). *International Journal of Morphology*, 25(4), 767-774.
- Gatesy, J., Matthee, C., DeSalle, R., & Hayashi, C. (2002). Resolution of a supertree/supermatrix paradox. *Systematic Biology*, 51(4), 652-664.



- Geoffroy Saint-Hilaire, I. (1859). *Histoire naturelle générale des règnes organiques*, vol. II. Masson, Paris
- Giardina, C. R., & Kuhl, F. P. (1977). Accuracy of curve approximation by harmonically related vectors with elliptical loci. *Computer Graphics and Image Processing*, 6(3), 277-285.
- Giribet, G., DeSalle, R., & Wheeler, W. C. (2002). 'Pluralism' and the aims of phylogenetic research. In *Molecular systematics and evolution: theory and practice* (pp. 141-146). Birkhäuser, Basel.
- Goldstein, P. Z., & DeSalle, R. (2011). Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *Bioessays*, 33(2), 135-147.
- Gonzales, D. T. T., & Saloma, C. P. (2014). A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. *Toxicon*, 92, 66-74.
- Gorson, J. (2018). *Characterizing Venom Gene Expression, Function and Species Diversity in Predatory Marine Snails of the Terebridae*. PhD at Hunter College, New-York
- Gualtieri, N. (1742). *Index testarvm conchyliorvm quae adservantvr in mvseo Nicolai Gualtieri (...) et methodice distribvta exhibentvr. Tabvlis CX. - pp. I-XXIII [= 1-23], I-CX [= 1-110], Tab. I-CX [= 1-110]. Florentinae. (Albizzini).*
- Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2007). GEIGER: investigating evolutionary radiations. *Bioinformatics*, 24(1), 129-131.
- Harvey, M. G., & Rabosky, D. L. (2018). Continuous traits and speciation rates: Alternatives to state-dependent diversification models. *Methods in Ecology and Evolution*, 9(4), 984-993.
- Hebenstreit, J. E. (1743). *Museum Richterianum: continens fossilia animalia, vegetabilia mar. Lipsiae: Caspar Fritsch.*
- Hebert, P. D., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313-321.
- Hedley, C. (1922) A revision of the Australian Turridae. *Records of the Australian Museum*, 13, 213-359, pls. 42-56.
- Helbling, G. S. 1779. *Beyträge zur Kenntniß neuer und seltener Konchylien. Aus einigen Wienerischen Sammlungen. - Abhandlungen einer Privatgesellschaft in Böhmen, zur Aufnahme der Mathematik, der vaterländischen Geschichte, und der Naturgeschichte* 4: 102-131, Tab. I-IV [= 1-4]. Prag.
- Hennig, W., (1966). *Phylogenetic systematics*. University of Illinois Press.
- Heralde III, F. M., Watkins, M., Ownby, J. P., Bandyopadhyay, P. K., Santos, A. D., Concepcion, G. P., & Olivera, B. M. (2007). Molecular phylogeny of some Indo-Pacific genera in the subfamily Turrinae H. Adams and A. Adams, 1853 (1838)(Gastropoda: Neogastropoda). *Nautilus*, 121(3), 131-138.

- Heralde III, F. M., Imperial, J., Bandyopadhyay, P. K., Olivera, B. M., Concepcion, G. P., & Santos, A. D. (2008). A rapidly diverging superfamily of peptide toxins in venomous *Gemmula* species. *Toxicon*, 51(5), 890-897.
- Heralde III, F. M., Kantor, Y. I., Astilla, M. A. Q., Lluisma, A. O., Geronimo, R., Alino, P. M., Watkins, M., Corneli, P. S., Olivera, B. M., Santos, A., Concepcion, G. (2010). The Indo-Pacific *Gemmula* species in the subfamily Turridae: Aspects of field distribution, molecular phylogeny, radular anatomy and feeding ecology.
- Herschel to Lyell 20 Feb. 1836, later published in C. Babbage, Ninth Bridgewater treatise (2nd edn) (Murray, London, 1838), pp. 226–227
- Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518-522.
- Holford, M., Puillandre, N., Terryn, Y., Cruaud, C., Olivera, B., & Bouchet, P. (2008). Evolution of the *Toxoglossa* venom apparatus as inferred by molecular phylogeny of the Terebridae. *Molecular Biology and Evolution*, 26(1), 15-25.
- Huang, X., & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome research*, 9(9), 868-877.
- Hull, D., L., (1988). *Science as a process*. University of Chicago Press
- Humphrey, G., (1797). *Museum Calonnianum*. Specification of the various articles which compose the magnificent museum of natural history collected by M. De Calonne in France (...). - pp. i-viii [= 1-8], 1-84. London.
- Igea, J., Miller, E. F., Papadopulos, A. S., & Tanentzap, A. J. (2017). Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLoS biology*, 15(7), e2002792.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, 491(7424), 444.
- Jukes-Browne, J. (1904). On Some Questions of Nomenclature. *J. Conch.* 11(4):97-103
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6), 587.
- Kantor, Y. I., & Sysoev, A. V. (1991). Sexual dimorphism in the apertural notch of a new species of *Gemmula* (Gastropoda: Turridae). *Journal of molluscan studies*, 57(2), 205-209.
- Kantor, Y. I. (2006). On the morphology and homology of the “central tooth” in the radulae of Turridae (Conoidea: Turridae). *Ruthenica*, 16(1-2), 47-52.
- Kantor, Y. I., Puillandre, N., Olivera, B. M., & Bouchet, P. (2008). Morphological proxies for taxonomic decision in turrids (Mollusca, Neogastropoda): a test of the value of shell and radula characters using molecular data. *Zoological science*, 25(11), 1156-1171.

- Kantor, Y. I., & Puillandre, N. (2012). Evolution of the radular apparatus in Conoidea (Gastropoda: Neogastropoda) as inferred from a molecular phylogeny. *Malacologia*, 55(1), 55-91.
- Kantor, Y. I., Fedosov, A. E., & Puillandre, N. (2018). New and unusual deep-water Conoidea revised with shell, radula and DNA characters. *Ruthenica*, 28(2).
- Katoh, K., Kuma, K. I., Miyata, T., & Toh, H. (2005). Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome informatics*, 16(1), 22-33.
- Kekkonen, M., & Hebert, P. D. (2014). DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Molecular ecology resources*, 14(4), 706-715.
- Kilburn, R. N. (1983). Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 1. Subfamily Turrinae. *Annals of the Natal Museum*, 25(2).
- Kilburn, R. N., Fedosov, A. E., & Olivera, B. M. (2012). Revision of the genus *Turris* (Gastropoda: Conoidea: Turridae) with the description of six new species. *Zootaxa*, 3244(1), 1.
- Klein, J. T. (1753). *Tentamen methodi ostracologicæ sive dispositio naturalis cochlidum et concharum in suas classes, genera et species, iconibus singulorum generum aeri incisus illustrata. Accedit lucubratiuncula de formatione, cremento et coloribus testarum quæ sunt cochlidum et concharum. Tum commentariolum in locum Plinii Hist. Nat. Libr. IX cap. XXXIII, de concharum differentiis. Denique sciagraphia methodi ad genus serpentium ordinate digerendum.* - pp. [1-9], 1-177, [1-27], [1-8], 1-44, 1-16, [1], Tab. I-XII [= 1-12]. Lugduni Batavorum. (Wishoff).
- Kohn, A. J., Saunders, P. T., & Wiener, S. (1960). Preliminary studies on the venom of the marine snail *Conus*. *Annals of the New York Academy of Sciences*, 90(3), 706-725.
- Kohn, A. J. (1990). Tempo and mode of evolution in Conidae. *Malacologia*, 32(1), 55-67.
- Kohn, A. J. (2016). Human injuries and fatalities due to venomous marine snails of the family Conidae. *International journal of clinical pharmacology and therapeutics*, 54(7), 524.
- Kornet, D. J. (1993). Permanent splits as speciation events: a formal reconstruction of the internodal species concept. *Journal of Theoretical Biology*, 164(4), 407-435.
- Kozak, K. M., Wahlberg, N., Neild, A. F., Dasmahapatra, K. K., Mallet, J., & Jiggins, C. D. (2015). Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies. *Systematic biology*, 64(3), 505-524.
- Kraus, N. J., Corneli, P. S., Watkins, M., Bandyopadhyay, P. K., Seger, J., & Olivera, B. M. (2011). Against expectation: A short sequence with high signal elucidates cone snail phylogeny. *Molecular phylogenetics and evolution*, 58(2), 383-389.
- Kraus, N. J., Watkins, M., Bandyopadhyay, P. K., Seger, J., Olivera, B. M., & Corneli, P. S. (2012). A very short, functionally constrained sequence diagnoses cone snails in several *Conasprella* clades. *Molecular phylogenetics and evolution*, 65(1), 335-338.
- Krug, P. J., Vendetti, J. E., Ellingson, R. A., Trowbridge, C. D., Hirano, Y. M., Trathen, D. Y., Rodriguez, A. K., Swennen, C., Wilson, N. G. & Valdés, Á. A. (2015). Species

- selection favors dispersive life histories in sea slugs, but higher per-offspring investment drives shifts to short-lived larvae. *Systematic Biology*, 64(6), 983-999.
- Lagomarsino, L. P., Condamine, F. L., Antonelli, A., Mulch, A., & Davis, C. C. (2016). The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist*, 210(4), 1430-1442.
- Landis, J. B., Soltis, D. E., Li, Z., Marx, H. E., Barker, M. S., Tank, D. C., & Soltis, P. S. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *American journal of botany*, 105(3), 348-363.
- Lange, Carl Nicolaus (1722). *Methodus nova & facilis testacea marina pleraque* available online at: <http://gdz.sub.unigoettingen.de/en/index.html>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), 357.
- Leal, J. H., Kohn, A. J., & Mensch, R. (2017). A Veliconcha Unveiled: Observations on the Larva and Radula of *Conus spurius*, with Implications for the Origin of Molluscivory in *Conus*. *American malacological bulletin*, 35(2), 111-119.
- Legendre, F., & Condamine, F. L. (2018). When Darwin's Special Difficulty Promotes Diversification in Insects. *Systematic biology*, 67(5), 873-887.
- 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.
- Li, J., Ó Foighil, D., & Strong, E. E. (2016). Commensal associations and benthic habitats shape macroevolution of the bivalve clade Galeommatoidea. *Proceedings of the Royal Society B: Biological Sciences*, 283(1834), 20161006.
- Linnæus, C. (1758). *Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata.* - pp. [1-4], 1-824. Holmiæ. (Salvius).
- Lister, M., Lister, S. & Lister, A. (1688). *Historiæ [sive synopsis methodicæ] conchyliorum [quorum omnium picturæ, ad vivum delineatæ, exhibētur.] Liber IV. qui est de buccinis marinis etiam vermiculj, dentalia, et patellæ numerantur.* ibidem. - pp. [1], Pl. [1-238]. Londini.
- Locard, A. (1897) *Expéditions Scientifiques du Travailleur et du Talisman. Mollusques testacés.* Vol. 1. Masson, Paris, 516 pp., 22 pls.
- López-Vera, E., De La Cotera, E. P. H., Maillo, M., Riesgo-Escovar, J. R., Olivera, B. M., & Aguilar, M. B. (2004). A novel structural class of toxins: the methionine-rich peptides from the venoms of turrid marine snails (Mollusca, Conoidea). *Toxicon*, 43(4), 365-374.
- Lorion, J., Buge, B., Cruaud, C., & Samadi, S. (2010). New insights into diversity and evolution of deep-sea Mytilidae (Mollusca: Bivalvia). *Molecular Phylogenetics and Evolution*, 57(1), 71-83.

- Machac, A. (2014). Detecting trait-dependent diversification under diversification slowdowns. *Evolutionary Biology*, 41(2), 201-211.
- MacNeil, F. S. (1960). Tertiary and Quaternary Gastropoda of Okinawa: US Survey Prof.
- Maddison, W. P., & FitzJohn, R. G. (2014). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic biology*, 64(1), 127-136.
- Maes, V. O. (1971). Evolution of the toxoglossate radula and methods of envenomation. *Annu. Rep. Am. Malacol. Union for*, 1970, 69-72.
- Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27(21), 2957-2963.
- Malmstrøm, M., Matschiner, M., Tørresen, O. K., Star, B., Snipen, L. G., Hansen, T. F., Baalsrud, H. T., Nederbragt, A. J., Hanel, R., Salzburger, W., Stenseth, N.C., Joakobsen, K. S., & Jentoft, S. (2016). Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, 48(10), 1204.
- Mary, L. (2018) Recherche de toxines dans les transcriptomes de glandes à venin de Conoidea (Gastropoda). Rapport de M2, Master SEP-SDUEE 2017-2018
- Mayden, R. L. (1997). A hierarchy of species concepts: the denouement in the saga of the species problem.
- Mayr, E. (1942). *Systematics and the Origin of Species.*, (Columbia University Press: New York.).
- McLean, J.H., (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, A. I., & Sysoev, A. V. (2001). The foregut anatomy of the genus *Xenuroturris* (Gastropoda, Conoidea, Turridae), with a description of a new genus. *Ruthenica*, 11(1), 7-14.
- Medinskaya, A. I. (2002a). Foregut anatomy of the Turrinae (Gastropoda, Conoidea, Turridae). *Ruthenica*, 12, 135-159.
- Medinskaya, A. I. (2002b). Structure of the venom gland-muscular bulb complex in the family Turridae (Gastropoda, Conoidea). *Ruthenica*, 12(2), 125-133.
- Meek, F.B. (1876) A report on the invertebrate Cretaceous and Tertiary fossils of the upper Missouri country. In Hayden, F.V. (Ed.) Report of the United States Geological Survey of the Territories. *Invertebrate Paleontology*, 9, lxiv + 1–629, 45 pls.
- Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), pdb-prot5448.
- Miller BA. (1970). Studies on the biology of Indo-Pacific Terebra. PhD diss., Univ. N.H., Durham, NH

## Bibliographie

- Mirarab, S., Bayzid, M. S., Boussau, B., & Warnow, T. (2014). Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science*, 346(6215), 1250463.
- Moen, D., & Morlon, H. (2014). Why does diversification slow down?. *Trends in Ecology & Evolution*, 29(4), 190-197.
- Monaghan MT , Wild R , Elliot M , Fujisawa T , Balke M , Inward DJG , Lees DC , Ranaivosolo R , Eggleton P , Barraclough TG , Vogler AP. (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation, *Syst. Biol.* vol. 58 (pg. 298 -311)
- Monnet, C., Zollikofer, C., Bucher, H., & Goudemand, N. (2009). Three-dimensional morphometric ontogeny of mollusc shells by micro-computed tomography and geometric analysis. *Paleontologia Electronica*, 12(3/12A), 1-13.
- Moore, B. R., Höhna, S., May, M. R., Rannala, B., & Huelsenbeck, J. P. (2016). Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences*, 113(34), 9569-9574.
- Morlon, H., Parsons, T. L., & Plotkin, J. B. (2011). Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*, 108(39), 16327-16332.
- Morlon, H. (2014). Phylogenetic approaches for studying diversification. *Ecology letters*, 17(4), 508-525.
- Morlon, H., Lewitus, E., Condamine, F. L., Manceau, M., Clavel, J., & Drury, J. (2016). RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods in Ecology and Evolution*, 7(5), 589-597.
- Morrison, J. P. E., (1966) On the families of Turridae. The American Malacological Union. Annual Reports, for 1965: 1–2.
- Moulines, C. U. (1998). Esbozo de ontoepistemosemántica (Sketch of Ontoepistemosemantics). *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia*, 13(1), 141-159.
- Nelson, G., & Platnick, N. (1981). *Systematic and biogeography*. Columbia Univ. Press, New York.
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, 32(1), 268-274.
- O'Hara, T. D., Hugall, A. F., Woolley, S. N., Bribiesca-Contreras, G., & Bax, N. J. (2019). Contrasting processes drive ophiuroid phylodiversity across shallow and deep seafloors. *Nature*, 565(7741), 636.
- Olivera, B. M., Hillyard, D. R., & Watkins, M. (2008). A new species of *Gemmula*, Weinkauff 1875; Evidence of two clades of Philippine species in the genus *Gemmula*. *Philipp Sci Lett*, 11, 11-15.

- Olivera, B. M., Showers Corneli, P., Watkins, M., & Fedosov, A. (2014). Biodiversity of cone snails and other venomous marine gastropods: evolutionary success through neuropharmacology. *Annu. Rev. Anim. Biosci.*, 2(1), 487-513.
- Olivera, B. M., Seger, J., Horvath, M. P., & Fedosov, A. E. (2015). Prey-capture strategies of fish-hunting cone snails: behavior, neurobiology and evolution. *Brain, behavior and evolution*, 86(1), 58-74.
- Olivera, B. M., Fedosov, A., Imperial, J. S., & Kantor, Y. (2017). Physiology of envenomation by conoidean gastropods. In *Physiology of Molluscs* (pp. 169-204). Apple Academic Press.
- Omaga, C. A. Carpio, L. D. Imperial, J. S. Daly, N. L. Gajewiak, J. Flores, M. S. Espino, S. S. Christensen, S. Filchakova, O. M. López-Vera, E. Raghuraman, S. Olivera B. M. and Concepcion, G. P. (2017) *Biochemistry* 56, 6051–6060
- Orme, D., Freckleton, R., Thomas, G., & Petzoldt, T. (2013). The caper package: comparative analysis of phylogenetics and evolution in R. *R package version*, 5(2), 1-36.
- Page, L. R., & Hookham, B. (2017). The gastropod foregut—evolution viewed through a developmental lens. *Canadian Journal of Zoology*, 95(4), 227-238.
- Palumbi, S. R., & Baker, C. S. (1994). Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, 11(3), 426-435.
- Pante, E., Schoelinck, C., & Puillandre, N. (2014). From integrative taxonomy to species description: one step beyond. *Systematic Biology*, 64(1), 152-160.
- Penchaszadeh, P. E. (1982). Reproductive aspects of *Polystira barrette* (Guppy, 1866) (Gastropoda, Turridae) from Golfo Triste, Venezuela. *Veliger* 25(2), 160.
- Pérez-Escobar, O. A., Cass, S., & Dodsworth, S. (2018). Digest: Drivers of coral diversification in a major marine biodiversity hotspot. *Evolution*.
- Perron, F. E. (1981). Larval growth and metamorphosis of *Conus* (Gastropoda: Toxoglossa) in Hawaii.
- Peters, H., O'Leary, B. C., Hawkins, J. P., Carpenter, K. E., & Roberts, C. M. (2013). *Conus*: First comprehensive conservation Red List assessment of a marine gastropod mollusc genus. *PLoS One*, 8(12), e83353.
- Phuong, M. A., & Mahardika, G. N. (2018). Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. *Molecular biology and evolution*, 35(5), 1210-1224.
- Phuong, M. A., Alfaro, M. E., Mahardika, G. N., Marwoto, R. M., Prabowo, R. E., von Rintelen, T., Vogt, P. W. H., Hendricks, J. R. & Puillandre, N. (2019). Lack of signal for the impact of conotoxin gene diversity on speciation rates in cone snails. *Systematic biology*.

- Postaire, B., Bruggemann, J. H., Magalon, H., & Faure, B. (2014). Evolutionary dynamics in the Southwest Indian Ocean marine biodiversity hotspot: a perspective from the rocky shore gastropod genus *Nerita*. *PloS one*, 9(4), e95040.
- Powell, A.W.B. (1942) The New Zealand Recent and fossil Mollusca of the family Turridae with general notes on turrid nomenclature and systematics. *Bulletin of the Auckland Institute and Museum*, 2, 1–188, 14 pls.
- Powell, A. W. B. (1964) The family Turridae in the Indo-Pacific. Part 1. The subfamily Turrinae. *Indo-Pacific Mollusca*, 1(5), 227–346, pls 172–262.
- Powell, A. W. B. (1967) The family Turridae in the Indo-Pacific, Part 1a. The Turrinae concluded. *Indo-Pacific Mollusca* ;1:409–444
- Puillandre, N., Samadi, S., Boisselier, M. C., Sysoev, A. V., Kantor, Y. I., Cruaud, C., Couloux, A., & Bouchet, P. (2008). Starting to unravel the toxoglossan knot: molecular phylogeny of the “turrids” (Neogastropoda: Conoidea). *Molecular Phylogenetics and Evolution*, 47: 1122–1134.
- Puillandre, N. (2008). Taxonomie intégrative des Turridae: phylogénie, délimitation d'espèces et barcoding. *Systématique Evolutive*. Thèse de doctorat du Muséum. Paris Muséum national D'Histoire naturelle, 229p.
- Puillandre, N., Baylac, M., Boisselier, M. C., Cruaud, C., & Samadi, S. (2009). An integrative approach to species delimitation in *Benthomangelia* (Mollusca: Conoidea). *Biological Journal of the Linnean Society*, 96(3), 696-708.
- Puillandre, N., & Holford, M. (2010). The Terebridae and teretoxins: Combining phylogeny and anatomy for concerted discovery of bioactive compounds. *BMC Chemical Biology*, 10(1), 7.
- Puillandre, N., Cruaud, C., & Kantor, Y. I. (2010). Cryptic species in *Gemmuloborsonia* (Gastropoda: Conoidea). *Journal of Molluscan Studies*, 76(1), 11-23.
- Puillandre, N. (2012). La «Révolution» ADN. *Techniques & Culture*. Revue semestrielle d'anthropologie des techniques, (59), 338-345.
- Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012a). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular ecology*, 21(8), 1864-1877.
- Puillandre, N., Modica, M. V., Zhang, Y., Sirovich, L., Boisselier, M. C., Cruaud, C., Holford, M., & Samadi, S. (2012b). Large-scale species delimitation method for hyperdiverse groups. *Molecular ecology*, 21(11), 2671-2691.
- Puillandre, N., Koua, D., Favreau, P., Olivera, B. M., & Stöcklin, R. (2012c). Molecular phylogeny, classification and evolution of conopeptides. *Journal of molecular evolution*, 74(5-6), 297-309.
- Puillandre, N., Fedosov, A. E., Zaharias, P., Aznar-Cormano, L., & Kantor, Y. I. (2017). A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess. *Zoological Journal of the Linnean Society*, 181(2), 243-271.



- Puttick, M. N., Clark, J., & Donoghue, P. C. (2015). Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proceedings of the Royal Society B: Biological Sciences*, 282(1820), 20152289.
- Rabosky, D. L. (2006). LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from
- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS one*, 9(2), e89543.
- Rabosky, D. L., Santini, F., Eastman, J., Smith, S. A., Sidlauskas, B., Chang, J., & Alfaro, M. E. (2013). Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature communications*, 4, 1958.molecular phylogenies. *Evolutionary Bioinformatics*, 2, 117693430600200024.
- Rabosky, D. L., Grudler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W. & Larson, J. G. (2014). BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5(7), 701-707.
- Rabosky, D. L., & Goldberg, E. E. (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic biology*, 64(2), 340-355.
- Rabosky, D. L., Mitchell, J. S., & Chang, J. (2017). Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic biology*, 66(4), 477-498.
- Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). Tracer 1.6. URL: <http://beast.bio.ed.ac.uk/tracer>.
- Raup, D. M. (1966). Geometric analysis of shell coiling: general problems. *Journal of Paleontology*, 1178-1190.
- Reeve, L.A. (1843–1846) Monograph of the Genus *Pleurotoma*. *Conchologia Iconica, or Illustrations of the Shells of Molluscous Animals*. Reeve Brothers, London, Vol. 1, 40 pls. + index and errata.
- Robinson, D. F., & Foulds, L. R. (1981). Comparison of phylogenetic trees. *Mathematical biosciences*, 53(1-2), 131-147.
- Röding, P. F. 1798. *Museum Boltenianum sive catalogus cimeliorum e tribus regnis naturæ quæ olim collegerat Joa. Fried Bolten, M. D. p. d. per XL. annos proto physicus Hamburgensis. Pars secunda continens conchyliæ sive testacea univalvia, bivalvia & multivalvia.* - pp. [1-3], [1-8], 1-199. Hamburgi. (Trapp).
- Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*, 22(5), 939-946.
- Ronai, I., & Shimek, I. (1984). The biology of the Northeastern Pacific Turridae. IV. Shell morphology and sexual dimorphism in *Aforia circinata* (Dall, 1873). *Veliger*, 26(4), 258-263.
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient

- Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, 61(3), 539-542.
- Rose, J. P., Kleist, T. J., Löffstrand, S. D., Drew, B. T., Schönenberger, J., & Sytsma, K. J. (2018). Phylogeny, historical biogeography, and diversification of angiosperm order Ericales suggest ancient Neotropical and East Asian connections. *Molecular phylogenetics and evolution*, 122, 59-79.
- Rosenberg, G. (1997). Reproducibility of results in phylogenetic analysis of mollusks: a reanalysis of the Taylor, Kantor, and Sysoev (1993) data set for conoidean gastropods. *American Malacological Bulletin*, 14, 219-228.
- Rumphius, G. E. 1711. *Thesaurus imaginum piscium testaceorum; quales sunt cancri, echini, echinometra, stellæ marinæ, &c. ut & cochlearum; inter quas numerantur lunares; lacinatae; trochi; valvatae, sive semilunares; valvatae striatae; cassides tuberosae, verrucosae, laeves & murices; globosae; buccina; strombi; volutae; alatae; porcellanae majores & minores; cylindri, &c. Quibus accedunt conchylia, ut nautilus, cornu ammonis, &c. Conchae univalviae & bivalviae; quarum species sunt solenes univalvii, Chamæ asperae, Chamæ laeves, pectines, pectunculi, tellinae, solenes bivalvii, musculi, pinnæ, ostrea, &c. Denique mineralia uti metalla, lapides & argillae, variis in locis reperta.* - pp. [1-15], LXI-LXVIII [= 61-68], Tab. I-LX [= 1-60]. Lugduni Batavorum. (Vander Aa).
- Samadi, S., & Barberousse, A. (2006). The tree, the network, and the species. *Biological Journal of the Linnean Society*, 89(3), 509-521.
- Sanders, M. T., Merle, D., Bouchet, P., Castelin, M., Beu, A. G., Samadi, S., & Puillandre, N. (2017). One for each ocean: revision of the *Bursa granularis* species complex (Gastropoda: Tonnoidea: Bursidae). *Journal of Molluscan Studies*, 83(4), 384-398.
- Sanders, M. T., Merle, D., & Puillandre, N. (2019). A review of fossil Bursidae and their use for phylogeny calibration. *Geodiversitas*, 41(1), 247-265.
- Santibañez Aguascalientes, N. A., Ortíz Ordóñez, E., Falcón Alcántara, A., & Heimer de la Cotera, E. P. (2013). Estudio histológico del tubo digestivo y aparato venenoso de *Gemmula periscelida* (Gastropoda: Turridae). *International Journal of Morphology*, 31(1), 7-14.
- Sayyari, E., Whitfield, J. B., & Mirarab, S. (2017). Fragmentary gene sequences negatively impact gene tree and species tree reconstruction. *Molecular biology and evolution*, 34(12), 3279-3291.
- Schepman, M.M. (1913) The Prosobranchia of the Siboga Expedition. Part 5. Toxoglossa. *Resultats Siboga-Expeditie*, 49-1, 365-452, pls. 25-30.
- Sheridan, R., Van Mol, J. J., & Bouillon, J. (1973). Étude morphologique du tube digestif de quelques Turridae (Mollusca-Gastropoda-Prosobranchia-Toxoglossa) de la région de Roscoff. *Cahiers de Biologie marine*, 14, 159-188.
- Shimek, R. L. & Kohn, A. J. (1981). Functional morphology and evolution of the toxoglossan radula. *Malacologia* 20: 423±438.

- Shi, J. J., & Rabosky, D. L. (2015). Speciation dynamics during the global radiation of extant bats. *Evolution*, 69(6), 1528-1545.
- Seronay, R. A., Muallil, R. N., & Alino, P. M. (2014). Turrid Fishery in Central Visayas, Philippines. *Asian Fisheries Science*, 27(1), 30-44.
- Seronay, R. A., Ali, P. M., & Olivera, B. M. (2016). Turrids in the Outer Barrier Reef, Northeast Bohol, Philippines from Lumun-Lumun Nets. *Annals of Studies in Science and Humanities*, 1(1), 95-110.
- Shen, X. X., Hittinger, C. T., & Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology & Evolution*, 1(5), 0126.
- Simakov O., Marletaz F., Cho S.-J., Edsinger-Gonzales E., Havlak P., Hellsten U., Kuo D.-H., Larsson T., Lv J., Arendt D., Savage R., Osoegawa K., de Jong P., Grimwood J., Chapman J.A., Shapiro H., Aerts A., Otiillar R.P., Terry A.Y., Boore J.L., Grigoriev I. V., Lindberg D.R., Seaver E.C., Weisblat D.A., Putnam N.H., Rokhsar D.S. 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* 493:526–531
- Siu-Ting, K., Torres-Sánchez, M., San Mauro, D., Wilcockson, D., Wilkinson, M., Pisani, D., O’Connell, M. J. & Creevey, C. J. (2019). Inadvertent Paralog Inclusion Drives Artifactual Topologies and Timetree Estimates in Phylogenomics. *Molecular biology and evolution*, 36(6), 1344-1356.
- Smit, A. F. A., Hubley, R., & Green, P. (2015). RepeatMasker Open-4.0. 2013–2015.
- Smith, E.A. (1877). Diagnoses of new species of Pleurotomidae in the British Museum. *Annals and Magazine of Natural History*, series 4, 19, 488–501.
- Smith, E.A. (1882a) Diagnoses of new species of Pleurotomidae in the British Museum. *Annals and Magazine of Natural History*, series 5, 10, 206–218.
- Smith, E.A. (1882b) Diagnoses of new species of Pleurotomidae in the British Museum. *Annals and Magazine of Natural History*, series 5, 10, 296–306.
- Smith, E.A. (1884) Diagnoses of new species of Pleurotomidae in the British Museum. *Annals and Magazine of Natural History*, series 5, 14, 317–329.
- Sohl, N. F. (1969). The fossil record of shell boring by snails. *American Zoologist*, 9(3), 725-734.
- Sollas, I. B. J. (1907). The Molluscan Radula. *Quar. Jour. Micros. Sci.*, 51: 115-136
- Sowerby, G.B., I. (1834). Characters of new species of Mollusca and Conchifera. *Proceedings of the Zoological Society of London*, 1833, 134–139.
- Su, M., Liu, H., Liang, X., Gui, L., & Zhang, J. (2018). Dietary analysis of marine fish species: enhancing the detection of prey-specific DNA sequences via high-throughput sequencing using blocking primers. *Estuaries and coasts*, 41(2), 560-571.
- Stein, L. D. (2010). The case for cloud computing in genome informatics. *Genome biology*, 11(5), 207.

- Steppan, S. J., & Schenk, J. J. (2017). Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. *PLoS One*, 12(8), e0183070.
- Statius Müller, P. L. (1766). *Deliciae Naturae Selectae oder auserlesenes Naturalien-Cabinet welches aus den drey Reichen der Natur zeigt, was von curiösen Liebhabern aufbehalten und gesammelt zu werden verdient*. Ehemals herausgegeben von Georg Wolfgang Knorr berühmten Kupferstecher in Nürnberg, fortgesetzt von dessen Erben, beschrieben von Phillip Ludwig Statius Müller (...), und in das Französische übersetzt von Matthäus Verdier de la Blaquiere (...). - pp. [1-26], 1-132, Pl. A-E5 [= 1-39]. Nürnberg.
- Streicher, J. W., Schulte, J. A., & Wiens, J. J. (2015). How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. *Systematic Biology*, 65(1), 128-145.
- Streicher, J. W., Miller, E. C., Guerrero, P. C., Correa, C., Ortiz, J. C., Crawford, A. J., & Wiens, J. J. (2018). Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. *Molecular phylogenetics and evolution*, 119, 128-143.
- Sysoev, A. V. (1991). Preliminary analysis of the relationship between turrids (Gastropoda, Toxoglossa, Turridae) with different types of radular apparatus in various Recent and fossil faunas. *Ruthenica*, 1, 53-66.
- Taylor, J. D. "Diets and habitats of shallow water predatory gastropods around Tolo Channel, Hong Kong." *The malacofauna of Hong Kong and southern China* (1980): 163-180.
- Taylor, J. D., Kantor, Y.I., Sysoev, A.V. (1993). Foregut anatomy, feeding mechanisms, relationships and classification of the Conoidea (= Toxoglossa)(Gastropoda). *Bull. nat. Hist. Mus., London (Zoology)*, 59(2), 125-170.
- Taylor, J. D. (1994). Foregut anatomy of the larger species of Turrinae, Clavatulinae and Crassispirinae (Gastropoda: Conoidea) from Hong Kong. In *Third International Workshop on the Malacofauna of Hong-Kong and Southern China*.
- Thiele, J. (1925) *Gastropoda der Deutschen Tiefsee-Expedition. II. Teil. Wissenschaftliche Ergebnisse Deutschen Tiefsee-Expedition auf dem Dampfer "Valdivia" 1898–1899*, 17, 1(35)– 348(382), pls. 1(13)–34(46).
- Thiele, J. (1929) *Handbuch der Systematischen Weichtierkunde. Vol. 1*. Gustave Fischer, Jena, 376 pp.
- Thomas, G. H., Hartmann, K., Jetz, W., Joy, J. B., Mimoto, A., & Mooers, A. O. (2013). PASTIS: an R package to facilitate phylogenetic assembly with soft taxonomic inferences. *Methods in Ecology and Evolution*, 4(11), 1011-1017.
- Trémaux, P. (1865) *Origine et transformations de l'homme et des autres êtres*. Hachette, Paris.
- Todd, J. A., & Johnson, K. G. (2013). Dissecting a marine snail species radiation (Conoidea: Turridae: Polystira) over 12 million years in the southwestern Caribbean. *Bulletin of Marine Science*, 89(4), 877-904.

- Todd, J. A., & Rawlings, T. A. (2014). A review of the *Polystira* clade—the Neotropic's largest marine gastropod radiation (Neogastropoda: Conoidea: Turridae sensu stricto). *Zootaxa*, 3884(5), 445-491.
- Troschel, F. H. (1866). Das Gebiss der schnecken zur Begründung einer natürlichen classification (Vol. 1). Nicolaische Verlagsbuchhandlung.
- Tryon, G. W. (1883). Structural and systematic conchology. An introduction to the study of the Mollusca, 2. Published by the author, Philadelphia. 430 pp., pls. 23–91.
- Tucker, J. K. (2004). Catalog of Recent and fossil turrids (Mollusca: Gastropoda). *Zootaxa*, 682(1), 1-1295.
- Tucker, J. K., & Tenorio, M. J. (2009). Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. *Conchbooks*.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, 40(15), e115-e115.
- Villanea, F. A., Parent, C. E., & Kemp, B. M. (2016). Reviving Galápagos snails: Ancient DNA extraction and amplification from shells of probably extinct endemic land snails. *Journal of Molluscan Studies*, 82(3), 449-456.
- Vincent, C. (2019) Identification des proies de gastéropodes venimeux (Conoidea) par approche de métabarcoding. Rapport de M2, Master SEP-SDUEE 2018-2019
- Watkins, M., Hillyard, D. R., & Olivera, B. M. (2006). Genes expressed in a turrid venom duct: divergence and similarity to conotoxins. *Journal of molecular evolution*, 62(3), 247-256.
- Watson, R.B. (1881) Mollusca of H.M.S. Challenger Expedition. Parts VIII-X. *Journal of the Linnean Society (London)*, 15, 388–475.
- Weber, M. G., & Agrawal, A. A. (2014). Defense mutualisms enhance plant diversification. *Proceedings of the National Academy of Sciences*, 111(46), 16442-16447.
- Wilkins, J. S. (2011). Philosophically speaking, how many species concepts are there?.
- Wilkinson, M., McInerney, J. O., Hirt, R. P., Foster, P. G., & Embley, T. M. (2007). Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends in ecology & evolution*, 22(3), 114-115.
- Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The perils of DNA barcoding and the need for integrative taxonomy. *Systematic biology*, 54(5), 844-851.
- Williams, S. T., & Duda Jr, T. F. (2008). Did Tectonic activity stimulate Oligo–Miocene speciation in the Indo-West Pacific?. *Evolution: International Journal of Organic Evolution*, 62(7), 1618-1634.
- Williams, S. T., Smith, L. M., Herbert, D. G., Marshall, B. A., Warén, A., Kiel, S., Linse, K., Vilvens, C. & Kano, Y. (2013). Cenozoic climate change and diversification on the continental shelf and slope: evolution of gastropod diversity in the family Solariellidae (Trochoidea). *Ecology and evolution*, 3(4), 887-917.

## *Bibliographie*

- Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, 107(20), 9264-9269.
- Yeates, D. K., Seago, A., Nelson, L., Cameron, S. L., Joseph, L. E. O., & Trueman, J. W. (2011). Integrative taxonomy or iterative taxonomy? *Systematic entomology*, 36(2), 209-217.
- Zaragüeta, R., & Pécaud, S. (2016). Why should cladograms be dichotomous?. *The Future of Phylogenetic Systematics: The Legacy of Willi Hennig*, 86, 230.
- Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinformatics*, 19(6), 153.
- Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29(22), 2869-2876.
- Zhang, S. Q., Che, L. H., Li, Y., Liang, D., Pang, H., Ślipiński, A., & Zhang, P. (2018). Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications*, 9(1), 205.
- Ziegler, A., & Menze, B. H. (2018). Accelerated acquisition, visualization, and analysis of zoo-anatomical data. In *Computation for Humanity* (pp. 258-287). CRC Press.
- Ziegler, A., Bock, C., Ketten, D. R., Mair, R. W., Mueller, S., Nagelmann, NPracht, E. D., & Schröder, L. (2018). Digital three-dimensional imaging techniques provide new analytical pathways for malacological research. *American Malacological Bulletin*, 36(2), 248-274.

### **ANNEXE 1 :**

Puillandre, N., Fedosov, A. E., **Zaharias, P.**, Aznar-Cormano, L., & Kantor, Y. I. (2017). A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess. *Zoological Journal of the Linnean Society*, 181(2), 243-271.

# A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): integrative taxonomy in a nomenclatural mess

NICOLAS PUILLANDRE<sup>1\*</sup>, ALEXANDER E. FEDOSOV<sup>2</sup>, PAUL ZAHARIAS<sup>1</sup>,  
LAETITIA AZNAR-CORMANO<sup>1</sup> and YURI I. KANTOR<sup>2</sup>

<sup>1</sup>Institut de Systématique, Évolution, Biodiversité ISYEB – UMR 7205 – CNRS, MNHN, UPMC, EPHE, Muséum national d'Histoire naturelle, Sorbonne Universités, 57 rue Cuvier, CP26, F-75005 Paris, France

<sup>2</sup>A.N. Severtzov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninski prospect 33, 119071 Moscow, Russian Federation

Received 11 October 2016; revised 23 January 2017; accepted for publication 1 February 2017

Integrative taxonomy, and in particular species delimitation using molecular data, often leads to the discovery of new species. However, these new species are not systematically turned into formally described species, because, among other reasons, linking molecularly defined groups with available taxonomic names can be tricky. Here we delimit species in the genus *Lophiotoma* (Gastropoda, Conoidea, Turridae) using two unlinked genetic markers (the mitochondrial *COI* gene and the nuclear 28S ribosomal RNA gene), shell and radula characters, and geographic and bathymetric distribution. Several methods of species delimitation (ABGD, GMYC and PTP) resulted in several alternate species partitions, discussed using an integrative approach. We ended up with ten different species, among which seven have been unequivocally linked to available species names. We designate neotypes for two of them (*L. acuta*, *L. jickelii*). The three remaining species are described as new: *L. semfala* sp. nov., *L. bratususa* sp. nov. and *L. kina* sp. nov. We discuss the difficulties encountered in locating type specimens and in linking them to recognized molecular species, in a context where the vast majority of mollusc types are empty, dried shells and consequently difficult to sequence.

ADDITIONAL KEYWORDS: ABGD – GMYC – neotype designation – PTP – species delimitation – species description.

## INTRODUCTION

While DNA and integrative taxonomy (Dayrat, 2005; Will, Mishler & Wheeler, 2005) certainly participated in the revival of taxonomic research in the last 10 years, their impact on species descriptions remains limited. Most species descriptions are still based on morphological characters only (Pante, Schoelink & Puillandre, 2014) and descriptions that include a molecular diagnosis remain scarce (Jörger & Schrödl, 2013; Renner, 2016). In the Mollusca collection of the Muséum National d'Histoire Naturelle (MNHN), Paris, the first holotype associated with a DNA sequence was

registered in 2008; since then, 2126 holotypes have been deposited in the MNHN collection of molluscs, but only 65 are linked to a DNA sequence. As quoted by Bouchet & Strong (2010), '80% of the new species descriptions of shelled marine gastropod species published in 2006 contained a description of the shell only [i.e. not only lacking mention of DNA characters, but also anatomy or radula]'

Why does the input of DNA characters remain so insignificant in the description of biodiversity, in spite of its growing popularity among biologists? One of the reasons lies probably in the dichotomy between taxonomists (including amateurs, particularly active in molluscs) and 'molecularists', people who actually produce the DNA sequences. Most species remain described based on morphological characters because these characters still remain largely more accessible than DNA

\*Corresponding author. E-mail: puillandre@mnhn.fr  
[Version of Record, published online 20 June 2017; <http://zoobank.org/> urn:lsid:zoobank.org:pub:27882D0C-5833-4DB4-98D9-77600EEC22C3]



characters. Conversely, most molecularists are not trained in taxonomy and nomenclature, and many of the new species they discover, some of them being undetectable with morphological characters, remain undescribed and thus virtually ignored by the scientific community (Goldstein & DeSalle, 2011). Nevertheless, both approaches should actually be encouraged and applied synergistically: on the one hand, many species are difficult to distinguish morphologically, and in these cases integrative taxonomy, including DNA characters, has proved its usefulness (Pante *et al.*, 2015); on the other hand, linking molecularly defined species to available names, and eventually proposing new names, requires knowledge of the nomenclatural rules, of the taxonomic literature and, in particular, of the type specimens.

However, even close examination of the type material may be of little use in marine molluscs, as many name-bearing types simply do not fulfil their function, being too worn and badly preserved to confidently link the species name to other, more recently collected, material (Bouchet & Strong, 2010). It is particularly true when several species share identical teleoconchs, differentiated only by protoconchs, radulae, anatomical or even DNA characters, as many of these characters are inaccessible on these types. Thus, a lost name-bearing type would actually be preferable, because in this case a neotype could be designated, which would provide an ultimate solution to a species identity problem. This, however, requires either a proof that the name-bearing types were lost or application to the Commission of Zoological nomenclature, both of which are time-consuming procedures.

To illustrate the benefit of a combination of molecularists and taxonomists, we applied an integrative taxonomic approach in a group of marine gastropods, *Lophiotoma* (Gastropoda, Conoidea, Turridae), which cumulates many of the difficulties listed above, plus some others, making it a good model to illustrate the link between species delimitation and species description: (1) preliminary results published in Puillandre *et al.* (2012b) suggest that several MOTUs can share very similar shells; (2) because of their shell variability, several described species have been synonymized in the literature, and many names are potentially applicable; and (3) type specimens of some species have been lost and are known by figures only and therefore are difficult to link to subsequently collected specimens. In this study, we apply the name *Lophiotoma* to the clade defined in Puillandre *et al.* (2012b) that includes the type species *L. acuta* (Perry, 1811), but exclude other species referred to as *Lophiotoma* [e.g. *L. albina* (Lamarck, 1822) or *L. natalensis* Bozzetti, 2016], or sometimes as 'larger *Lophiotoma*' (Olivera, 2004), but which are not phylogenetically related to *L. acuta*. These shallow-water turrids, restricted to the Indo-Pacific, have been known since the early 19th century. Like most other conoideans,

they are characterized by a venom apparatus, producing toxins used to capture their prey (most likely polychaetes). Their taxonomy was revised by Powell (1964), and although they are regularly sampled by shell collectors, only one additional species (*L. vezzaroi* Cossignani, 2015) referable to the *Lophiotoma* group, as circumscribed here, has been described since.

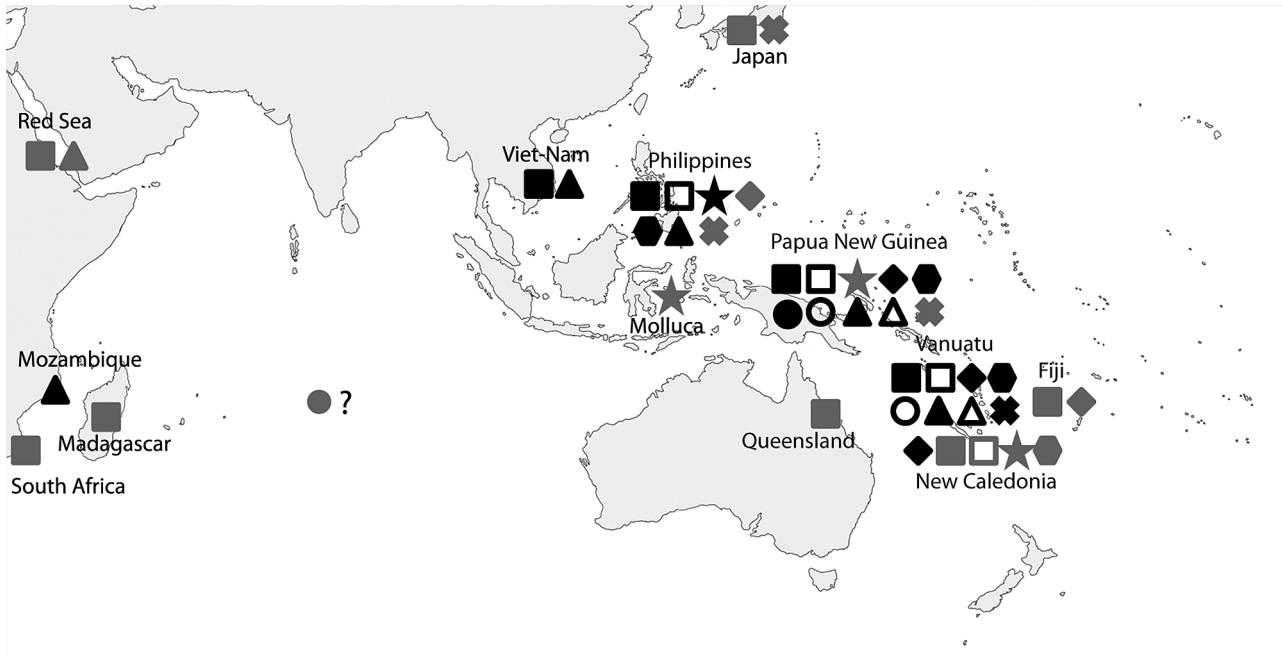
To delimit species in this genus, we followed the general workflow of Puillandre *et al.* (2012b): species hypotheses are proposed in an integrative framework, based on a unified species concept in which species are considered as definitely diverging lineages (De Queiroz, 2007; Samadi & Barberousse, 2009). First, primary species hypotheses (PSH) were proposed using part of the mitochondrial *COI* gene and three of the most widely used methods based on monolocus data: ABGD (automatic barcode gap discovery; Puillandre *et al.*, 2012a), GMYC (general mixed Yule coalescent model; Pons *et al.*, 2006; Monaghan *et al.*, 2009) and PTP (Poisson tree processes; Zhang *et al.*, 2013). Second, monophyly of the PSH was tested performing maximum likelihood and Bayesian inference analyses on both *COI* and nuclear *28S* genes, two unlinked genetic markers, to check whether each PSH corresponds to an independent lineage in both gene trees. Finally, morphological variability and geographic and bathymetric distributions were integrated to turn the PSH into secondary species hypotheses (SSH). In the final step, and after a deep search of the literature and in museum collections, available names were tentatively applied to the SSH, relying on shell characters and type localities; when no available name was found, the SSH was described as a new species.

## MATERIAL AND METHODS

### SAMPLING

The material was collected during several expeditions to the Indo-Pacific: PANGLAO 2004 and AURORA 2007 in the Philippines, SANTO 2006 in Vanuatu, INHACA 2011 in Mozambique, Nha-Trang in Vietnam (2010–2016) and PAPUA NIUGINI (2012) and KAVIENG 2014 in Papua New Guinea (expeditions.mnhn.fr) (Fig. 1). All material is stored in the MNHN.

Until 2012, live specimens for molecular analysis were anaesthetized with an isotonic solution of MgCl<sub>2</sub> and fixed in 96% ethanol. Specimens collected during later expeditions were processed with a microwave oven (Galindo *et al.*, 2014): the living molluscs in small volumes of sea water were exposed to microwaves for 7–30 s, depending on specimen size. Bodies were immediately removed from shells and dropped in 96% ethanol. Specimens are registered in the MNHN collection and sequences were deposited in BOLD (Barcode of Life Datasystem) and GenBank (Table S1).



**Figure 1.** Map showing the species distributions. Filled squares: *L. acuta*; empty squares: *L. semfala* sp. nov.; stars: *L. polytropia*; diamonds: *L. abbreviate*; hexagons: *L. brevicaudata*; full circles: *L. picturata*; empty circles: *L. bratasusa* sp. nov.; full triangles: *L. jickelii*; empty triangles: *L. kina* sp. nov.; crosses: *L. vezzaroi*; black symbols: species presence confirmed with sequenced specimens; grey symbols: species presence reported in the literature or confirmed by studied (non-sequenced) specimens. ?: 'Indian Ocean'.

#### DNA SEQUENCING

DNA was extracted using the Epmotion 5075 robot (Eppendorf), following the manufacturers' recommendations. A fragment of the cytochrome oxidase subunit I (*COI*) and of the 28S rRNA genes was amplified using universal primers LCO1490/HCO2198 (Folmer *et al.*, 1994) and either C1/D3 (Jovelin & Justine, 2001) or C2/CONO (GAAAAGAAGCTTTGAAGAGAGAGT) / D3 (Ober, 2002), respectively. PCRs were performed in 25  $\mu$ L, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM MgCl<sub>2</sub>, 0.26 mM dNTP, 0.3 mM each primer, 5% DMSO and 1.5 units of Qbiogene Q-Bio Taq. For the *COI* fragment, amplification consisted of an initial denaturation step at 94 °C for 4 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, followed by extension at 72 °C for 1 min. The final extension was at 72 °C for 5 min. The 28S PCRs were performed in 20  $\mu$ L reaction volumes, containing a final concentration of 1X SsoAdvanced Universal SYBR Green Supermix, 0.3 mM primers and 0.5  $\mu$ g/ $\mu$ L of BSA and 1  $\mu$ L of DNA extract. The amplification thermal profiles consisted of an initial denaturation for 3 min at 94 °C, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 1 min and a final extension at 72 °C for 5 min. PCR products were purified and sequenced by the Eurofins sequencing facility.

#### SPECIES DELIMITATION

*COI* sequences were aligned manually; 28S sequences were aligned using Muscle (Edgar, 2004) and alignments were checked by eye. Only 47 (over 907) nucleotides were variable in the 28S alignment, and only a few indels (of one nucleotide each) were detected. Pairwise genetic distances (*p*-distances) were calculated using MEGA 6 (Tamura *et al.*, 2013), following Srivathsan & Meier (2012). ABGD, GMYC, PTP and the phylogenetic methods were applied to the *COI* and 28S alignments, plus a concatenation of the *COI* and 28S alignments (for a subset of specimens – see Results section). For ABGD, the web version (<http://www.wabi.snv.jussieu.fr/public/abgd>) and the default parameters were used, with a *p*-distance model. Bayesian trees were reconstructed using BEAST v1.8.3 (Drummond *et al.*, 2012), running 100 000 000 (for the 28S and *COI* + 28S datasets) or 200 000 000 (for the *COI* dataset) generations with a sampling frequency of one tree each 4000 generations. Relative divergence times were estimated using a relaxed log-normal clock with a coalescent prior and a constant population size, following the recommendations of Monaghan *et al.* (2009). Both the 'single' (one single threshold is defined for the whole tree to delimit species) and 'multiple' (multiple thresholds in the tree can be eventually defined) methods of GMYC were applied using the trees obtained

with BEAST. Maximum likelihood trees, using RaxML v8.2.8 (Stamatakis, 2006), with the robustness of the nodes assessed using 1000 bootstraps, and a Bayesian tree, using Mr.Bayes 3.2.6 (Huelsenbeck, Ronquist & Hall, 2001), were reconstructed. For the MrBayes analyses, each of the two runs consisted of eight Markov chains and 20 000 000 generations, with eight chains, five swaps at each generation, a sampling frequency of one tree each 2000 generations and a chain temperature set at 0.02. For the Bayesian analyses (BEAST and MrBayes), convergence of each run was evaluated using TRACER 1.6 (Rambaut & Drummond, 2014) to check that all effective sample size values exceeded 200. Consensus trees were calculated after omitting the first 25% trees as burn-in. All phylogenetic analyses were performed on the Cipres Science Gateway (<http://www.phylo.org/portal2>). In all cases, a GTR + I + G substitution model was used, and the *COI* gene was divided into three partitions corresponding to the three codon positions (as suggested using the BIC score calculated by PartitionFinder – Lanfear *et al.*, 2016). For the concatenated datasets, four partitions were defined (three codon positions of the *COI* and *28S* gene). PTP was run with default parameters using the RAxML trees. Two specimens of closely related taxa were used as outgroups for phylogenetic analyses: *Turris babylonia* (Linnaeus, 1758) and *Iotyrria musivum* Kantor, Puillandre, Olivera & Bouchet, 2008 (Conoidea, Turridae).

The R package SPIDER 1.4-1 (Brown *et al.*, 2012) was used to identify pure diagnostic sites in each delimited species in the *COI* and *28S* alignments.

#### SHELL MORPHOLOGY AND RADULA

Radulae were prepared by standard methods (Kantor & Puillandre, 2012) and examined by scanning electron microscope TeScan TS5130MM in the Institute of Ecology and Evolution of Russian Academy of Sciences (IEE RAS). Protoconchs were measured in standard position and the number of whorls counted according to Bouchet & Kantor (2004).

#### ABBREVIATIONS USED IN TEXT

##### *Abbreviations of museums and repositories*

MMM: Mostra Mondiale Malacologia.  
 MHNG: Muséum d'Histoire Naturelle, Geneva, Switzerland.  
 MNHN: Muséum National d'Histoire Naturelle, Paris, France.  
 NHMUK: Natural History Museum of United Kingdom, London, UK.  
 SMF: Forschungsinstitut Senckenberg, Frankfurt, Germany.  
 USNM: National Museum of Natural History, Smithsonian Institution, Washington, DC, USA.

ZMB: Museum für Naturkunde, Humboldt-Universität, Berlin, Germany.

##### *Other abbreviations*

AL: Shell aperture length;  
 R/V: research vessel;  
 SL: shell length;  
 SW: shell width;  
 st.: station.

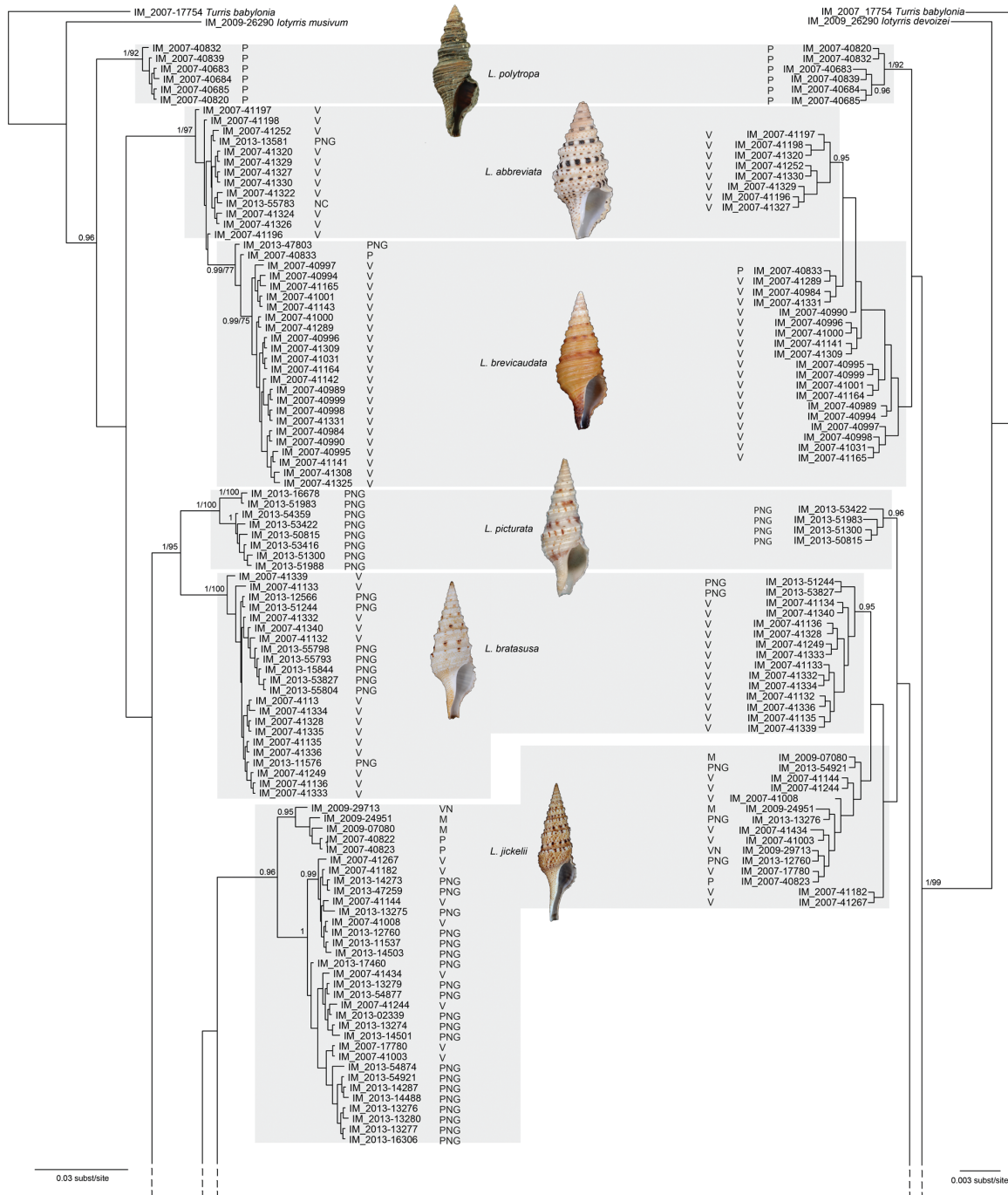
## RESULTS

Based on the partition with the highest number of PSH obtained with the *COI* gene, we built a reduced concatenated (*COI* + *28S*) dataset to limit computational time: a maximum of five specimens per PSH and per geographic region were kept. All the partitions obtained with ABGD, GMYC and PTP for the three datasets are shown in the Table 1. For the *COI* and *COI* + *28S* datasets, two partitions are discussed among the partitions proposed by ABGD: the partitions with the highest ('splitter' partition) and lowest ('lumper' partition) number of PSH (for the *28S* dataset, only one partition was proposed by ABGD). The results of the GMYC 'multiple' analyses are not shown, nor are the results of the GMYC 'single' analysis for *28S*, because they proposed an unrealistic number of PSH, that is not in agreement with either of the other methods, the phylogenetic trees or the other characters (111 PSH with the dataset *COI* for GMYC 'multiple', 27 for the *COI* + *28S* dataset for the GMYC 'multiple', and 79 and 78 for the *28S* dataset for the GMYC 'single' and 'multiple', respectively). In all cases, the GMYC 'multiple' partition was not significantly better than the GMYC 'single' partition ( $P \gg 0.05$ ). The *28S* gene is much less variable than the *COI* gene, and ABGD provided very few PSH with this gene (only five): this partition will be ignored in the rest of the text. In all the other cases, the number of PSH delimited varies from 8 to 16, all of them being compatible (i.e. they correspond to more or less inclusive PSH). In several cases (*L. picturata* 1, *L. semifala* 2 and *L. kina* 2 see Table 1 and below), these splits correspond to a single specimen isolated from the others, in PSH including few specimens (fewer than five).

By comparing the PSH obtained with the different datasets (Table 1), the results of the phylogenetic analyses (Figs 2, 3), the morphological variation and the bathymetrical and geographical distributions (Table S1), we turned the PSH into SSH and attributed available names to them or described them as new. Two PSH are found with all the genes and methods, have very distinct shells and in our material are restricted to a single archipelago: they were identified as *L. polytropa* (Helbling, 1779), restricted to the Philippines, and *L. vezzaroi*, in Vanuatu. Those two PSH also always correspond to highly supported clades in the phylogenetic analyses. The PSH identified as *L. abbreviata* (Reeve,



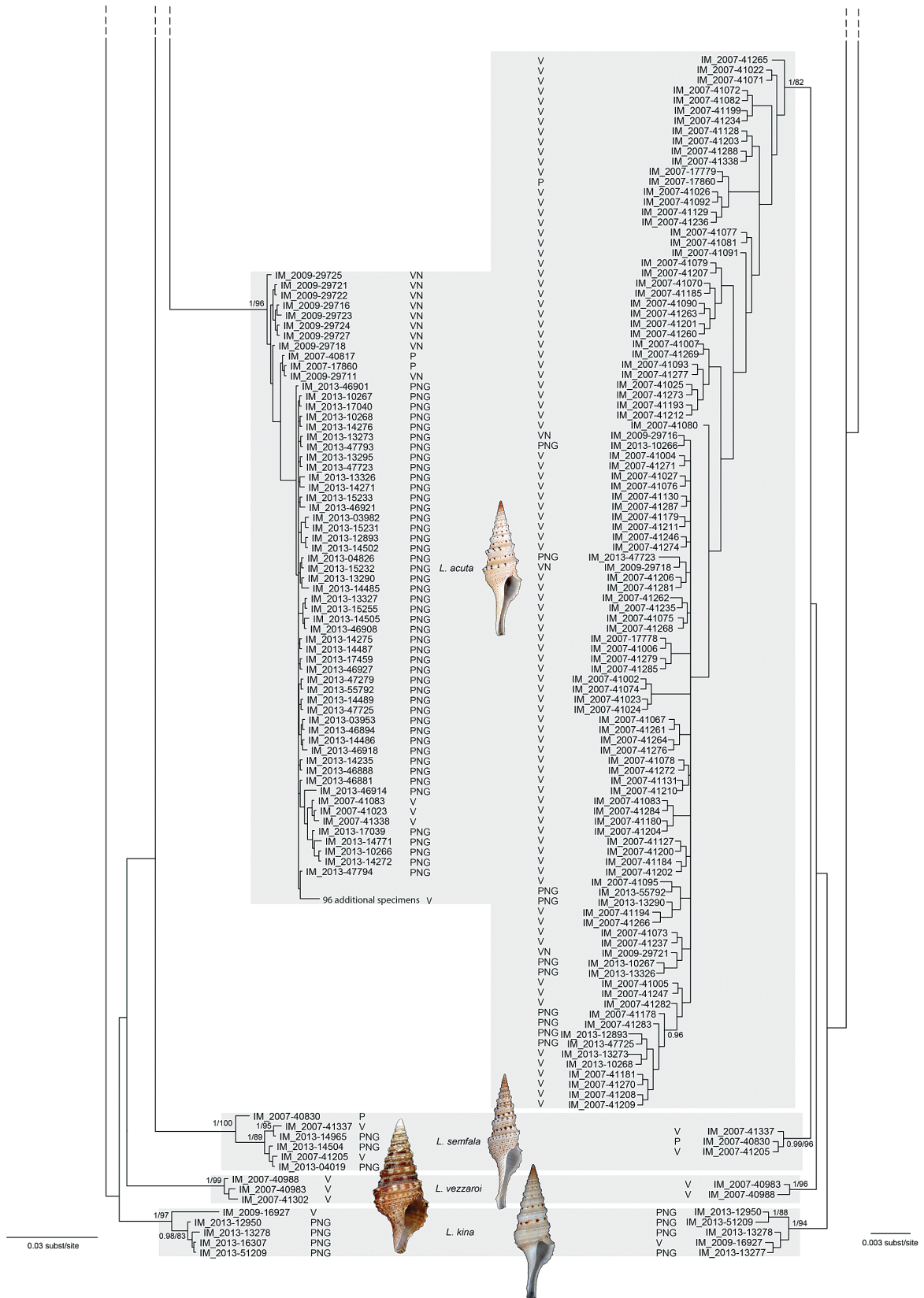




**Figure 2.** Bayesian trees (Mr Bayes) for the *COI* (left) and *28S* (right) genes. Posterior probabilities (>0.95) and bootstrap values (>75) are shown for each node. Letters next to each specimen number refer to the locality: M: Mozambique; VN: Vietnam; P: Philippines; PNG: Papua New Guinea; V: Vanuatu; NC: New Caledonia.

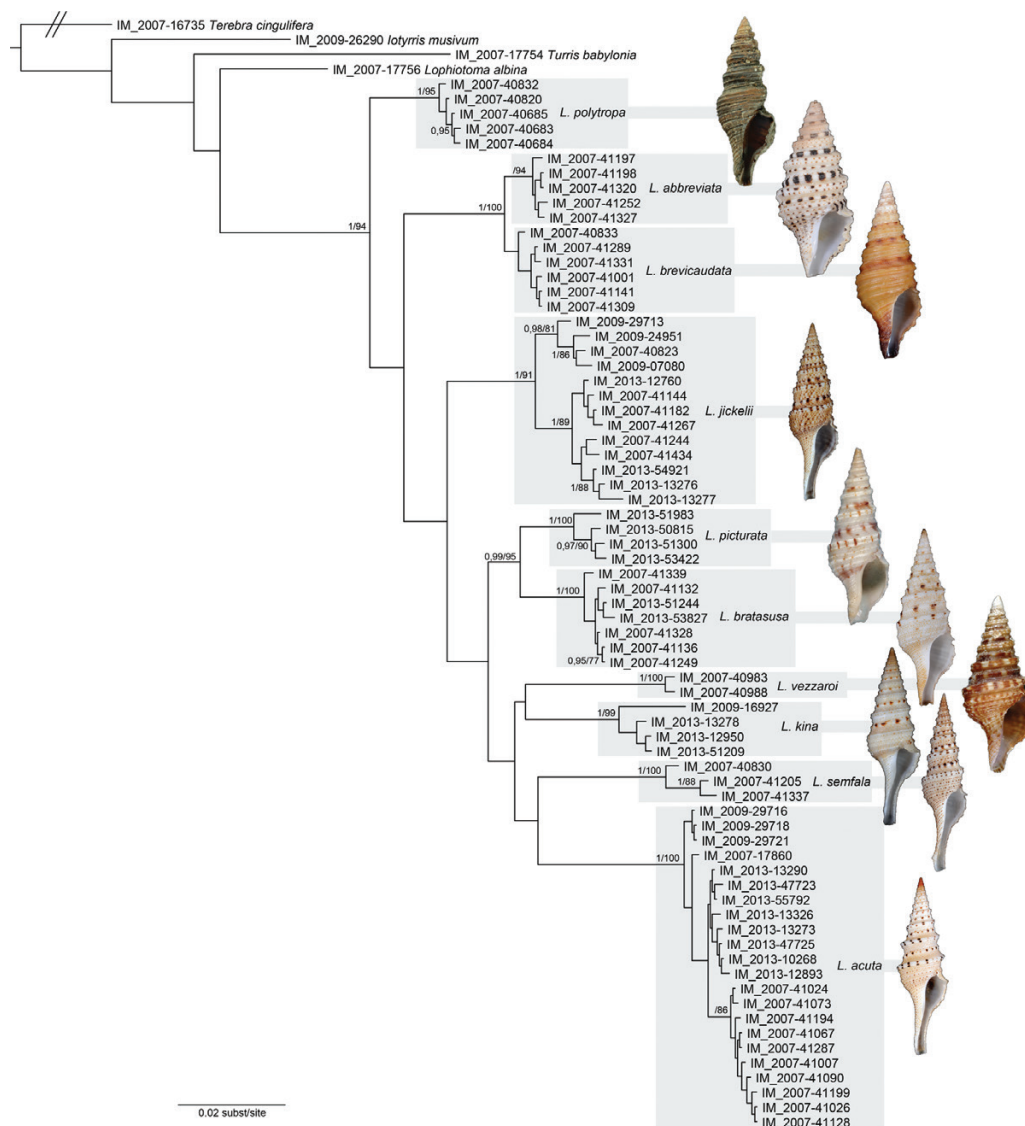
1843) and *L. brevicaudata* (Reeve, 1843), again with very distinct shells, are either found as a single PSH or as two different PSH (in one case – *COI* – PTP, *L. abbreviata* is divided in three groups, each corresponding to an unsupported clade). Their association generally corresponds to a highly supported clade. With the *28S* gene,

*L. abbreviata* is monophyletic and (moderately) supported and *L. brevicaudata* is not monophyletic; it is the opposite with the *COI* gene, and with both genes both PSH are reciprocally monophyletic. Both are found in sympatry, sometimes even at the same station. It is the only species pair that seems to have distinct bathymetric



Downloaded from https://academic.oup.com/zoolinnean/article-abstract/181/2/243/3873498 by Museum National d'Histoire Naturelle user on 31 July 2019

Figure 2. Continued.



**Figure 3.** Bayesian tree of the *COI* and *28S* genes concatenated. Posterior probabilities (>0.95) and bootstraps values (>75) are shown for each node.

preferences, *L. abbreviata* being found at an average depth of 2.9 m down to 7 m and *L. brevicaudata* at 14.8 m ( $\pm 9$  m). One supported clade, found in Papua New Guinea and Vanuatu, is constantly defined as a separate PSH (except with *28S* – PTP): *L. bratasusa* sp. nov. It is morphologically very similar to (see Taxonomy section), but distinguishable from, another PSH restricted to Papua New Guinea, sometimes co-occurring with it: *L. picturata* (Weinkauff, 1876). The latter is sometimes separated into two PSH ('splitter' partition of ABGD and PTP with the *COI* gene, PTP with the *COI* + *28S* dataset), morphologically undistinguishable and phylogenetically less supported (or even not recognized as monophyletic) than the whole PSH *L. picturata*. A similar situation is also found for a group of specimens with

shells preliminarily identified as *L. acuta*. The first PSH, *L. acuta*, is abundant and widely distributed and sometimes divided into two PSH ('splitter' partition of ABGD with the *COI* and *COI* + *28S* datasets). The second, *L. semifala* sp. nov., contains fewer specimens, also widely distributed (Philippines, Papua New Guinea and Vanuatu), and once again is sometimes divided into two PSH ('splitter' partition of ABGD and PTP with the *COI* gene, PTP with the *COI* + *28S* dataset). However, as for *L. acuta* and *L. picturata*, the support is lower for the subgroups. Finally, the two last PSH are also morphologically similar: *L. jickelii* (Weinkauff, 1875) and *L. kina* sp. nov. Once again, these were sometimes separated into two PSH each ('splitter' partition of ABGD and PTP with the *COI* gene, GMYC 'single' partition and PTP with the



*COI* + 28S dataset for *L. kina* and ‘splitter’ partition of ABGD, ‘single’ partition of GMYC and PTP with the *COI* and *COI* + 28S datasets for *jickelii*), less supported than the more inclusive PSH. *Lophiotoma picturata* and *L. bratasusa* sp. nov., *L. semfala* sp. nov. and *L. acuta*, on the one hand, and *L. jickelii* and *L. kina* sp. nov., on the other hand, are also found in sympatry, sometimes co-occurring at the same station.

## TAXONOMY

We provide descriptions for the new or newly defined taxa (as in case of neotype designation) and diagnoses for the species for which the status and scope do not change (compared to the generally accepted scope of the species). In addition to the type material, see Table S1 for the other material examined. For each species, the list of pure diagnostic sites for both *COI* and 28S genes is provided in Table 2.

### SUPERFAMILY CONOIDEA FLEMING, 1822

#### FAMILY TURRIDAE H. & A. ADAMS, 1853 (1838)

#### GENUS *LOPHIOTOMA* CASEY, 1904

*Type species: Pleurotoma acuta* Perry, 1811, OD.

**Diagnosis:** Shell medium-sized to large, narrow to broad fusiform, with attenuated, usually long and nearly straight canal. Protoconch multispiral or paucispiral. Teleoconch whorls usually angulated at shoulder. Sculpture of sharp pronounced cords, including sinus area. Anal sinus deep, with nearly parallel sides. Operculum with apical nucleus.

Marginal radular teeth duplex. Anterior (inner) half solid, narrowly lanceolate, dorso-ventrally compressed

with sharp lateral cutting edges. In posterior half major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation [*sensu* Kantor (2006)] either absent or very weak, represented by central tooth in shape of flat poorly developed cusp.

**Included species:** *Lophiotoma abbreviata* (Reeve, 1843); *L. acuta* (Perry, 1811); *L. bratasusa* sp. nov.; *L. brevicaudata* (Reeve, 1843); *L. jickelii* (Weinkauff, 1875); *L. kina* sp. nov.; *L. picturata* (Weinkauff, 1876); *L. polytropa* (Helbling, 1779); *L. ruthveniana* (Melvill, 1923); *L. semfala* sp. nov.; *L. vezzaroi* Cossignani, 2015.

**Remarks:** The genus was revised by Powell (1964) who recognized two subgenera (nominative one and *Lophioturris* Powell, 1964) differing on the basis of the protoconch – multispiral in the former and blunt paucispiral in the latter. Powell attributed five Recent species to *Lophiotoma* s.s. As specified in the Introduction section, previous analyses revealed that among those included species *Lophiotoma albina* should be excluded as it is more closely related to *Gemmula*-like species, while on the contrary *L. polytropa* attributed by Powell to *Lophioturris* is confidently included in *Lophiotoma* on the basis of an earlier phylogenetic analysis (Puillandre *et al.*, 2012b). The protoconch of *L. polytropa* is unknown so far. *Lophioturris*, with the type species *Turris indica* (Röding, 1798), clusters in one clade with *Unedogemmula* MacNeil, 1960 (type species *Pleurotoma unedo* Kiener, 1839), not related to *Lophiotoma* as defined here, and thus becomes junior subjective synonym of the latter. Among species treated as *Lophiotoma* by Powell (1964), only one species, *L. ruthveniana* (Melvill, 1923), is absent from our material and its position remains unconfirmed. The recently described *Lophiotoma vezzaroi* Cossignani, 2015 was sequenced and falls within the *Lophiotoma*

**Table 2.** List of diagnostic sites (character state – position) for both *COI* and 28S gene for each species

Species	<i>COI</i>	28S
<i>L. polytropa</i>	T – 290; G – 292; C – 334; A – 376; C – 424; G – 553	A – 854; T – 860
<i>L. abbreviata</i>	G – 331	C – 396
<i>L. brevicaudata</i>	G – 535	
<i>L. jickelii</i>	C – 158; A – 313; C – 457; T – 598	
<i>L. picturata</i>	C – 100; G – 181; T – 508; C – 529; C – 538	G – 858
<b><i>L. bratasusa</i> sp. nov.</b>	C – 151; G – 211; C – 238; C – 451	C – 833
<i>L. vezzaroi</i>	G – 37; C – 92; C – 259; C – 271; C – 347; T – 418; C – 533; T – 562	T – 541; A – 680; A – 696
<b><i>L. kina</i> sp. nov.</b>	A – 22; G – 232; G – 574; C – 613	
<b><i>L. semfala</i> sp. nov.</b>	C – 74; A – 85; T – 127; G – 208; T – 295; C – 307; C – 319; C – 328; C – 428	C – 404; T – 855; G – 860
<i>L. acuta</i>	T – 169; C – 287; G – 298; C – 364; C – 407	T – 496

The positions refer to the alignments provided in Appendices 1 and 2.



clade as defined here. This species was described from the Philippines and found by us in Vanuatu; conchologically it is rather similar to *L. ruthveniana*.

*LOPHIOTOMA ACUTA* (PERRY, 1811)

(FIG. 4)

*Pleurotoma acuta* Perry, 1811: pl. 5, fig. 5.

*Pleurotoma marmorata* Lamarck, 1816: pl. 439, fig. 6 (non *Pleurotoma marmorata* Link, 1807).

*Pleurotoma tigrina* Lamarck, 1822: 95 (nom. nov. pro *Pleurotoma marmorata* Lamarck, 1816, non *Pleurotoma marmorata* Link, 1807).

*Pleurotoma punctata* Schubert & Wagner, 1829: 155, pl. 234, figs 4103 a, b (no locality).

*Lophiotoma microsticta* Casey, 1904: 130.

*Lophiotoma acuta* Perry, 1811 – Powell, 1964 (part.): 303–305, colour plate 180, figs 1–10, 15–18 (non plate 180, figs 14, 19).

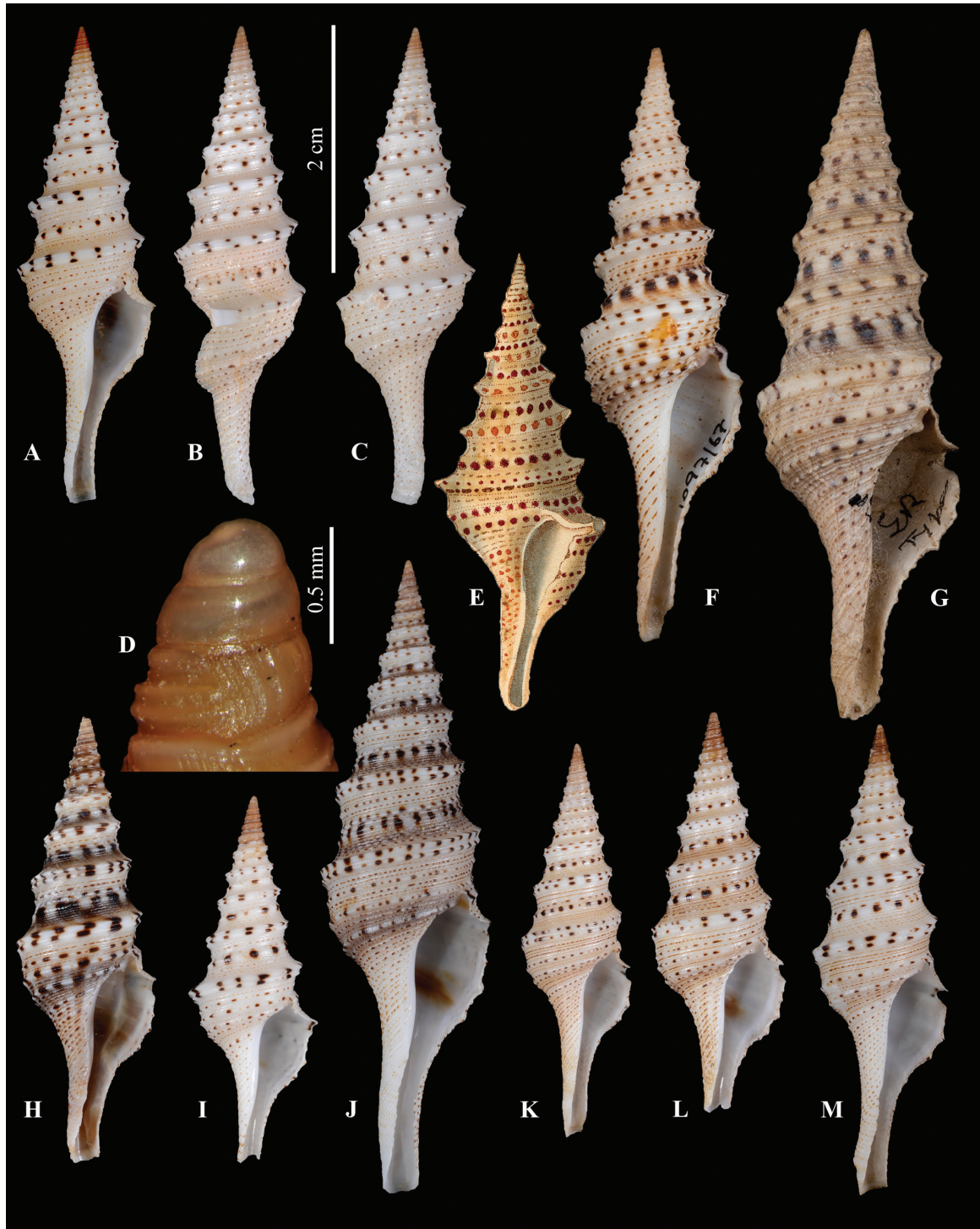
*Type material*: Neotype of *Lophiotoma acuta* (here designated), MNHN IM-2007-41179, the same specimen is designated as a neotype of *Pleurotoma punctata* (Schubert & Wagner, 1829). Three syntypes of *Pleurotoma tigrina*, MHNG (MHNG-MOLL-51664). Syntypes of *Lophiotoma microsticta*, ?USNM [fide Powell (1964), see below]. Type locality Cebu, Philippines.

*Type locality*: Vanuatu, E Malo Island, 15°43.4'S, 167°15'E, flat sand and dead corals, 6 m (Expedition SANTO 2006, st. DR84, R/V *Aldric*).

*Description (neotype)* (Fig. 4A–D): Shell medium thick, narrow fusiform, spire high, siphonal canal long narrow, slightly inclined to left. Protoconch (Fig. 4D) conical, of nearly three evenly convex whorls, smooth first whorls, posteriormost half whorl with nine axial nearly straight riblets, more densely spaced in posterior part of protoconch. Protoconch diameter 0.78 mm, height 0.85 mm. Teleoconch whorls strongly angulated at shoulder, ten in total. Suture shallow, subsutural region wide, distinctly concave, subsutural cord low, triangular in profile, with 3 weak angular ridges, central one strongest. Subsutural region smooth on upper teleoconch whorls, with one spiral ridge appearing on fourth, two on sixth, three on seventh and seven on last whorl. Paired sinus cords strongest and form strong angulated shoulder. On upper whorls both cords similar in size and rounded on top, on penultimate and last whorls cords distinctly triangular in profile, upper much stronger. Base of spire whorls smooth on first whorl, with one spiral cord on two to sixth whorls, starting from seventh whorl number of cords gradually increases, and penultimate whorl with six slightly different in size narrow cords; interspaces three to four times broader than cords. Base of last whorl with five

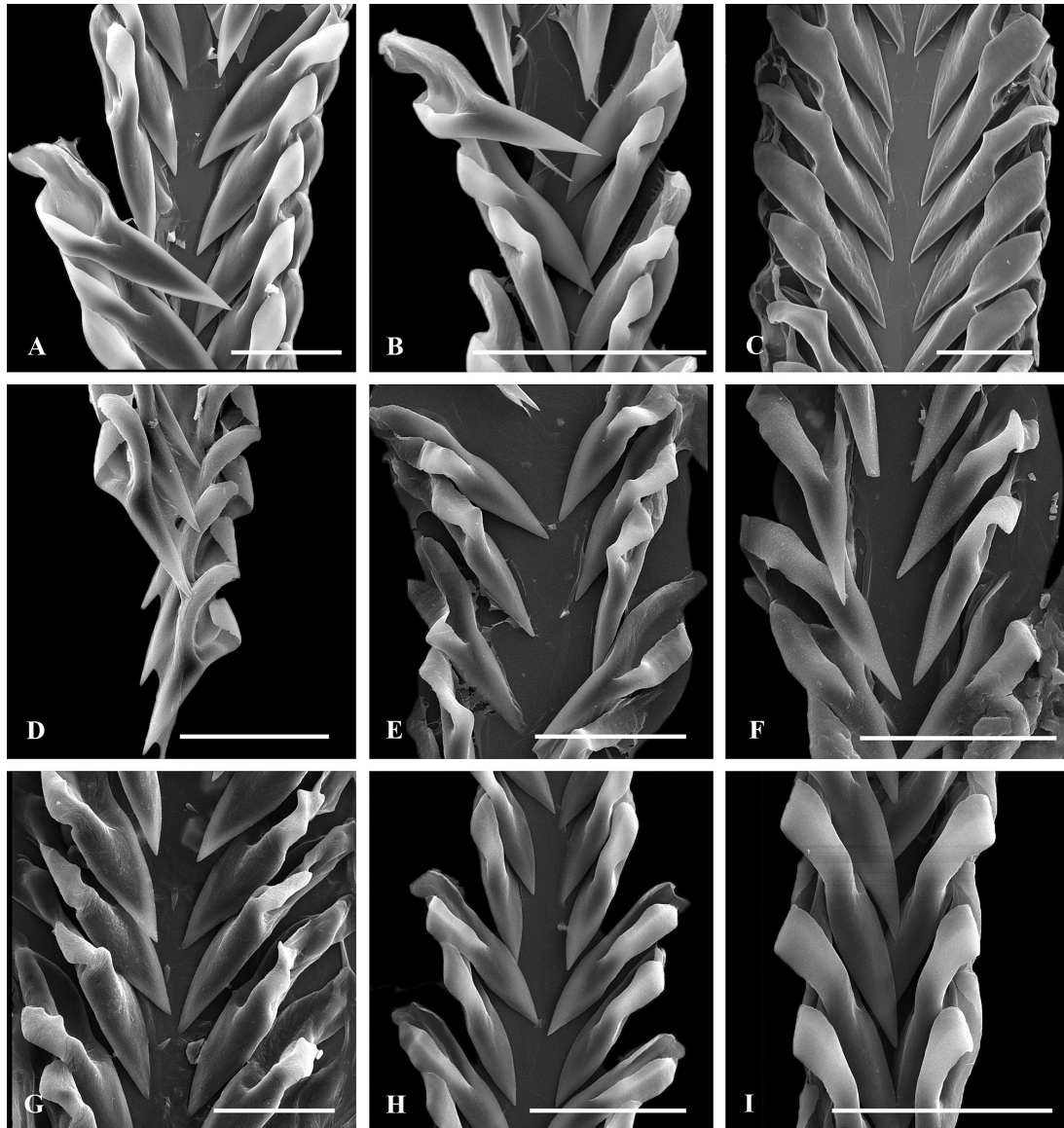
major spiral cords and several riblets between them, canal with 20 cords, becoming gradually broader, lower and more closely spaced anteriorly. Shell base gradually narrowing towards narrow and long nearly straight siphonal canal. Aperture pear shaped, outer lip concave in upper part and weakly convex below shoulder, gradually passing into canal. Anal sinus deep, with nearly parallel sides, with straight posterior margin parallel to shell axis; outer lip in side-view rounded and opisthocline, stromboid notch ill-defined. Growth lines indistinct, closely spaced. Shell creamy, protoconch and two first teleoconch whorls light brown. Subsutural cord with regularly spaced brown spots, not extending beyond cord. Sinus cords with distinct dark brown regularly spaced spots occupying whole width of cord and separate on each cord, minor spiral cords with dense brown flecks. Aperture creamy inside. Measurements: SL 38.8 mm, AL (with canal) 19.7 mm, SW 11.0 mm. Radula examined in five specimens, all from Papua New Guinea, very similar in all specimens (Fig. 5A, B). Radula membrane long, of 55–80 rows of teeth of which 25–30 not fully formed. Marginal teeth duplex. Anterior (inner) half solid, narrowly lanceolate, dorsoventrally compressed with sharp lateral cutting edges. In posterior half major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation absent or very weak, of flat poorly developed regularly positioned cusp, looking like folds of membrane.

*Remarks*: The species is very variable in terms of coloration and shell shape. The base colour can be from pure white to light orange and even light brown (subsutural region, shell base and canal) with lighter sinus area. With some reservation two colour forms can be distinguished, although intermediate specimens can also be found. In the light form, the brown spots are more scarce and usually confined to major cords, especially to subsutural and sinus ones, while the smaller cords have separate brown speckles. In the dark form (Fig. 4H), the entire shell can be light brown, with a lighter band along the sinus cords. The large brown spots on the subsutural cord dissolve in the lower part into brown band, occupying the entire subsutural zone. The brown spots on minor cords can be as large as those on sinus cords. The canal and anterior part of the aperture can also be brownish. Transitional specimens between forms can be found. The dark form was found within the entire distribution area of the species. In Vanuatu, which is most rich in sequenced material, 66% of specimens were represented by the light form, 24% by the dark form and 10% were attributed to intermediate forms (total number of checked specimens 94). A rather distinct form is found in Vietnam and the Philippines (Fig. 4J) – the shells are large (reaching 51 mm in our material), relatively heavy and with a less pronounced sinus cord, and the spots and speckles are rather fine, except those



**Figure 4.** *Lophiotoma acuta* (Perry, 1811). (A–D) Neotype, MNHN IM-2007-41179, SANTO 2006, st. DR84, SL 38.8 mm. (D) Lateral view of the protoconch. (E) Original illustration from Perry (1811). (F–G) Syntypes of *Pleurotoma tigrina* Lamarck, 1822 (MHNG-MOLL-51664). (F) SL 48.1 mm; (G) SL 56 mm. (H) Dark form, MNHN IM-2007-41007, SANTO 2006, st. FR10, SL 35.9 mm. (I) MNHN IM-2007-41025, SANTO 2006, st. LD01, SL 29.7 mm. (J) MNHN IM-2009-29711, Vietnam, Nha Trang Bay, st. ND7, SL 50.9 mm. (K) MNHN IM-2013-10267, PAPUA NIUGINI, st. PR07, SL 31.9 mm. (L) MNHN IM-2013-17040, PAPUA NIUGINI, st. PR152, SL 32.5 mm. (M) MNHN IM-2013-46888, KAVIENG 2014, st. KR06, SL 38.2 mm. All shells at the same scale.

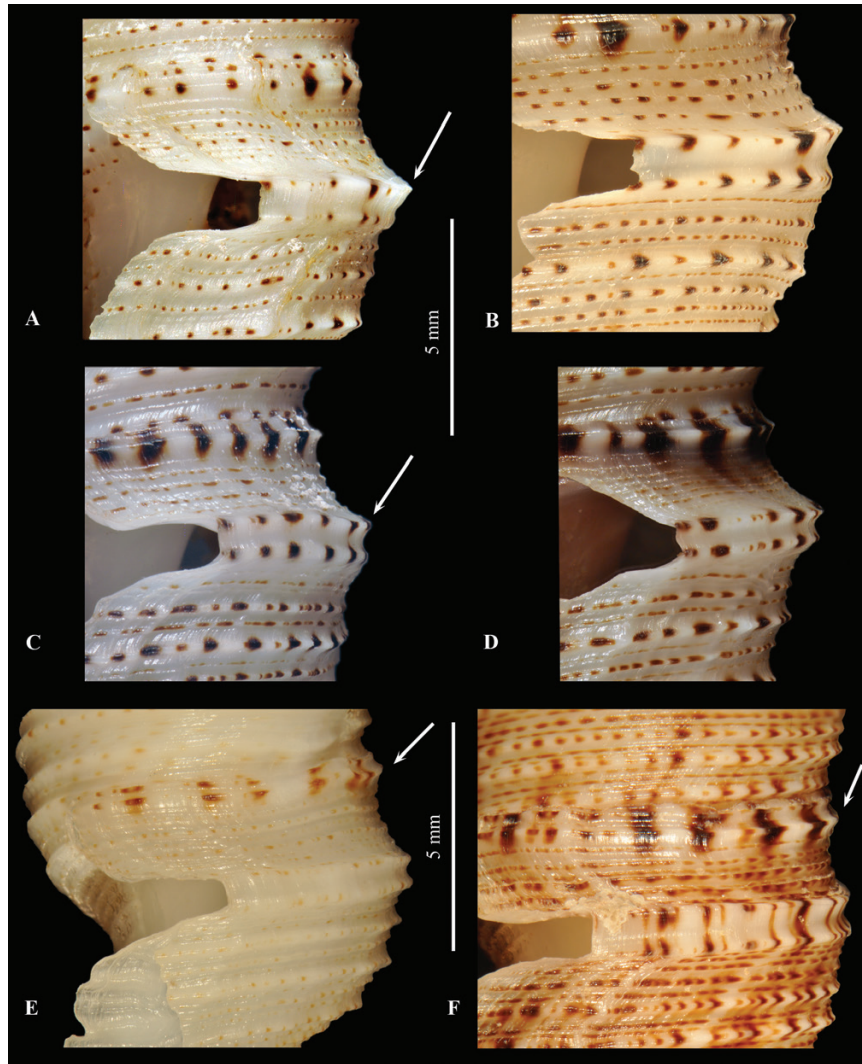




**Figure 5.** Radulae of studied *Lophiotoma*. (A–B) *Lophiotoma acuta* (Perry, 1811). (A) MNHN IM-2013-14235, PAPUA NIUGINI st. PD33, SL 32.1 mm. (B) MNHN IM-2013-14505, PAPUA NIUGINI st. PD41, SL 21.6. (C) *Lophiotoma polytropa* (Helbling, 1779), MNHN uncatalogued, PANGLAO 2004, st. M50. (D) *Lophiotoma brevicaudata* (Reeve, 1843), MNHN IM-2007-40994, SANTO 2006, st. DB12, SL 16.8 mm. (E) *Lophiotoma picturata* (Weinkauff, 1876), MNHN IM-2013-53422. (F) *Lophiotoma bratasusa* *sp. nov.*, holotype. (G) *Lophiotoma jickelii* (Weinkauff, 1875), neotype. (H) *Lophiotoma semifala* *sp. nov.*, MNHN IM-2013-14504. (I) *Lophiotoma kina* *sp. nov.*, holotype. Scale bars 50  $\mu$ m.

on the subsutural cord. In the molecular tree based on *COI* they are sister to the rest of *Lophiotoma acuta*, but do not form a monophyletic group. The syntype of *Lophiotoma microsticta* Casey, 1904 [illustrated by Powell (1964): pl. 233, figs 4–5], with shell of 59.7 mm in length, is rather similar to this form. Protoconchs studied in eight specimens are rather uniform, consisting of 2.75 whorls. Number of axial riblets varies from 6 to 11, protoconch height 0.88–0.95, diameter 0.8–0.83 mm.

The species is most similar to *L. semifala* *sp. nov.* and some specimens can hardly be distinguished; nevertheless, the morphology of the sinus cords seems to be rather uniform in *L. acuta* – on the last whorl (in adult specimens) the upper cord is much more pronounced than the lower and has a distinct triangular shape with sharp upper edge, while in *L. semifala* *sp. nov.*, the cords are nearly similar to each other and are more obtuse and rounded on top (Fig. 6).



**Figure 6.** Anal sinus and spiral sculpture of different species of *Lophiotoma*. (A) *Lophiotoma acuta* (Perry, 1811), MNHN IM-2007-41179. (B) *Lophiotoma acuta*, MNHN IM-2009-29711, SL 50.9 mm. (C) *Lophiotoma semifala* sp. nov., holotype, MNHN IM-2007-41337, SL 41.8 mm. (D) *Lophiotoma semifala* sp. nov., dark form, MNHN IM-2007-40830, SL 35.7 mm. (E) *Lophiotoma kina* sp. nov., holotype, MNHN IM-2013-16307, SL 31.0 mm. (F) *Lophiotoma jickelii* (Weinkauff, 1875), neotype, MNHN IM-2013-13275, SL 39.4 mm. Arrows indicate diagnostic details of the sculpture.

The species was treated as broadly distributed and strongly variable. Powell (1964) listed a number of nominal taxa in the synonymy of this species, including *Pleurotoma jickelii* Weinkauff, 1875 and *Pleurotoma picurata* Weinkauff, 1876. On the basis of molecular and morphological analysis, these two species appeared to be valid. *Pleurotoma acuta* Perry, 1811 was described without locality data or shell measurements. The original shell illustration is a bit grotesque, although suitable for positive identification. Few existing types described by Perry (1811) are stored in the NHMUK (Dance, 1986) and the type of *P. acuta* is not among them. Due to the complicated taxonomic situation with the *L. acuta* complex, a neotype is here

designated. The name *Pleurotoma marmorata* (non *Pleurotoma marmorata* Link, 1807 = *Turris chaldeia* Kilburn, Fedosov & Olivera, 2012) was listed by Lamarck (1816) (pl. 439, fig. 6, included in references, p. 8). Later Lamarck (1822: 95) renamed the species *P. tigrina*, citing his own figure, but still proposed the name *Pleurotoma marmorata* for another species, which became the homonym for the third time. Three syntypes of *Pleurotoma tigrina* are in MHNG (MHNG-MOLL-51664) (Fig. 4F–G herein) and it is seemingly conspecific with *L. acuta* in our current understanding, being closer to the ‘dark’ form. Judging from the syntypes of *P. marmorata* Lamarck, 1822 (MHNG-MOLL-51663) the species belongs to the



genus *Unedogemmula* and was listed in synonymy of *Lophiotoma (Lophioturris) indica* (Röding, 1798) by Powell (1964). The syntype of *Lophiotoma microsticta* Casey, 1904 was illustrated by Powell (1964: pl. 233, figs 4, 5) and claimed to be deposited in USNM. Nevertheless, we were not able to find it in the collections. Judging from the photo it has the same sculpture pattern as *L. acuta*, that is, the dominating upper sinus cord; therefore, we confirm the opinion of Powell (1964), that it is a synonym of *L. acuta*. The type material of *Pleurotoma punctata* was not traced despite queries in the corresponding museums and the original illustration is rather crude, although the general outline is similar to that of *L. acuta*. In order to fix the problem and to stabilize the nomenclature, we designate the neotype of *Pleurotoma acuta* Perry, 1811 as the neotype of *P. punctata* as well; thus, the latter name is now a junior objective synonym of *P. acuta*. *Pleurotoma peaseana* Dunker, 1871 [*Pleurotoma (Turris) peaseana* Dunker, 1871: 154 (Indian Ocean)] is another species of doubtful affinity, which was synonymized with *L. acuta* by Powell (1964). It was illustrated only in Weinkauff (1876, in Weinkauff & Kobelt, 1875–1887: 66, pl. 2, fig. 10). The illustration depicts a rather stout shell with moderately elongate canal, much shorter than in both *L. acuta* and *L. semfala*. The species may not be closely related to *L. acuta*. We were not able to trace the type despite querying museums where Dunker's type material might be stored. Powell (1964) synonymized the species with *L. acuta* without providing any arguments, an opinion followed by Oyama (1966) and Higo, Callomon & Gotō (1999). Moreover, Weinkauff (1876, in Weinkauff & Kobelt, 1875–1887) described the protoconch of *P. peaseana* as consisting of three smooth semitranslucent whorls with poorly visible suture, not mentioning the characteristic axial ribs in the posteriormost part of the protoconch. This seems more similar to the protoconch of *Unedogemmula* and we exclude the species from synonymy of *L. acuta*.

**Distribution:** Confirmed distribution of the species (based on sequenced specimens) – tropical Indo-west Pacific (from Vanuatu to Vietnam). Judging from published data, it also includes South Africa (Kilburn, 1983), Red Sea (Verbinnen & Dirks, 2007), Japan (Okutani, 2000), Fiji, Queensland (Australia) (Powell, 1964), New Caledonia (uncatalogued MNHN material).

***LOPHIOTOMA SEMFALA* SP. NOV.**

(FIG. 7)

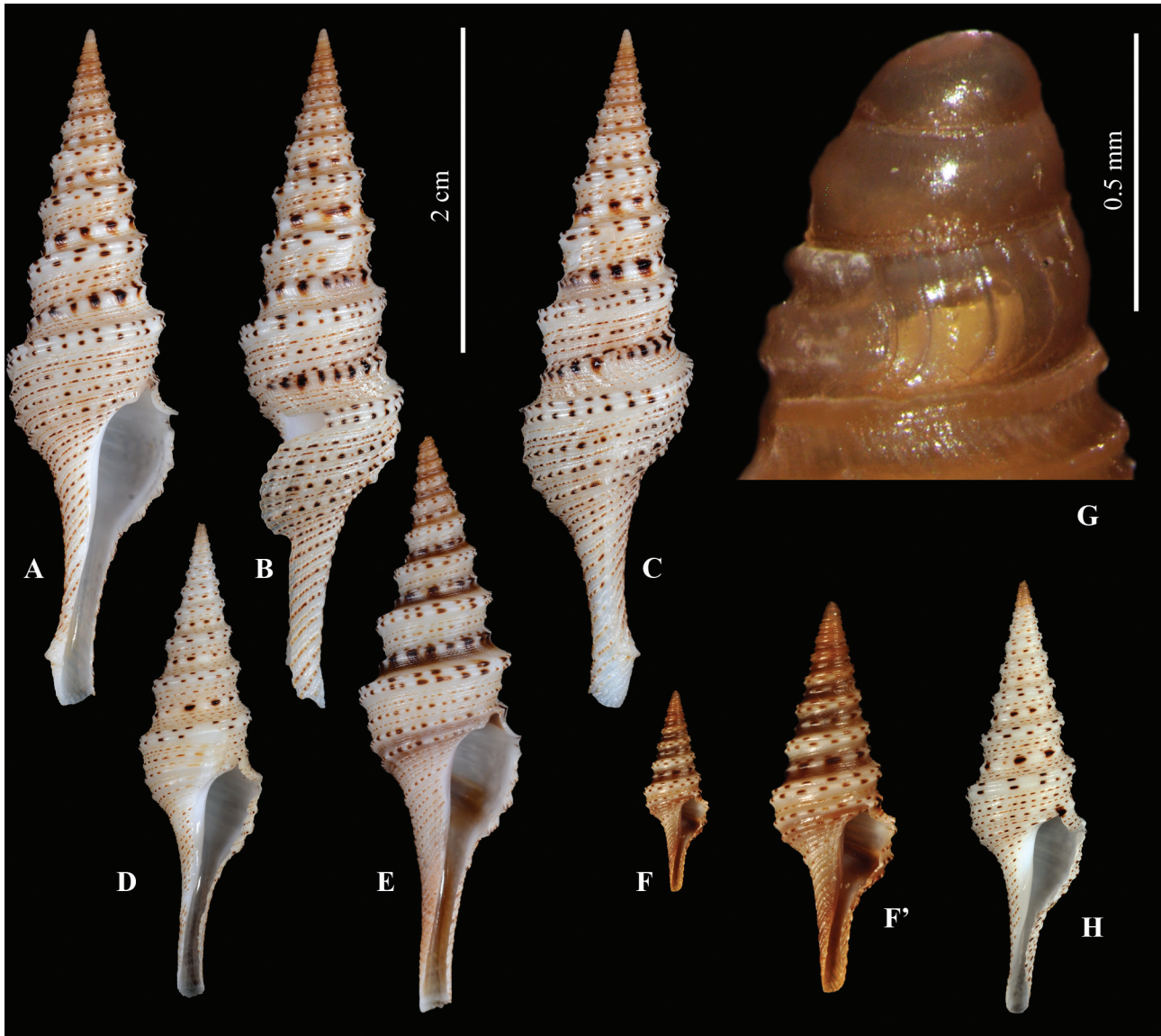
urn:lsid:zoobank.org:act:9E586542-B03B-4133-93F7-55B79FFD5A33

**Type material:** Holotype MNHN IM-2007-41337.

**Type locality:** Vanuatu, Aoré I. Aimbuei Bay, 15°32.8'S, 167°11.6'E, white coral sand, 3–8 m (Expedition SANTO 2006, st. LD35, R/V *Alis* annex).

**Etymology:** *semfala* – the 'same' in Bislama, the creole language, one of the official languages of Vanuatu. Used as noun in apposition to reflect the similarity to *Lophiotoma acuta*.

**Description (holotype) (Fig. 7A–C):** Shell medium thick, narrow fusiform, spire high, siphonal canal long narrow, slightly inclined to left. Protoconch conical, eroded, rendering exact whorl count and sculpture examination doubtful, of about three evenly convex whorls. Protoconch diameter 0.73 mm, height 0.85 mm. Teleoconch whorls angulated at shoulder, ten in total. Suture very shallow, indistinct, subsutural region wide, distinctly concave, subsutural cord low, triangular in profile, with three angular ridges on last whorl, central one strongest. On upper teleoconch whorls, only central ridge persists. Subsutural region smooth on upper teleoconch whorls, with one spiral ridge appearing on fourth, two on sixth, three on seventh and five on last whorl. Paired sinus cords strongest forming angulated shoulder. On upper whorls both cords nearly equal in size, obtusely triangular, on penultimate and last whorls cords more angulate, although still rounded on top, only on last whorl upper cord distinctly stronger than lower. Base of spire whorls smooth on first four whorl, with one spiral cord on fifth to sixth whorls, starting from seventh whorl number of cords gradually increases, and penultimate whorl with seven narrow cords of slightly different size, median much stronger; interspaces three to four times broader than cords. Base of last whorl with three major spiral cords and several riblets between them, canal with 20 cords, becoming gradually broader, lower and more closely spaced anteriorly. Shell base gradually narrowing towards narrow and long nearly straight siphonal canal. Aperture pear shaped, outer lip concave in upper part and weakly convex below shoulder, gradually passing into canal. Anal sinus deep, with nearly parallel sides, with straight posterior margin, parallel to shell axis; outer lip in side-view rounded and opisthocline, stromboid notch well defined. Growth lines indistinct, closely spaced. Shell creamy, protoconch and three first teleoconch whorls very light brown. Subsutural cord with regularly spaced brown spots, not extending beyond cord, broader on last three whorls. Sinus cords with distinct dark brown regularly spaced spots occupying whole width of cord and separate on each cord, minor spiral cords with dense brown flecks. Aperture creamy inside. Measurements: SL 41.8 mm, AL (with canal) 20.9 mm, SW 10.6 mm. Radula examined in three specimens, two from Papua

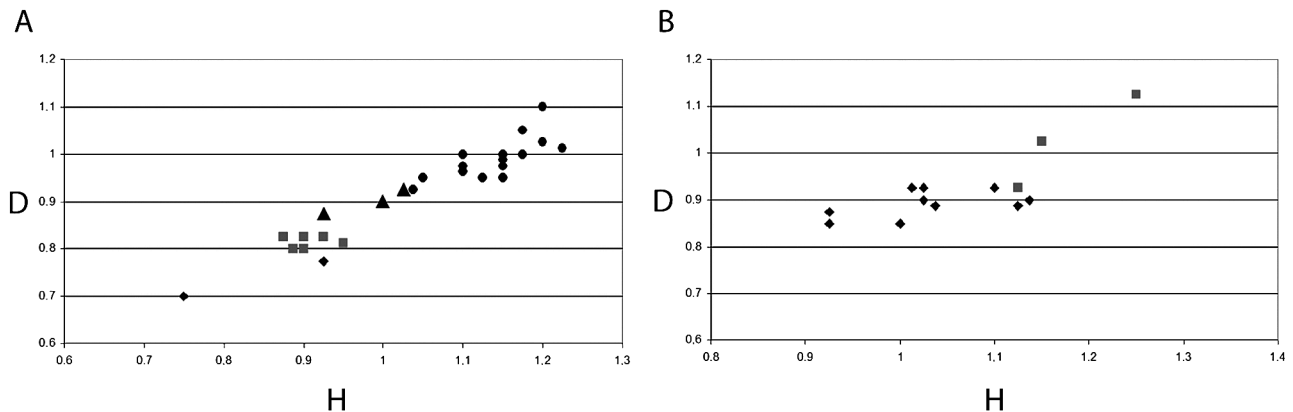


**Figure 7.** *Lophiotoma semifala* sp. nov. (A–C) Holotype, MNHN IM-2007-41337, SANTO 2006, st. LD35, SL 41.8 mm. (D) MNHN IM-2013-14504, PAPUA NIUGINI, st. PD41, SL 29.5 mm. (E) MNHN IM-2007-40830, Philippines, PANGLAO 2004, st. R62, SL 35.7 mm. (F–G) Dark form, MNHN IM-2013-4019, PAPUA NIUGINI, st. PD39, SL 12.4 mm (F – at the same scale as other shells, F' enlarged). (G) Lateral view of the protoconch. (H) MNHN IM-2013-14965, PAPUA NIUGINI, st. PD45, SL 26.8 mm. All shells (except F') at the same scale.

New Guinea and one from the Philippines, very similar in all examined specimens (Fig. 5H). Radula membrane medium long, of 33–50 rows of teeth of which 9–16 not fully formed. Marginal teeth duplex. Anterior (inner) half solid, narrowly lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In posterior half major and accessory limbs rather thin, bifurcate at about 45° angle. Central formation absent.

**Remarks:** The new species is represented only by six specimens, including the holotype and despite the

limited material, two rather distinct forms can be recognized. The 'light' form that includes the holotype has fewer brown spots and the base colour is uniformly creamy. The brown spots on the subsutural cord are in most specimens confined to the cord itself and do not extend beyond, but in the holotype on some whorls there are brownish blurred extensions of the spots to the subsutural region. Available specimens other than the holotype are smaller and less speckled. The 'dark' form is represented by two specimens only, one being juvenile (Fig. 7F–G). It has



**Figure 8.** Scatterplot of protoconch measurements in studied species of *Lophiotoma*. D: protoconch diameter, mm; H: exposed height, mm. (A) Diamonds: *L. semifala* sp. nov.; squares: *L. acuta*; triangles: *L. kina* sp. nov.; circles: *L. jickelii*. (B) Diamonds: *L. bratasusa* sp. nov.; squares: *L. picturata*.

slightly darker base colour, with a light brown shell base and canal and with the subsutural region below the subsutural cord uniformly brown. There was no correlation between geographic distributions, since one specimen of the dark form was collected in the Philippines, while another in Papua New Guinea at similar depths. The sinus cords of the adult specimen of the dark form are also sharper on top on the last whorl. On most parts of the teleoconch whorls, the sinus cords are either similar in size, or the lower even slightly more pronounced, than the upper, but on the last whorl the situation is reversed. An intact protoconch persists only in the juvenile of the dark form (Fig. 7G), it consists of 2.75 whorls, diameter 0.68 mm, height 0.73, which is significantly smaller than in holotype, although the existing material is insufficient for estimates of variation. The species is extremely similar to *Lophiotoma acuta*, which also has dark and light forms. It can be distinguished in most cases by being less pronounced and more rounded on the top sinus cords, providing a less angulated appearance to the shell shoulder, as well as the cords being more similar in size (Fig. 6, compare A, B with C, D), and domination of the lower cord over the higher one on the teleoconch whorls. The protoconch of *Lophiotoma acuta* is slightly larger (Fig. 8), while the radula is longer (consists of 55–80 rows of teeth vs. 33–50 rows in *L. semifala* sp. nov.).

**Distribution:** The species was found in the Philippines, Papua New Guinea and Vanuatu. In all these localities, it is sympatric with *L. acuta*. Judging from available material (only six sequenced specimens), it is much more rare than *L. acuta*, for which we had sequenced more than 160 specimens. Although we did not sequence any specimens from New Caledonia, judging

from the shell characters the species is also found in New Caledonia (uncatalogued MNHN material).

*LOPHIOTOMA POLYTROPA* (HELBLING, 1779)

(FIG. 9E)

*Murex (Fusus) polytropus* Helbling, 1779: 119, pl. 2, figs 24, 25.

*Pleurotoma fascialis* Lamarck, 1822: 93; – Kiener, 1840: 27, pl. 4, fig. 2.

*Lophiotoma (Lophioturris) polytropus*. – Powell, 1964: 313–314, pl. 244.

*Lophiotoma polytropus*. – Poppe, 2008: 770, pl. 680, fig. 4.

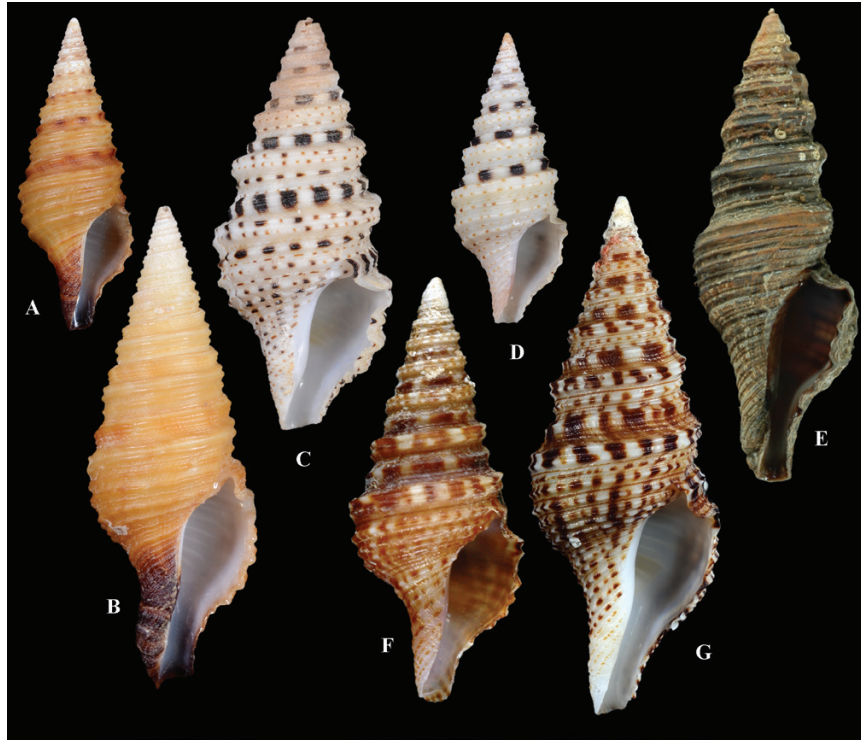
*Lophioturris polytropus*. – Lozouet & Plaziat, 2008: 134, pl. 31, figs 5–9.

**Type material:** *Murex (Fusus) polytropus*, whereabouts unknown; syntypes of *Pleurotoma fascialis*, MHNG (personal communication of P. Stahlschmidt, not seen).

**Type locality:** Not stated.

**Diagnosis:** Shell medium-sized, exceeding 50 mm, thick, turritiform, with thick brown periostracum, shell dark-purplish brown. Sculpture of strong spiral elements, with rounded or angulate subsutural cord followed by notably elevated paired and broadly spaced sinus cords. Shell periphery and base with dense elevated cords, similar in size to sinus cords and with intermediate finer ridges. Siphonal canal medium long, nearly straight; aperture rather wide, purplish to greyish inside. Radula (Fig. 5C) with duplex marginal teeth. Anterior (inner) half solid, lanceolate, slightly asymmetrical, with nearly straight anterior margin and convex posterior margin, dorso-ventrally compressed with





**Figure 9.** Shells of examined species of *Lophiotoma*. (A–B) *Lophiotoma brevicaudata* (Reeve, 1843). (A) MNHN IM-2007-40994, SANTO 2006, st. DB12, SL 16.7 mm. (B) MNHN IM-2013-47803 KAVIENG 2014, st. KS15, SL 26.0 mm. (C–D) *Lophiotoma abbreviata* (Reeve, 1843). (C) MNHN IM-2013-55783, New Caledonia, Nouméa, SL 22.4 mm. (D) MNHN IM-2007-41197, SANTO 2006, st. FB52, SL 15.8 mm. (E) *Lophiotoma polytropa* (Helbling, 1779), MNHN IM-2007-40832, PANGLAO 2004, st. M30, SL 43.0 mm. (F–G) *Lophiotoma vezzaroi* Cossignani, 2015. (F) MNHN IM-2007-40983, SANTO 2006, st. DS04, SL 14.4 mm. (G) Radula voucher, Tinina, Balut Island, Philippines, SL 34.7 mm.

sharp lateral cutting edges. In posterior half major and accessory limbs bifurcate at about 45° angle, rather thin. Accessory limb narrowing interiorly, where it fuses with major limb. Central formation absent.

**Remarks:** The species is rather distinct from all other congeners in having a strong, tightly adhered periostracum and uniformly coloured dark shell.

*Pleurotoma fascialis* Lamarck, 1822, was considered as a synonym of *L. polytropa* by Powell (1964). Although we have not seen the type material of the species, judging from the illustration of Kiener (1840: 27, pl. 4, fig. 2), the type of *Pleurotoma fascialis* is morphologically similar to the illustration of Helbling and therefore we follow Powell's opinion. Because the recognition of this species is not an issue, even though we were not able to locate the types of Helbling, we do not designate a neotype for *L. polytropa*.

**Distribution:** Powell (1964) recorded the species from the Philippines, Moluccas, New Britain and New Caledonia. The species is considered rare. Nevertheless, Lozouet & Plaziat (2008) found it common in the

mangrove environments of the lower estuary of the Abatan River (Bohol, Philippines). All the sequenced specimens originated from this locality. The species was successfully recollected several years later in the mentioned biotope (Kantor, Fedosov, unpublished).

#### *LOPHIOTOMA ABBREVIATA* (REEVE, 1843)

(FIG. 9C, D)

*Pleurotoma abbreviata* Reeve, 1843 (in 1843–1846): pl. 10, fig. 86.

*Lophiotoma abbreviata*. – Powell, 1964: 309, pls. 237–238, figs 1, 2; Poppe, (2008): pl. 683, fig. 5.

**Type material:** Lectotype [designated by Powell (1964)] and three paralectotypes in NHMUK.

**Type locality:** Masbate Island, Philippines, reefs at low tide.

**Diagnosis:** Shell small, turritiform, with contrasting black spots on white background colour, and short siphonal canal, giving shell stout appearance. Sculpture of strong

spiral elements, with rounded or angulate subsutural cord followed by notably elevated bisected sinus cord, and one fainter ridge on spire whorls. Shell base with dense elevated cords, sometimes interchanged by fine ridges. Microsculpture of dense very fine spiral treads throughout shell surface. Siphonal canal short and rather robust; aperture rather wide with moderately deep anal sinus. Inside of outer lip with distinct lirae.

*Remarks:* The small and robust-looking shell of *L. abbreviata* differs from notably more elongated, with long siphonal canal *L. jickelii*, *L. vezzaroi*, *L. semfala* sp. nov. and *L. kina* sp. nov. In turn, the variegated colour pattern readily distinguishes *L. abbreviata* from tan *L. brevicaudata* and dark-brown *L. polytropa*. While being distinctive among congeners, *L. abbreviata* resembles small species of the genus *Iotyrris*, *I. devoizei* and *I. kingae*, primarily in colour pattern. However, both mentioned *Iotyrris* species have an even shorter siphonal canal and thus proportionally much higher spire. Besides, the spiral elements are denser, and the whorl profile is less angulate, because of the lower sinus cord in *Iotyrris* species. Powell recognized two subspecies in addition to the nominotypical: *L. abbreviata lifouensis* (Sowerby, 1907) known only from Lifou, Loyalty Islands; and *L. abbreviata ustulata* (Reeve, 1846) with unknown type locality. The latter subspecies differs markedly in shell from the nominotypical one and its status remains unclear, as suggested by Powell (1964). We also did not have specimens from Lifou available for sequencing and therefore the status of *L. abbreviata lifouensis* remains unresolved. Concerning the latter, Cernohorsky (1972) claimed that the shells corresponding to both nominotypical and *lifouensis* subspecies were collected sympatrically in Fiji.

*Distribution:* Confirmed distribution of the species (based on sequenced specimens) is Papua New Guinea, New Caledonia and Vanuatu. According to published data, it is also found in the Philippines (Springsteen & Leobrera, 1986) eastward to Fiji (Cernohorsky, 1972).

*LOPHIOTOMA BREVICAUDATA* (REEVE, 1843)

(FIG. 9 A, B)

*Pleurotoma brevicaudata* Reeve, 1843 (in 1843–1846): pl. 15, fig. 126.

*Lophiotoma brevicaudata* – Powell, 1964: 406.

*Type material:* Lectotype and two paralectotypes in the NHMUK [designated by Powell (1964)] (not illustrated).

*Type locality:* Ticao Island, Philippines, H. Cuming collection.

*Diagnosis:* Shell small, turritiform, with prominent spiral sculpture; spire coloured light-brown or tan, siphonal canal dark-brown. Whorl outline indistinctly convex, as subsutural cord separated from succeeding cords by wide and deep depression. Sinus cord wide, composed of two ridges with rather shallow interspace, followed by two cords on whorl's base. Interspaces between cords sculptured by fine treads. Shell base convex, constricted to rather slender siphonal canal, sculptured with dense spiral to oblique cords. Aperture elongate, anal sinus moderately deep, wide, angulated at tip. Outer aperture lip with white callus, distinctly lirate within.

*Remarks:* *Lophiotoma brevicaudata* is one of the easily recognizable species, primarily because of its characteristic colour pattern with tan or light brown background colour, and dark siphonal canal. Crests of spiral ridges are sometimes dark-brown as well. In particular, rather monotonous coloration of the spire readily sets *L. brevicaudata* apart from the most closely related *L. abbreviata*. At the same time, *L. brevicaudata* is notably lighter, and in maturity smaller than *L. polytropa*. In addition to colour pattern, a rather short siphonal canal, as compared to *L. acuta*, *L. jickelii*, *L. vezzaroi*, *L. semfala* sp. nov. and *L. kina* sp. nov., allows rather straightforward identification of *L. brevicaudata* among congeners. The radula was examined in one sequenced specimen from Vanuatu (Fig. 5D). The radula is very similar to other congeners, with duplex marginal teeth. The anterior (inner) half is solid, narrowly lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In the posterior half the major and accessory limbs bifurcate at an angle of about 45°, rather thin. The central formation was not studied due to radula preparation.

*Distribution:* Confirmed distribution of the species (based on sequenced specimens) is from Philippines to Vanuatu. According to MNHN material, also New Caledonia.

*LOPHIOTOMA PICTURATA* (WEINKAUFF, 1876)

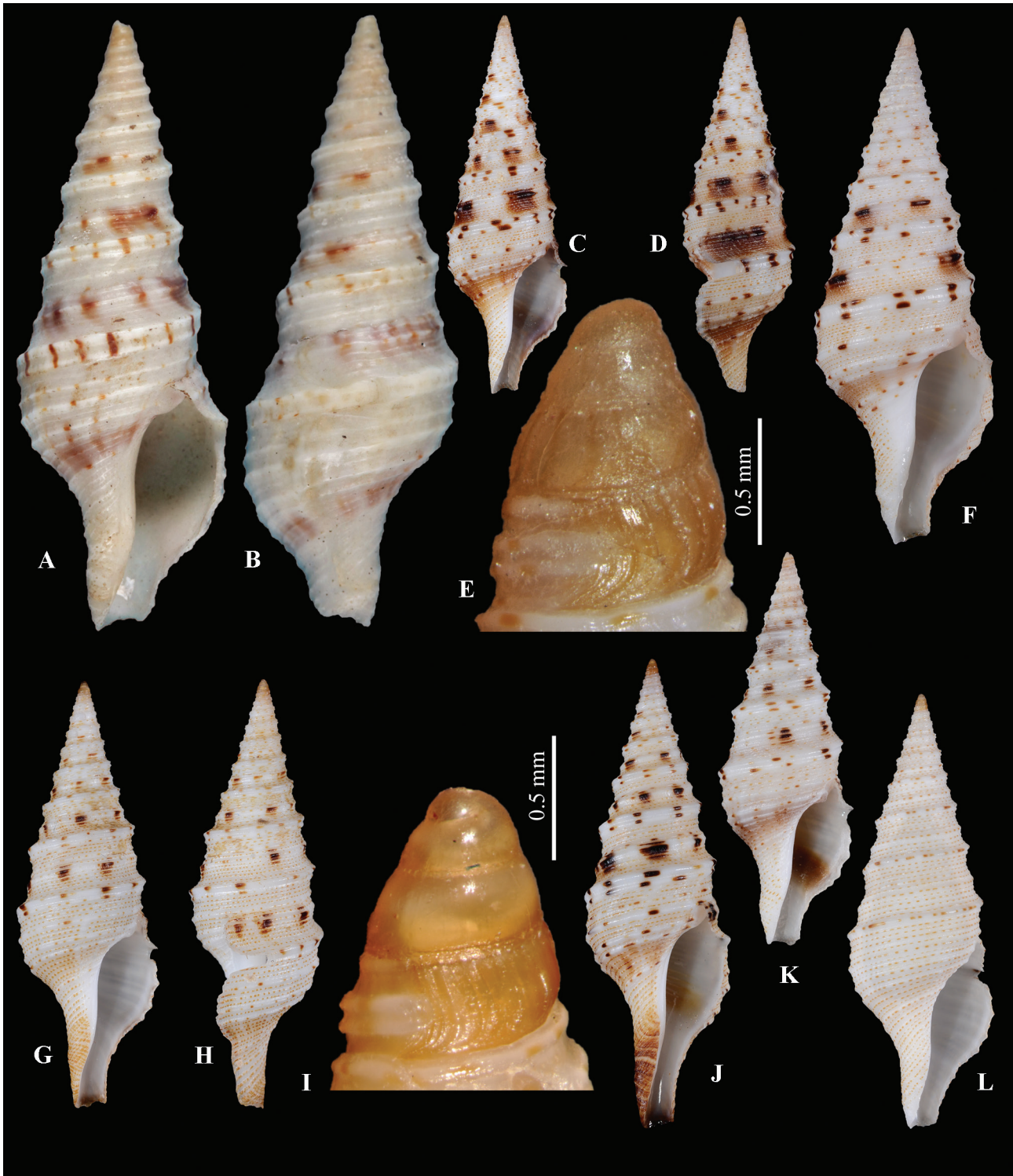
(FIG. 10A–F)

*Pleurotoma picturata* Weinkauff, 1876 in Weinkauff & Kobelt, 1875–1887: 66, pl. 2, fig. 10.

*Type material:* Lectotype (here designated) ZMB Moll 112610, ex-Paetel collection, Philippines, SL 41 mm; paralectotype ZMB Moll 112610.

*Type locality:* Philippines (originally Indischer Ocean).

*Diagnosis:* Shell solid, narrow turritiform, with high spire and moderately long siphonal canal. Protoconch



**Figure 10.** Shells of examined species of *Lophiotoma*. (A–F) *Lophiotoma picturata* (Weinkauff, 1876). (A, B) Lectotype of *Pleurotoma picturata* ZMB Moll 112610, SL 41 mm. (C–E) MNHN IM-2013-53422, KAVIENG 2014, st. KZ02, SL 24.5 mm. (E) Lateral view of the protoconch. (F) MNHN IM-2013-51988, KAVIENG 2014, st. KR62, SL 30.3 mm. (G–I) *Lophiotoma bratasusa* sp. nov. (G–I) Holotype, MNHN IM-2013-51988, KAVIENG 2014, st. KR54, SL 26.0 mm. (J) MNHN IM-2013-15844, PAPUA NIUGINI, st. PM41, SL 30.5 mm. (K) MNHN IM-2007-41339, SANTO 2006, st. DR106, SL 28.2 mm. (L) MNHN IM-2007-41132, SANTO 2006, st. LD04, SL 23.8 mm.



of 3.75–4 slightly convex whorls; early three whorls smooth and glossy, latest whorl sculptured with 14–17 axial riblets (Fig. 10E). Protoconch diameter 0.93–1.12 mm, height 1.13–1.25 mm. Teleoconch whorls distinctly angulated; spire whorls sculptured with fine subsutural cord, and strong bifurcated sinus cord, and fine threads on subsutural area and whorl base. Apical whorl portion between subsutural cord and sinus cord distinctly concave. Shell base shortly constricted to slender siphonal canal. Shell base with eight to nine fine threads interchanging with sharp narrow spiral ridges, canal with 13–15 threads. Aperture elongate. Anal sinus wide and rather deep, quadrangular in its apex. Aperture usually with 9–12 distinct lirae inside. Background colour cream, with distinct dark-brown spots on subsutural and sinus cords. Brown spots on subsutural cords surrounded by somewhat nebulous lighter brown or reddish blotches. Shell base with indistinct light-brown band. Spiral threads with regular light-brown dots, protoconch light-brown; aperture cream inside. Radula examined in one sequenced specimen from New Ireland (MNHN IM-2013-53422, Fig. 5E). Radula membrane long, of about 50 rows of teeth, of which 20 not fully formed. Radula very similar to other congeners, with duplex marginal teeth. Anterior (inner) half solid, narrowly lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In posterior half, major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation indistinct.

**Remarks:** The species is represented in our material by eight specimens from Bismarck Sea (Madang lagoon and New Ireland), ranging in height from 24.5 to 32.1 mm, showing modest variation in conchological characters. The only feature that is found to vary notably is the shape of the anal sinus. It is moderately deep and wide with an angulated outline in specimen MNHN IM-2013-53422 (Fig. 10C, D), and even wider in Weinkauff's type, collected from the Philippines. The sinus is U-shaped, and very deep in some other sequenced specimens. Despite the fact that no specimens of *L. picturata* from the Philippines were sequenced in the present study, we confidently apply the name to this clade of our molecular tree, based on conchological features that are shared by the studied type specimen from ZMB and sequenced specimens. No other specimens of *L. picturata*, mentioned by Weinkauff, were studied. Since a species morphologically close to *L. picturata* – *L. bratasusa* sp. nov. – was recognized in our analysis, in order to fix the identity of *Lophiotoma picturata*, we here designate the studied syntype ZMB Moll 112610 as a lectotype, thereby setting the type locality as the Philippines. Morphologically *L. picturata* is very close to *L. bratasusa* sp. nov.; however, there

are some minor, but rather stable, characters that allow unmistakable differentiation of the two species. Firstly, the two species differ in the number of protoconch whorls – the former species has a protoconch of 3.75–4 whorls, while the latter – with 3.25 whorls only. Correspondingly the diameter and height of the protoconchs are slightly larger in *L. picturata* (Fig. 8). Shell proportions and coloration also offer some minute differences. *Lophiotoma picturata* is more turritiform in outline (due to comparatively shorter siphonal canal), and the black or dark brown spots on the subsutural region are surrounded by less contrasting light-brown or reddish blotches. On the contrary, *L. bratasusa* has a more fusiform outline, and the dark spots on the subsutural region are more contrasting in appearance. Weinkauff (1876 in Weinkauff & Kobelt, 1875–1887), when describing the species, cited *Pleurotoma variegata sensu* Reeve (1843), non Kiener (1840). The illustration of Reeve (1843: pl. 1, species 2) depicts the shell from the dorsal side which has a vague resemblance to *P. picturata*, although positive identification is hardly possible. Powell (1964) synonymized *Pleurotoma picturata* with *Lophiotoma acuta* and this viewpoint was accepted by subsequent authors.

**Distribution:** Confirmed distribution of the species (based on sequenced specimens) is Papua New Guinea. The lectotype was collected in the Philippines and the original type locality was 'Indian Ocean', so its range should be broader, but this needs confirmation.

#### **LOPHIOTOMA BRATASUSA SP. NOV.**

(FIG. 10G–L)

urn:lsid:zoobank.org:act:768A32A5-678A-4F03-8FC4-270EB8DE197C

**Type material:** Holotype MNHN IM-2013-51244, SL 26.0 mm; paratype 1, MNHN IM-2013-12566, paratype 2, MNHN IM-2013-53827.

**Type locality:** Papua New Guinea, Kavieng Lagoon, E of Kulinus I., Silver Sound, 02°42.3'S, 150°39.1'E, 7–10 m, coarse sand, coral patches (Expedition KAVIENG 2014, st. KR54).

**Etymology:** *bratasusa* – sibling in Pidgin English, refers to the revealed sister relationship between the new species and morphologically similar *L. picturata*. Used as a noun in apposition.

**Description (holotype):** Shell solid, narrow fusiform with high spire and rather long siphonal canal. Protoconch of 3.25 slightly convex whorls. Earlier 2.75 whorls smooth and glossy; latest 0.5 whorl sculptured with fine arcuate riblets, widely set at earlier portion

and more dense at transition to teleoconch. Protoconch diameter 0.89 mm, height 1.13 mm. Teleoconch of nine angulated whorls, suture shallow and inconspicuous. Subsutural region distinctly concave; suture immediately bordered by fine thread, followed by typically low subsutural cord, and three to seven regularly set spiral threads. Sinus cord bifurcated, formed by two subequal ridges on early whorls, adapical ridge notably stronger on penultimate and last teleoconch whorls. Abapical whorls portion (=whorl's base) sculptured with four fine threads, fourth slightly stronger than preceding. Shell base shortly constricted to slender siphonal canal, sculpture of shell base of 11 fine threads, fourth and sixth elevated to form sharp spiral ridges. Siphonal canal with 15 threads, spirally oriented and widely set adapically and dense, weakly delineated from one-another and oblique towards canal's tip. Aperture elongate; outer aperture lip convex adapically, rounded in side view. Anal sinus typically deep and rather narrow with rounded apex. Aperture smooth inside, or bearing 8–9 weak lirae. Background colour cream, with distinct contrast dark-brown spots on subsutural and smaller dots on sinus cords. Spiral threads with regular light-brown dots, giving them appearance of dashed lines. Protoconch orange; inside of aperture cream. Radula (holotype) (Fig. 5F) long, of about 55 rows of teeth, of which 25 nascent. Radula very similar to other congeners, with duplex marginal teeth. Anterior (inner) half solid, narrowly lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In posterior half major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation absent.

*Remarks:* *Lophiotoma bratasusa* sp. nov. varies notably in shell shape, sculpture pattern and coloration. The two ridges of bisected sinus cord may be equally strong, subequal or differ notably, to the extent that the lower ridge is not stronger than succeeding spiral threads. Dark spots on the subsutural cord, typically well developed, may be lacking entirely in the light form (Fig. 10L), or on the contrary the light brown band on the shell base may be pronounced, and the tip of the siphonal canal coloured dark-brown (Fig. 10J). The species is undoubtedly closest to *L. picturata*, although some differences between the two exist (see remarks under *L. picturata*), of which key are the number of protoconch whorls (4 in *L. picturata* v. 3.25 in *L. bratasusa* sp. nov.) and the colour pattern on the subsutural cord (with extended lighter blotches in *L. picturata* or without in *L. bratasusa* sp. nov.).

*Distribution:* Confirmed distribution of the species (based on sequenced specimens) is Vanuatu and Papua New Guinea.

*LOPHIOTOMA JICKELII* (WEINKAUFF, 1875)

(Fig. 11)

*Pleurotoma jickelii* Weinkauff, 1875 in Weinkauff & Kobelt, 1875–1887: 20, pl. 4, figs 2, 3 (Massaua, Red Sea).

*Lophiotoma acuta* form *jickelii*. – Powell, 1964: 305, pl. 180, figs 14, 19.

*Type material:* Neotype MNHN IM-2013-13275 (here designated).

*Type locality:* Papua New Guinea, Tab Island, inner slope, 05°10.2'S, 145°50.3'E (Expedition PAPUA NIUGINI, st. PR42).

*Description (neotype):* Shell thin, fusiform (Fig. 11A–C), with high spire and long narrow siphonal canal very slightly inclined to left. Protoconch conical (Fig. 11D), of about 3.75 evenly convex whorls, posteriormost 0.75 whorl before transition to teleoconch with ten distinct arcuate ribs, more closely spaced towards transition to teleoconch. Protoconch diameter 1.0 mm, height 1.22 mm. Teleoconch whorls weakly angulated at shoulder, 10.5 in total. Suture moderately deep, distinct, subsutural region wide, distinctly concave. Subsutural cord distinct, narrow on upper four teleoconch whorls, rounded on top, with two additional angular ridges appearing in upper part of cord on 5th and subsequent whorls. Ridges become progressively stronger and on last whorl cord of three distinct sharp triangular in profile ridges, middle one most elevated. Subsutural region smooth on upper teleoconch whorls, with one spiral ridge appearing on third whorl, two on the fourth, three on fifth, up to six on the last whorl. Paired sinus cords strongest, separated by interspace four times wider than cords, broadly obtuse triangular in profile and of same strength on last whorl. On upper whorls both cords similar in size, very closely spaced on upper four whorls progressively broader spaced on later whorls. Base of spire whorls smooth on upper two whorls, with one spiral cord on the third to fourth whorl, two on fifth, and then fast enlarging in number up to 11, strongly different in size cords on penultimate whorl. Base of last whorl with 15 cords, five of which much more prominent; canal with 34 cords, becoming gradually lower anteriorly. Cords slightly nodulose on intersections with growth lines. Shell base sharply narrowing towards narrow and long nearly straight siphonal canal. Aperture pear shaped, strongly constricted posteriorly, with parietal callus producing distinct tooth, outer lip concave in upper part and strongly convex below shoulder, gradually passing into canal. Anal sinus deep, narrow, with nearly parallel sides, and nearly straight



**Figure 11.** *Lophiotoma jickelii* (Weinkauff, 1875). (A–D) Neotype, MNHN IM-2013-13275, PAPUA NIUGINI, st. PR42, SL 39.4 mm. (D) Lateral view of the protoconch. (E) MNHN IM-2013-54874, KAVIENG 2014, st. KR136, SL 36.2 mm. (F) MNHN IM-2007-41003, Vanuatu, SANTO 2006, SE corner of Espiritu Santo, 33.8 mm. (G) MNHN IM-2013-11537, PAPUA NIUGINI, st. PD08, SL 33.3 mm. (H) MNHN IM-2007-41144, SANTO 2006, st. AT38, SL 35.0 mm. (I) MNHN IM-2007-41182, SANTO 2006, st. ZR11, SL 32.9 mm. (J) MNHN IM-2009-7080, Mozambique, Inhaca Island, SL 29.4 mm. (K) MNHN IM-2009-29713, Vietnam, Nha Trang Bay, st. ND7, SL 44.9 mm. (L) MNHN IM-2013-12760, PAPUA NIUGINI, st. PR33, SL 37.9 mm. (M, N) Original illustration of the species (Weinkauff, 1875: pl. 4, figs 2, 3). (O) Specimen from Egypt, Brother Island, 10–35 m (collection of P. Stahlschmidt). All shells (except M, N) at the same scale.

posterior margin, parallel to shell axis; outer lip in side view rounded and opisthocline, stromboid notch well-defined. Shell light brown, protoconch and two first teleoconch whorls slightly darker. Subsutural cord(s) with light brown irregularly shaped spots. Sinus cords with narrow and irregularly spaced brown spots, minor spiral cords with spots sometimes chevron shaped and smaller flecks. Aperture light creamy, lirate deep inside. Measurements (neotype largest of

our specimens): SL 39.4 mm, AL (with canal) 19.8 mm, SW 10.7 mm. Radula (neotype) (Fig. 5G) long, of about 65 rows of teeth, of which 25 nascent. Radula similar to other congeners, with duplex somewhat stout marginal teeth. Anterior (inner) half is solid, lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In posterior half major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation distinct, of small sharp narrow cusp.



**Remarks:** The species is rather variable in terms of sculpture and coloration. All intermediate specimens can be found from very light, hardly speckled specimens from Vietnam (Fig. 11K) to very dark ones from Mozambique, similar to the dark form of *L. acuta* (Fig. 11J). Interestingly, the dark form was found only in Mozambique and the only two studied specimens from this region were dark. The degree of development of spiral cords (other than subsutural and sinus cords) can also be rather different: there can be as few as four subequal cords on the subsutural zone, up to six strongly unequal cords in the neotype. In all studied specimens, there are two or even sometimes three closely spaced cords immediately below the suture. On the contrary, in *L. acuta* and *L. semifala* sp. nov. the subsutural cord is single, sometimes with two much weaker additional threads running along it. This allows a reliable differentiation of *L. jickelii* from both *L. acuta* and *L. semifala* sp. nov. There seems to be geographically determined shell variability, with only dark forms sampled in Mozambique, and very light ones in Vietnam; however, very limited material available from the mentioned localities does not allow us to draw final conclusions. The species was for a long time considered to be a synonym of *L. acuta* (Powell, 1964: 305 and many others), or a Red Sea subspecies of *L. acuta*. The name was used as a valid one recently for specimens from the Philippines (Heralde *et al.*, 2007; Fedosov *et al.*, 2011), but its validity was never addressed from the viewpoint of taxonomy. The type of *Pleurotoma jickelii* Weinkauff, 1875 originated from C. Jickeli's collection, which is now partially stored in the Humboldt Museum, Berlin (<http://www.conchology.be/?t=9001&id=21727>). Nevertheless, the types were not found in the Berlin Museum, nor in SMF, where the material of some other Weinkauff species is kept. Therefore, we consider them to be lost. The species was described from Massawa (presently Eritrea) based on a beach-collected specimen. The illustration of Weinkauff & Kobelt (1875–1887): pl. 4, figs 2, 3) is a bit ambiguous and depicts the large shell (SL 53 mm) with poorly pronounced sinus cords and nearly straight sided bases of spire whorls, similar to those in our specimens. Powell (1964: pl. 180, fig. 19) illustrated a specimen of 'form *jickelii*' from the Red Sea very similar to ours and provided an adequate and accurate description of *Lophiotoma acuta* form *jickelii*. Finally, Verbinnen & Dirckx (2007) discussed the occurrence of *Lophiotoma acuta* in the Red Sea and the status of *L. acuta jickelii* (Weinkauff, 1875). They illustrated the shell of *acuta* (fig. 21) as well as two shells which represent *L. jickelii* (21a, 21b). We were able to examine one shell, collected in Egypt (Fig. 11O) and it, as well as specimens illustrated by Verbinnen & Dirckx, falls within

intraspecific variability of a single species as defined herein by molecular data. In the absence of sequenced material from the Red Sea and due to the confusing situation with the taxonomy of the species, we designate herein the specimen collected in Tab Island, Papua New Guinea, Madang Lagoon (Fig. 11A–C) as the neotype of *Lophiotoma jickelii*. The species is most similar to *Lophiotoma kina* sp. nov., found in Vanuatu and Papua New Guinea. For differences see the remarks for *Lophiotoma kina* sp. nov. The species can be readily distinguished from *L. acuta* by its less pronounced subequal sinus cords rounded on top, while in *L. acuta* the upper sinus cord is much more pronounced than the lower and both sinus cords have a sharp upper edge. *Lophiotoma jickelii* also differs from both *L. acuta* and *L. semifala* sp. nov. in that the subsutural cord is subdivided into several cords on the last and penultimate whorls in the former species while in the latter two it is uniform with a sharp upper edge and very weak additional ridges. The studied radula of *L. jickelii* has a broader anterior solid part of marginal teeth and a more pronounced cusp on the central formation.

**Distribution:** Confirmed distribution of the species (based on sequenced specimens) – tropical Indo-West Pacific from Mozambique to Vietnam, Philippines, Papua New Guinea, Vanuatu. Based on published data also the Red Sea.

#### *LOPHIOTOMA KINA* SP. NOV.

(Fig. 12)

urn:lsid:zoobank.org:act:BB330B09-7334-4F5D-B7FF-30F43737680A

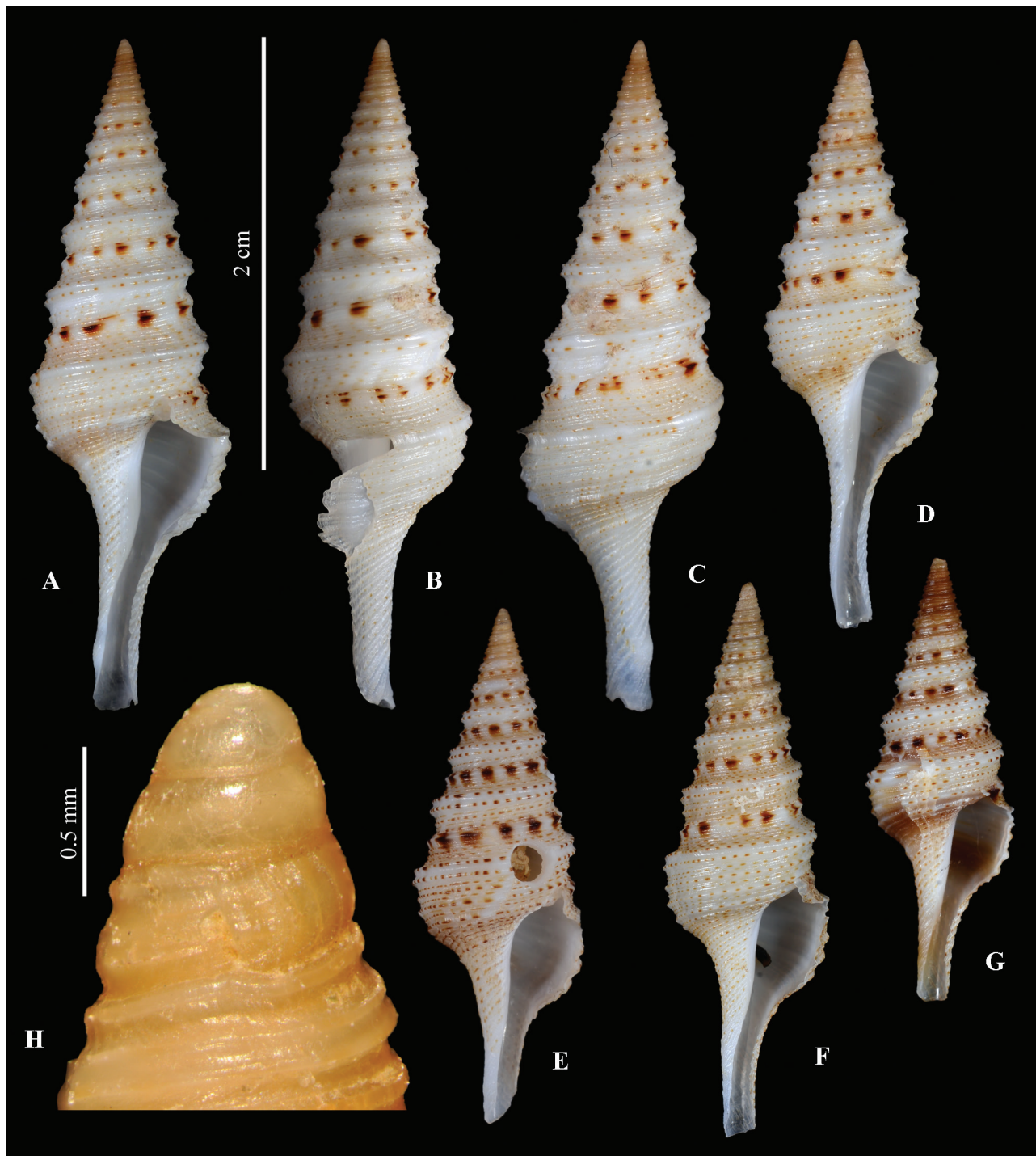
**Type material:** Holotype MNHN IM-2013-16307, paratype MNHN IM-2013-13278.

**Type locality:** Papua New Guinea, Madang Lagoon, W Tab Island, inner slope, 05°10.1'S, 145°50.2'E, 3–6 m (Expedition PAPUA NIUGINI, st. PR237).

**Etymology:** *kina* – the shell in Pidgin English, one of the official languages of Papua New Guinea. Used as a noun in apposition.

**Description (holotype):** Shell medium thick, fusiform, with high spire and long narrow siphonal canal very slightly inclined to left (Fig. 12A–C). Protoconch (intact in the specimen MNHN IM-2013-12950) conical, eroded of about 2.75 evenly convex whorls, posteriormost half whorl before transition to teleoconch with nine axial riblets (Fig. 12H). Protoconch diameter





**Figure 12.** *Lophiotoma kina* sp. nov. (A–C) Holotype, MNHN IM-2013-16307, PAPUA NIUGINI, st. PR237, SL 31.0 mm. (D) MNHN IM-2013-13278, PAPUA NIUGINI, st. PR42, SL 27.2 mm. (E) MNHN IM-2009-16927, SANTO 2006, st. FR10, SL 23.8 mm. (F) MNHN IM-2013-51209, KAVIENG 2014, st. KD13, SL 25.4 mm. (G–H) MNHN IM-2013-12950, PAPUA NIUGINI, st. PD23, SL 20.3 mm. (H) Lateral view of the protoconch. All shells at the same scale.

0.88 mm, height 0.93 mm. Teleoconch whorls weakly angulated at shoulder, 9.5 in total. Suture shallow, subsutural region wide, distinctly concave, subsutural

cord low, on upper five whorls narrow, rounded on top. On sixth whorl, additional angular ridge appearing in upper cord part, which becomes progressively

stronger and on last whorl cord consists of two distinct ridges, adapical one being twice lower than abapical ridge. Subsutural region smooth on upper teleoconch whorls, with one spiral ridge appearing on fourth, three on fifth, four on sixth and eight on last whorl. Paired sinus cords strongest, separated by interspace three times wider than cords, obtuse triangular in profile and nearly of same strength on last whorl. On early whorls both cords similar in size, with upper one being more pronounced on last and penultimate whorls. Base of spire whorls smooth on upper three whorls, with one spiral cord on fourth whorl, two on fifth, three on the sixth and seven on penultimate. Base of last whorl with three major spiral cords and two smaller ones between them, canal with 22 subequal cords, becoming gradually lower anteriorly. Shell base sharply narrowing towards narrow and long nearly straight siphonal canal. Aperture pear shaped, strongly constricted posteriorly with parietal callus producing distinct tooth, outer lip concave in upper part and weakly convex below shoulder, gradually passing into canal. Anal sinus deep, V-shaped, posterior margin nearly straight, parallel to shell axis; outer lip in side view rounded and opisthocline, stromboid notch well-defined. Growth lines indistinct, closely spaced. Shell light creamy, protoconch and three first teleoconch whorls slightly darker. Subsutural cord(s) with light brown irregularly shaped spots. Sinus cords with very weak light brown regularly spaced flecks, as well as minor spiral cords; spots occupying whole width of cord. Aperture light creamy, lirated deep inside. Measurements (holotype largest specimen): SL 31.0 mm, AL (with canal) 15.7 mm, SW 9.3 mm. Radula (Fig. 5I) is similar to other congeners, with duplex marginal teeth. Anterior (inner) half solid, narrowly lanceolate, dorso-ventrally compressed with sharp lateral cutting edges. In posterior half, major and accessory limbs bifurcate at about 45° angle, rather thin. Central formation was not examined due to radula preparation.

**Remarks:** The species is most similar to *L. jickelii* and can be distinguished by the more pronounced sinus cords and correspondingly more angulated whorls, generally less intensively coloured shell, with only very weak brown flecks on the sinus cords and other spiral elements. It also has a smaller protoconch (although the protoconch was available only in three specimens), consisting of 2.75–3 whorls in *L. kina* sp. nov. vs. 3.5–4.0 in *L. jickelii* (3.75 in most specimens) (Fig. 8).

**Distribution:** Confirmed distribution of the species (based on sequenced specimens) is Vanuatu and Papua New Guinea.

*LOPHIOTOMA VEZZAROII* COSSIGNANI, 2015

(FIG. 9F–G)

*Lophiotoma abbreviata*. – Okutani, 2000: pl. 313, fig. 54 (not of Reeve, 1843).

*Lophiotoma* cf. *ruthveniana* Melvill, 1923. – Poppe, 2008: pl. 683, fig. 4.

*Lophiotoma vezzaroii* Cossignani, 2015: 30–31, text figs.

**Type material:** Holotype MMM – Cupra Marittima.

**Type locality:** New Place Birat Samal Island, Philippines. Tangle net at 100–200 m.

**Material examined:** Three specimens sequenced (Table S1), one specimen Tinina Balut Island, Philippines, tangle net at 100–200 m.

**Diagnosis:** Shell medium sized (up to 39 mm), turritiform, with prominent spiral sculpture; shell coloured with dense irregularly shaped brown to dark brown spots, siphonal canal off-white to tan. Whorl outline moderately convex, angulated at sinus. Sinus cords paired, subequal in size, with narrow interspace. Subsutural ramp and shell base sculptured with varying in width and prominence cords and finer riblets. Shell base convex, strongly constricted to rather slender siphonal canal, sculptured with dense spiral to oblique cords. Aperture elongate, anal sinus moderately deep, wide, angulated at tip. Aperture distinctly lirated inside. Radula examined in one poorly preserved specimen from Tinina, Balut Island (Fig. 9G), in all respects similar to other studied herein species of *Lophiotoma*.

**Remarks:** The species was confused previously with *Lophiotoma ruthveniana*. Okutani (2000) illustrated a specimen very similar to *Lophiotoma abbreviata*. Although described from the Philippines, our material and the record of Okutani suggest that its distribution extends from Japan to Papua New Guinea and Vanuatu from 10 to 15 to more than 100 m depth.

**Distribution:** Vanuatu (sequenced specimens), Japan, Philippines and Papua New Guinea.

## DISCUSSION

Following an integrative taxonomic approach, we applied several criteria and methods of species delimitation to identify species boundaries within *Lophiotoma*. The three exploratory methods used (ABGD, GMYC and PTP) do not always agree on the species delimitation, but the use of other criteria and characters allowed the most robustly supported species partitions

to be chosen: the ten SSH retained are recognized as genetically (based on both distances and phylogenetic relationships) and morphologically distinct. The application of the GMYC ‘multiple’ method to our dataset resulted in notably oversplit partitions; a similar tendency was found, for example by [Kekkonen & Hebert \(2014\)](#). On the contrary, genes less variable than *COI*, such as *28S*, tend to cluster some partitions that are later found worthy of recognition as species.

Among the ten delimited species, species in three pairs (*L. acuta* – *L. semfala* sp. nov., *L. picturata* – *L. bratasusa* sp. nov. and *L. jickelii* – *L. kina* sp. nov.) are barely distinguishable morphologically: without molecular evidence they would hardly be suspected to be separate species. Moreover, the intraspecific morphological variability exceeds the interspecific variability, particularly in shell coloration, with the presence of ‘light’ and ‘dark’ forms within each species of the pairs *L. acuta* – *L. semfala* sp. nov. and *L. jickelii* – *L. kina* sp. nov. The radular characters that sometimes can be useful for species delimitation ([Kantor et al., 2008](#)) were of no help in the case of *Lophiotoma*. All examined species had extremely similar radular morphology, and only in one species, *L. jickelii* the central formation of the radula in the shape of a weak but distinct cusp was observed, while in all others it was either absent, or indistinct. However, we confidently recognize them as distinct species, based on the following considerations: (1) both genes recognized them as six distinct clades, (2) only in two cases (with the *COI* gene for *L. picturata* and *L. bratasusa* sp. nov. and with the *28S* gene for *L. acuta* and *L. semfala* sp. nov.) are ‘morphological pairs’ also recovered as sister species, and (3) remarkably, morphologically similar species always occur sympatrically, which tends to support the hypothesis of genetic isolation between species in ‘morphological pairs’.

The integrative taxonomic approach followed here was thus efficient to propose robust species hypotheses. It represents one additional example of the value of molecular characters when species can hardly be distinguished morphologically, a common situation in gastropods ([Jörger & Schrödl, 2013](#)) and, in particular, in conoideans (e.g. [Duda et al., 2008](#); [Puillandre et al., 2010](#)). However, if proposing putative species using DNA sequences is now common, linking the SSH to available names, most often attached to nonsequenced type specimens, remains problematic. Until now, all species of *Lophiotoma* have been described using conchological characters only. Moreover, locating type specimens to tentatively attribute their associated names to the defined SSH, based on morphological resemblance, was probably the most difficult task. Among the seven species already described before the present work, the type

material was located for four species only (*L. abbreviata*, *L. brevicaudata*, *L. picturata* and *L. vezzaroi*). For *L. vezzaroi*, the holotype was properly designated, and for *L. abbreviata* and *L. brevicaudata* lectotypes were designated in previous studies. For the last one, *L. picturata*, we located two syntypes and designated a lectotype. For the three other species, *L. acuta*, *L. jickelii* and *L. polytropha*, we were unable to locate the type material (see details in the taxonomic section). In the absence of type material, we had to rely on the illustrations in the original descriptions to link the SSH to these names. For *L. acuta* and *L. jickelii*, because these names are associated with species complexes that include morphologically similar species, we choose one of the sequenced specimens in each species as a neotype. We also examined, when possible, the type material and/or the original illustrations of the species synonymized with *L. acuta* in the literature, and concluded that none of these names can confidently be attributed to one of the three remaining SSH. Consequently, we described these three SSH as new species: *L. semfala* sp. nov., *L. bratasusa* sp. nov. and *L. kina* sp. nov.

More generally, most species of molluscs were described before the molecular revolution, and the identity of most newly described species still remains based on dry material and/or nonsequenced specimens ([Bouchet & Strong, 2010](#)). When dealing with species complexes, attributing names to molecular groups is thus tricky. When the type specimens are lost, designating a sequenced specimen as a neotype solves the problem. However, when the types are still available, only morphological resemblance can be used to decide to which of the molecular groups the name will be attributed. We applied this strategy earlier to unravel relationship between forms of *Xenuroturris cingulifera* (Lamarck, 1822), for which molecular analyses and studies of radula revealed presence of two species that were very similar conchologically ([Kantor et al., 2008](#)). An available name *Xenuroturris legitima* Iredale, 1929 was unearthed in the synonymy of *Iotyrris cingulifera* and applied to the nontypical form of *Xenuroturris cingulifera*. This decision was enabled by the fact that the types of the two mentioned names persist, both as empty shells, badly worn in the case of the Lamarck’s syntypes, and shell similarity helped to attribute each name to one of the two molecular groups. This solution was practical, because it allowed a stabilization of the nomenclature without the designation of neotypes, which in case of persisting types requires lengthy consideration by the Commission of Zoological Nomenclature. Here we used the same approach for *L. picturata*: we applied the name *picturata* to the species that was morphologically more similar to the lectotype.



Because type specimens remain the only way to unambiguously link names and genetic groups, one could suggest that sequencing type specimens, when available, is the ultimate solution. Traditionally, shell-bearing mollusc types are kept dried in collections, which does not ensure a correct DNA conservation. Recently developed NGS techniques would clearly help to sequence fragmented DNA, but a high proportion of name-bearing types are empty shells, which do not contain any remains of tissue inside. Recently published articles (Geist, Wunderlich & Kuehn, 2008; Andree & López, 2013; Villanea, Parent & Kemp, 2016) suggest that DNA can actually be extracted from shells, but whether such techniques are applicable to specimens kept dried for tens, or even hundreds of years, and for which the periostracum is potentially absent, remains to be tested. It also implies that a piece of the shell (Andree & López, 2013) of the holotype will be destroyed, a condition that will need to be accepted by museum curators.

In any case, providing DNA sequences should become the gold standard in species delimitation and description in groups where morphological characters are misleading, such as in Conoidea, to avoid erroneous species hypotheses based on shell characters only and to facilitate the attribution of names to genetic sequences in the future. Even if sequencing DNA from empty shells seem conceivable, it is difficult to imagine that all types of shelled molluscs will be sequenced in the future (for technical and financial reasons), and in most cases linking these names to molecular groups will be subject to controversy. Paraphrasing Marshall (1983), who said that ‘under absolutely no circumstances should further new species [of Triphoridae] be proposed unless a complete, unworn protoconch can be illustrated,’ ideally, under absolutely no circumstances should further new species of turrids be proposed without any molecular data.

#### ACKNOWLEDGEMENTS

A large part of the molecular material in this paper originates from various shore-based expeditions and deep sea cruises, conducted, respectively, by MNHN (Panglao 2004, Aurora 2007, Inhaca 2011); by MNHN, Pro-Natura International (PNI) and Institut de Recherche pour le Développement (IRD) as part of the Our Planet Reviewed programme (SANTO 2006, PAPUA NIUGINI, KAVIENG 2014); and by MNHN and Institut de Recherche pour le Développement (IRD) as part of the Tropical Deep-Sea Benthos program. In-country partners include the University of San Carlos, Cebu City; Maritime College, Luganville; Universidade Eduardo Mondlane, Maputo; University of Papua New Guinea, Port Moresby, and National

Fisheries College, Kavieng. Funders and sponsors include the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, Richard Lounsbery Foundation, Vinci Entrepose Contracting, Fondation EDF, the Philippines Bureau of Fisheries and Aquatic Resources (BFAR), the French Ministry of Foreign Affairs, the French Fonds Pacifique and the Government of New Caledonia. 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. Collection of material in Vietnam was supported by the Russian–Vietnamese Tropical Center. We are thankful to the staff of the Tropical Center for assistance in organization of the field sampling and loan of some laboratory equipment. We thank, among others, Virginie Héros, Philippe Maestrati, Barbara Buge, Julien Brisset and Gilberto Marani for their role in specimens processing during the expeditions and their help in curating the specimens. The photo of the type of *Pleurotoma tigrina* was kindly provided by Emmanuel Tardy (MHNG), and the photo of *Lophiotoma jickelii* from Egypt by Peter Stahlschmidt. We thank Catherine Rausch (MNHN) and Nadezda Surovenkova (IEE) for kindly granting access to SEM facilities. We are grateful to Marco Oliverio, Louise Allcock and an anonymous reviewer for their critical comments on the manuscript.

This project was partly supported by the Service de Systématique Moléculaire (UMS 2700 CNRS-MNHN), by the CONOTAX project funded by the French ANR (grant number ANR-13-JSV7-0013-01), and grant of the Russian Science Foundation (16-14-10118, PI Yu. Kantor).

#### REFERENCES

- Andree KB, López MA. 2013.** Species identification from archived snail shells via genetic analysis: a method for DNA extraction from empty shells. *Molluscan Research* **33**: 1–5.
- Bouchet P, Kantor YI. 2004.** New Caledonia: the major centre of biodiversity for volutomitrid molluscs (Mollusca: Neogastropoda: Volutomitridae). *Systematics and Biodiversity* **1**: 467–502.
- Bouchet P, Strong E. 2010.** Historical name-bearing types in marine molluscs: an impediment to biodiversity studies? In: Polaszek A, ed. *Systema naturae*, Vol. **250**. London: CRC Press, 63–74.
- Brown SDJ, Collins RA, Boyer S, Lefort MC, Malumbres-Olarte J, Vink CJ, Cruickshank RH. 2012.** Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Molecular Ecology Resources* **12**: 562–565.
- Casey TL. 1904.** Notes on the Pleurotomidae with description of some new genera and species. *Transactions of the Academy of Science of St. Louis* **14**: 123–170.

- Cernohorsky WO. 1972.** *Marine shells of the Pacific, Vol. 2.* Sydney: Pacific Publications.
- Cossignani T. 2015.** *Lophiotoma vezzaroi* sp. nov. *Malacologica Mostra Mondiale. Cupra Marittima* **88**: 30–31.
- Dance SP. 1986.** *A history of shell collecting.* Leiden: Brill and Backhuys.
- Dayrat B. 2005.** Towards integrative taxonomy. *Biological Journal of the Linnean Society* **85**: 407–415.
- De Queiroz K. 2007.** Species concepts and species delimitation. *Systematic Biology* **56**: 879–886.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.
- Duda TF Jr, Bolin MB, Meyer CP, Kohn AJ. 2008.** Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. *Molecular Phylogenetics and Evolution* **49**: 867–876.
- Dunker W. 1871.** Mollusca nova Musei Godeffroy Hamburgensis. *Malakozoologische Blätter* **18**: 150–175.
- Edgar RC. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.
- Fedosov A, Watkins M, Heralde FM III, Showers Corneli P, Concepcion GP, Olivera BM. 2011.** Phylogeny of the genus *Turris*: correlating molecular data with radular anatomy and shell morphology. *Molecular Phylogenetics and Evolution* **59**: 263–270.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994.** DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**: 294–299.
- Galindo LA, Puillandre N, Strong EE, Bouchet P. 2014.** Using microwaves to prepare gastropods for DNA barcoding. *Molecular Ecology Resources* **14**: 700–705.
- Geist J, Wunderlich H, Kuehn R. 2008.** Use of mollusk shells for DNA-based molecular analyses. *Journal of Molluscan Studies* **74**: 337–343.
- Goldstein PZ, DeSalle R. 2011.** Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *Bioessays* **33**: 135–147.
- Helbling S. 1779.** Beiträge zur Kenntnis neuer seltener Konchylien. *Abhandlungen einer Privatgesellschaft in Böhmen, zur Aufnahme der Mathematic, der Vaterland, Bischen Geschichte, und der Naturgeschichte.* **4**: 102–131.
- Heralde FM, Watkins M, Ownby JP, Bandyopadhyay PK, Santos AD, Concepcion GP, Olivera BM. 2007.** Molecular phylogeny of some Indo-Pacific genera in the subfamily Turridae H. Adams and A. Adams, 1853 (1838) (Gastropoda: Neogastropoda). *Nautilus* **121**: 131–138.
- Higo S, Callomon P, Gotō Y. 1999.** *Catalogue and bibliography of the marine shell bearing mollusca of Japan: Gastropoda, Bivalvia, Polyplacophora, Scaphopoda.* Yao, Japan: Elle Scientific Publications.
- Huelsenbeck JP, Ronquist F, Hall B. 2001.** MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17**: 754–755.
- Jörger KM, Schrödl M. 2013.** How to describe a cryptic species? Practical challenges of molecular taxonomy. *Frontiers in Zoology* **10**: 59.
- Jovelin R, Justine JL. 2001.** Phylogenetic relationships within the polyopisthocotylean monogeneans (Platyhelminthes) inferred from partial 28S rDNA sequences. *International Journal for Parasitology* **31**: 393–401.
- Kantor YI. 2006.** On the morphology and homology of the ‘central tooth’ in the radula of Turridae (Conoidea: Turridae). *Ruthenica* **16**: 47–52.
- Kantor YI, Puillandre N. 2012.** Evolution of the radular apparatus in Conoidea (Gastropoda: Neogastropoda) as inferred from a molecular phylogeny. *Malacologia* **55**: 55–90.
- Kantor YI, Puillandre N, Olivera BM, Bouchet P. 2008.** Morphological proxies for taxonomic decision in turrids (Mollusca, Neogastropoda): a test of the value of shell and radula characters using molecular data. *Zoological Science* **25**: 1156–1170.
- Kekkonen M, Hebert PD. 2014.** DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Molecular Ecology Resources* **14**: 706–715.
- Kiener L. 1840.** *Species general et iconographie des Coquilles vivantes Comprenant la Collection du Muséum d’Histoire Naturelle de Paris, Collection Lamarck, celle du Prince Masséna et les Découvertes Récentes des Voyageurs (Genre Pleurotome).* Paris: Rousseau.
- Kilburn RN. 1983.** Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 1. Subfamily Turridae. *Annals of the Natal Museum* **25**: 549–585.
- Lamarck JBPA. 1816.** *Tableau encyclopédique et méthodique des trois règnes de la nature. Vingt troisième partie. Mollusques et polypes divers.* Paris: Veuve Agasse, Imprimeur.
- Lamarck JBPA. 1816.** Liste des objets représentés dans les planches de cette livraison. In Lamarck JBPA. *Tableau Encyclopédique et Méthodique des Trois Règnes de la Nature. Mollusques et Polypes Divers.* Paris: Veuve Agasse, Imprimeur.
- Lamarck JBPA. 1822.** *Histoire naturelle des animaux sans vertèbres.* 6(2). Paris: Verdrière.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017.** PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**: 772–773.
- Lozouet P, Plaziat JC. 2008.** *Mangrove environments and molluscs.* Hackenheim: Conchbooks.
- Marshall B. 1983.** A revision of the Recent Triphoridae of southern Australia (Mollusca: Gastropoda). *Records of the Australian Museum, Supplement* **2**: 1–119.
- Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJ, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP. 2009.** Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology* **58**: 298–311.
- Ober KA. 2002.** Phylogenetic relationships of the carabid subfamily *Harpalinae* (Coleoptera) based on molecular sequence data. *Molecular Phylogenetics and Evolution* **24**: 228–248.

- Okutani T. 2000.** *Marine mollusks in Japan*. Tokyo, Japan: Tokai University Press.
- Olivera BM. 2004.** Larger forms in *Lophiotoma*: four new species described in the Philippines and three from elsewhere in the Indo-Pacific. *Science Diliman* **16**: 1–28.
- Oyama K. 1966.** On living Japanese Turridae (1). *Venus* **25**: 1–20.
- Pante E, Puillandre N, Viricel A, Arnaud-Haond S, Aurelle D, Castelin M, Chenuil A, Destombe C, Forcioli D, Valero M, Viard F, Samadi S. 2015.** Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology* **24**: 525–544.
- Pante E, Schoelincx C, Puillandre N. 2015.** From integrative taxonomy to species description: one step beyond. *Systematic Biology* **64**: 152–160.
- Perry G. 1811.** *Conchology, or the natural history of shells: containing a new arrangement of the genera Ans species, illustrated by coloured engravings, executed from the natural specimens, and including the latest discoveries*. London: W. Bulmer & Co.
- Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006.** Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55**: 595–609.
- Poppe GT. 2008.** *Philippine marine mollusks, Vol. II*. Hackenheim, Germany: Conchbooks.
- Powell AWB. 1964.** The family Turridae in the Indo-Pacific. Part I. The subfamily Turrinae. *Indo-Pacific Mollusca* **1**: 227–411.
- Puillandre N, Lambert A, Brouillet S, Achaz G. 2012a.** ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* **21**: 1864–1877.
- Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S. 2012b.** Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* **21**: 2671–2691.
- Puillandre N, Sysoev A, Olivera BM, Couloux A, Bouchet P. 2010.** Loss of planktotrophy and speciation: geographical fragmentation in the deep-water gastropod genus *Bathytoma* (Gastropoda, Conoidea) in the western Pacific. *Systematics and Biodiversity* **8**: 371–394.
- Rambaut A, Drummond AJ. 2014.** *Tracer v1.6*. Available at: <http://beast.bio.ed.ac.uk/Tracer>
- Reeve L. 1843.** Monograph of the genus *Pleurotoma*. *Conchologia Iconica* **1**: plates 1–18.
- Renner SS. 2016.** A return to Linnaeus's focus on diagnosis, not description: the use of DNA characters in the formal naming of species. *Systematic Biology* **65**: 1085–1095.
- Samadi S, Barberousse A. 2009.** Species: towards new, well-grounded practices. A response to Velasco. *Biological Journal of the Linnean Society* **96**: 696–708.
- Schubert G, Wagner J. 1829.** Neues systematisches Conchylien-Cabinet angefangen von Martini und Chemnitz. *Nürnberg, Bauer & Raspe* **12**: 1–196.
- Springsteen FJ, Leobrera FM. 1986.** *Shells of the Philippines*. Manila, Philippines: Carfel Seashell Museum.
- Srivathsan A, Meier R. 2012.** On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. *Cladistics* **28**: 190–194.
- Stamatakis A. 2006.** RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.
- Verbinnen G, Dirckx M. 2007.** Red Sea Mollusca. Part 17. Class: Gastropoda. Family: Turridae. *Gloria Maris* **43**: 7–27.
- Villanea FA, Parent CE, Kemp BM. 2016.** Reviving Galápagos snails: ancient DNA extraction and amplification from shells of probably extinct Galápagos endemic land snails. *Journal of Molluscan Studies* **82**: 449–456.
- Weinkauff H, Kobelt W. 1875–1887.** Die Familie Pleurotomidae. *Systematisches Conchylien-Cabinet von Martini und Chemnitz* **4**: 1–248.
- Will KW, Mishler BD, Wheeler QD. 2005.** The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* **54**: 844–851.
- Zhang J, Kapli R, Pavlidis P, Stamatakis A. 2013.** A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29**: 2869–2876.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Table S1.** List of specimens analysed, with MNHN number, species name, geographic locality, depth, and BOLD and GenBank accession numbers.

**Appendix 1.** COI alignment.

**Appendix 2.** 28S alignment.

## **ANNEXE 2 :**

Abdelkrim J., Aznar-Cormano L., Fedosov A.E., Kantor Y.I., Lozouet P., Phuong M.A., **Zaharias P.**, Puillandre N. (2018a). Exon-capture based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Mol. Biol. Evol.* 35:2355–2374.



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

Jawad Abdelkrim,<sup>1,2</sup> Laetitia Aznar-Cormano,<sup>1,2</sup> Alexander E. Fedosov,<sup>3</sup> Yuri I. Kantor,<sup>3</sup> Pierre Lozouet,<sup>4</sup> Mark A. Phuong,<sup>5</sup> Paul Zaharias,<sup>2</sup> and Nicolas Puillandre<sup>\*2</sup>

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

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

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

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

<sup>5</sup>Department 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  $\sim 10$ – $20,000$ —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

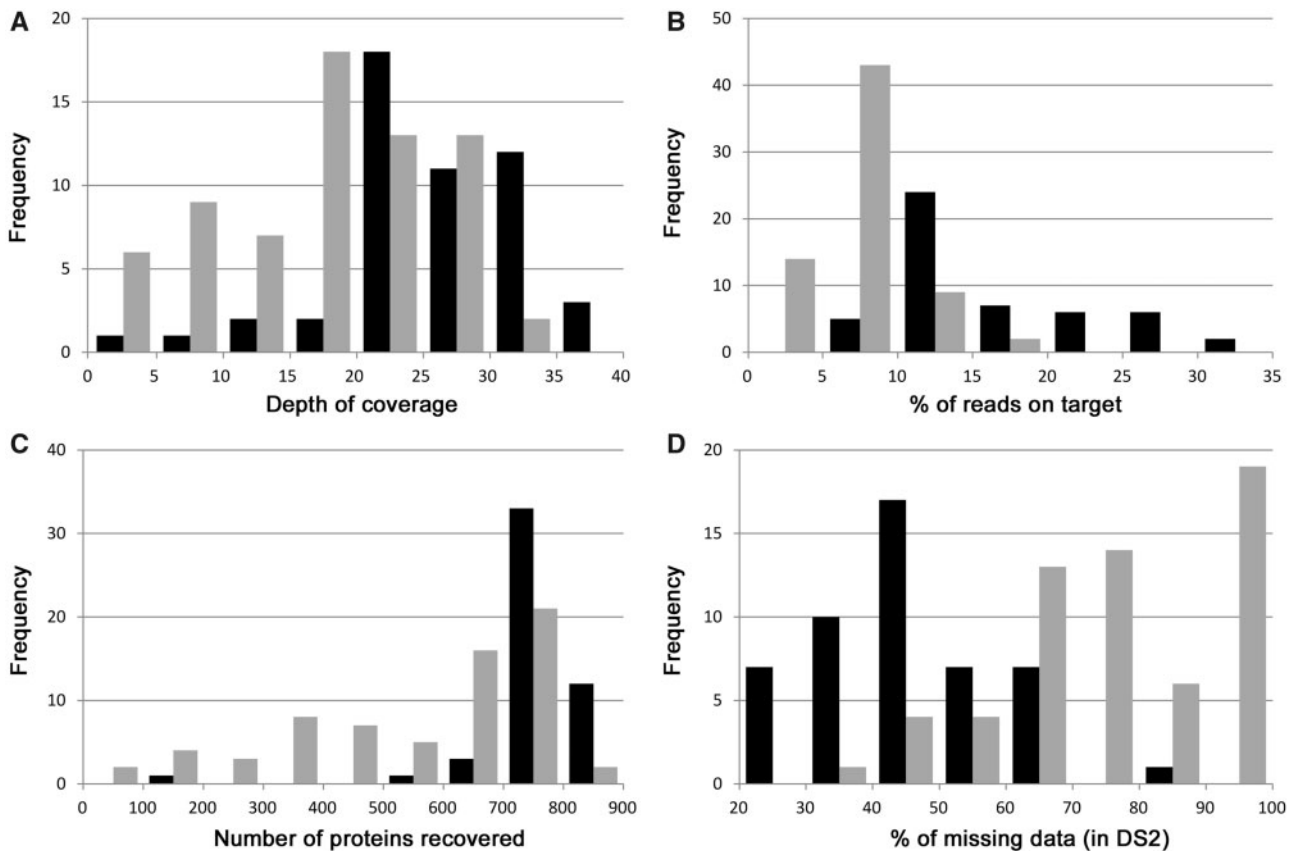
We used available Conoidean transcriptome data from the literature and online resources to identify 850 protein coding genes. We tentatively captured 120 samples representing the diversity of the Conoidea, as well as a selection of non-Conoidean neogastropod lineages. Two samples failed at the library 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 *Cruziturricula 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 (*Heteroturris 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+Fascioliidae being sister group to the clade Mitridae+Pyramimitridae+Conoidea (100/.73/.83).

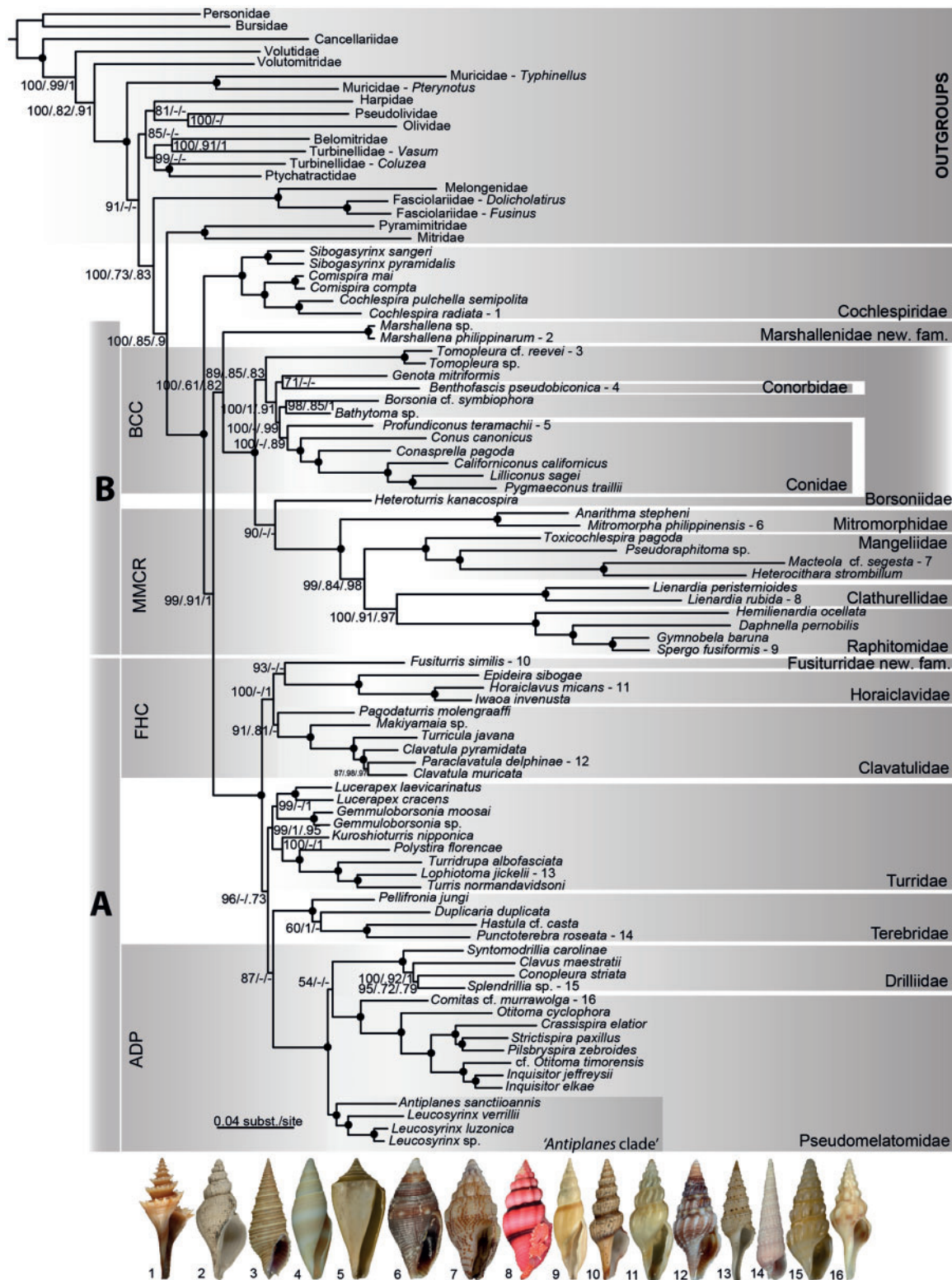
We found Conoidea to be monophyletic, with Cochlespiridae being sister group to the rest of the superfamily (99/.91/1), further divided into two main clades. The first one (100/0.61/0.82) includes the Marshallenidae new fam., sister group to a clade including on one hand the MMCR clade, and on the other 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. *Heteroturris* (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 *Heteroturris*, 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, *Kurishioturris* 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



**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.

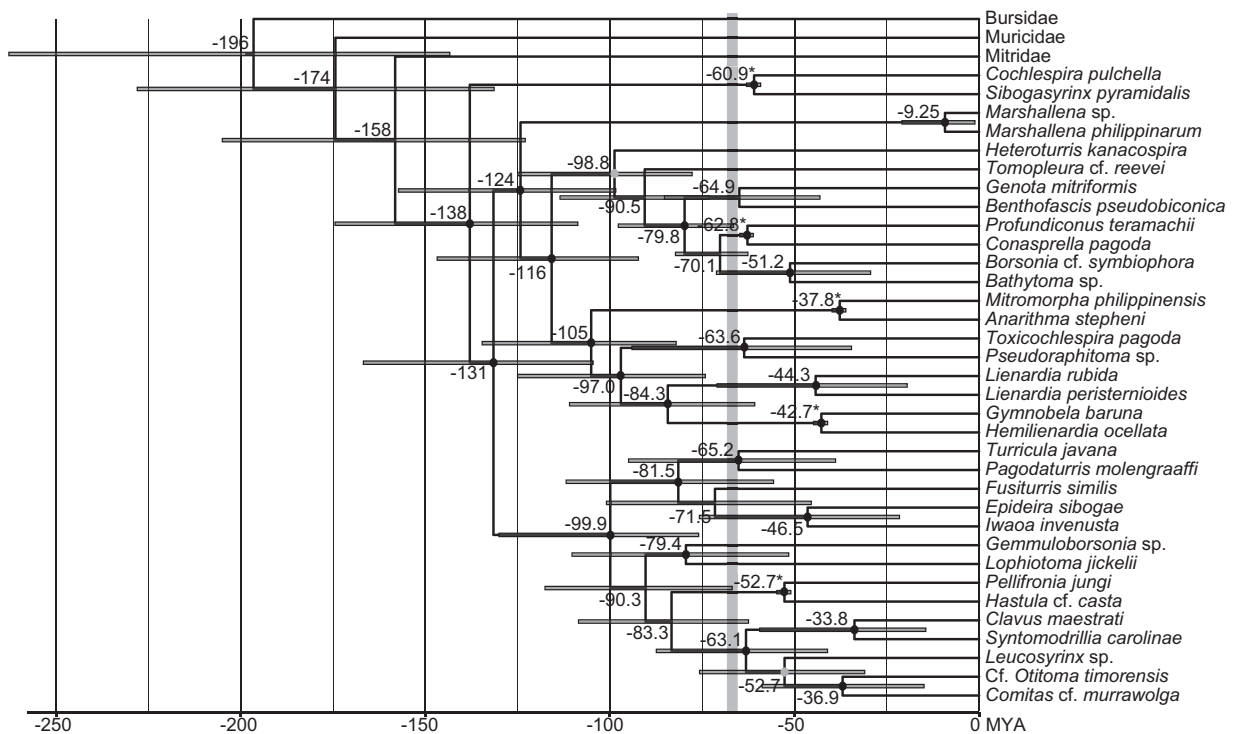
Marshallenidae, Fusiturridae, and Conorbidae were detected as significantly decreasing. The relationships between the diversification rates and the venom gland was never significant, irrespective of the number of species and the extinction rate considered. Among the five pairs of clades for which one

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.

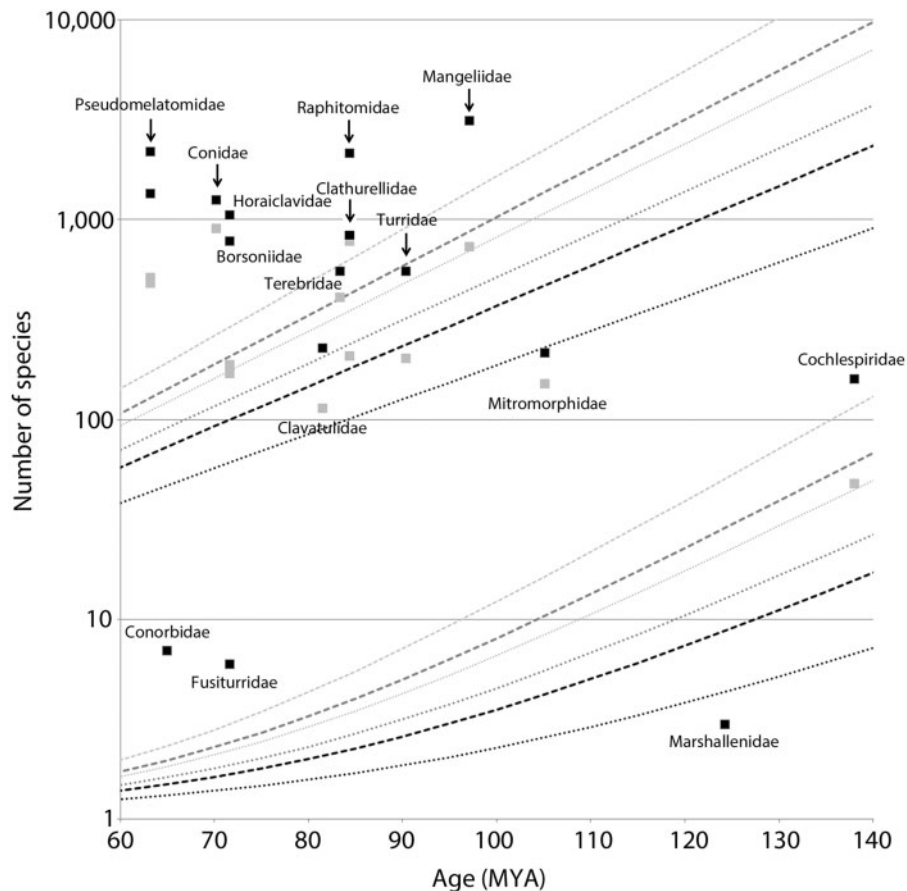
	DS2aML	DS2bML	DS3aML	DS3bML	DS3aAS	DS3bAS	DS3aWSB	DS3bWSB
CONOIDEA+MITRIDAE+PYRAMIMITRIDAE	49	100	100	100	0.89	0.85	0.57	0.9
CONOIDEA	100	100	100	100	1	1	1	1
COCHLESPIRIDAE	100	100	100	100	1	1	1	1
A+B	97	100	100	99	0.94	0.91	1	1
B	100	100	100	100	0.87	0.61	0.99	0.82
MARSHALLENIDAE	100	100	100	100	1	1	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

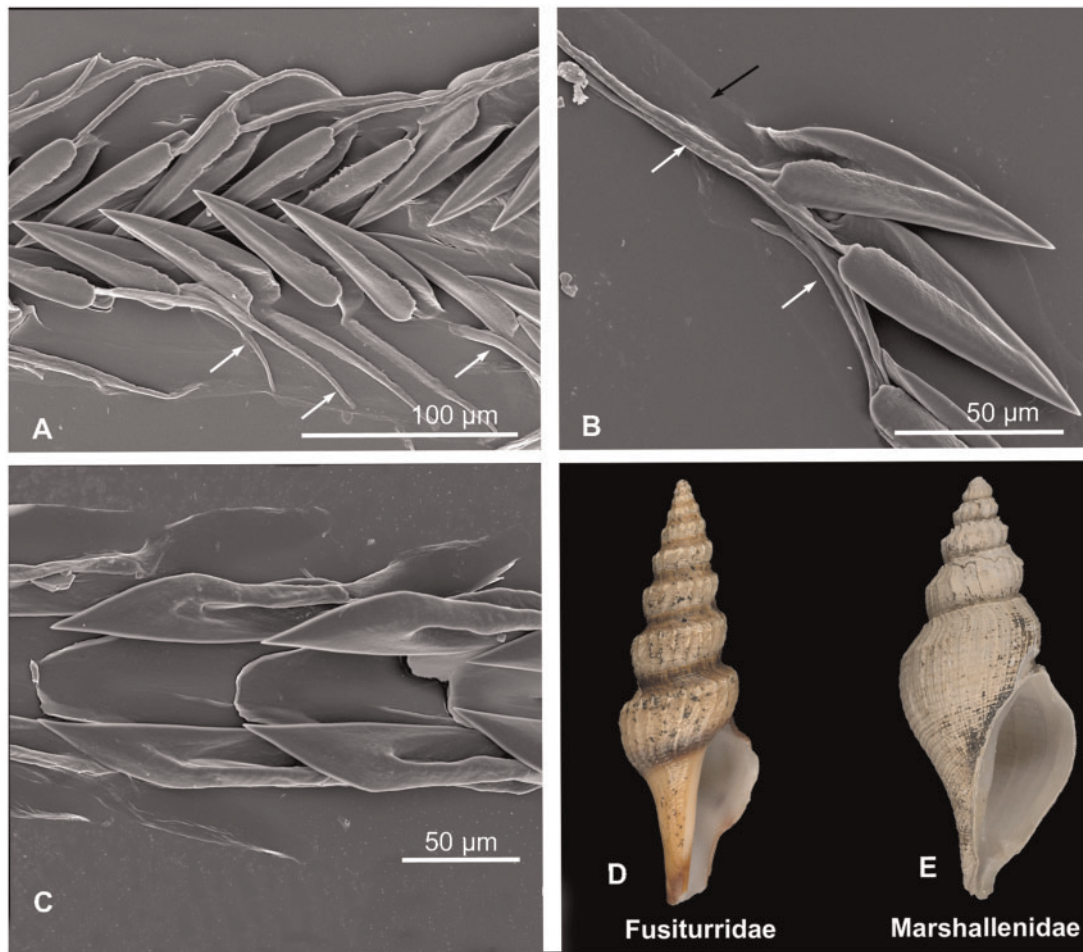
**Table 3.** Results of the PGLS Analysis, with the Correlation Coefficient and the  $P$  Value for Each Comparison.

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)—Marshalleniidae, (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).

Marshalleniidae 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

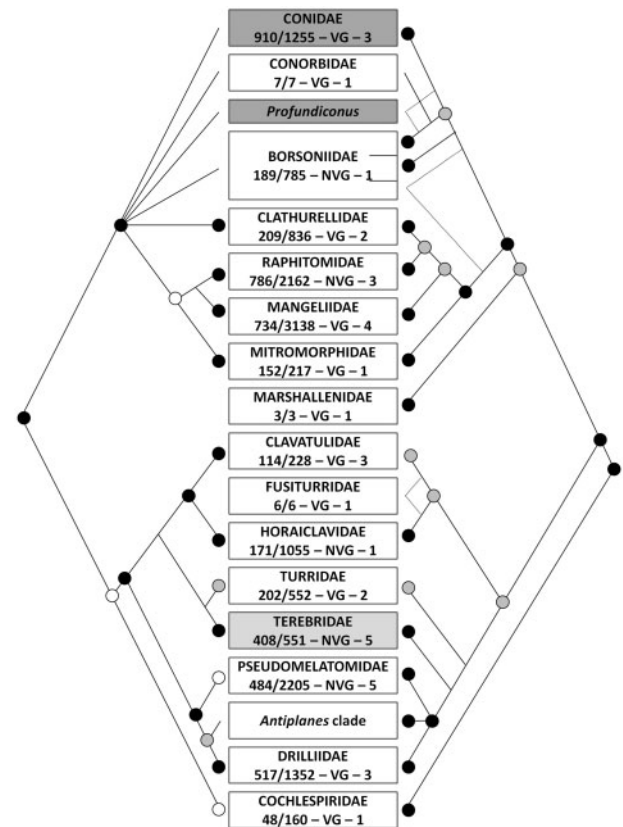


data set with 60 samples minimum per locus, in the ML and WSB trees. The differences in the phylogenetic trees were also limited between the data sets DS2 and DS3, the later (i.e., without the samples with >90% of missing data) resulting in a few better supported nodes (table 2).

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 ANT FLOCKS 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.



### **ANNEXE 3 :**

Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., **Zaharias, P.**, & Puillandre, N. (2018b). Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Molecular ecology*, 27(22), 4591-4611.

# Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework

Jawad Abdelkrim<sup>1,2</sup> | Laetitia Aznar-Cormano<sup>1</sup> | Barbara Buge<sup>3</sup> | Alexander Fedosov<sup>4</sup> | Yuri Kantor<sup>4</sup> | Paul Zaharias<sup>1</sup> | Nicolas Puillandre<sup>1</sup> 

<sup>1</sup>Institut de Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Paris, France

<sup>2</sup>Service de Systématique Moléculaire SSM-UMS2700 - Muséum National d'Histoire Naturelle, Paris, France

<sup>3</sup>Muséum National d'Histoire Naturelle, Paris, France

<sup>4</sup>A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia

## Correspondence

Nicolas Puillandre, Institut de Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 26, 75005 Paris, France.  
Email: puillandre@mnhn.fr

## Funding information

Russian Science Foundation, Grant/Award Number: 16-14-10118; Agence Nationale de la Recherche, Grant/Award Number: ANR-13-JSV7-0013-01

## Abstract

Species delimitation in poorly known and diverse taxa is usually performed based on monolocus, DNA-barcoding-like approaches, while multilocus data are often used to test alternative species hypotheses in well-studied groups. We combined both approaches to delimit species in the *Xenuroturris/lotyrris* complex, a group of venomous marine gastropods from the Indo-Pacific. First, COI sequences were analysed using three methods of species delimitation to propose primary species hypotheses. Second, RAD sequencing data were also obtained and a maximum-likelihood phylogenetic tree produced. We tested the impact of the level of missing data on the robustness of the phylogenetic tree obtained with the RAD-seq data. Alternative species partitions revealed with the COI data set were also tested using the RAD-seq data and the Bayes factor species delimitation method. The congruence between the species hypotheses proposed with the mitochondrial nuclear data sets, together with the morphological variability of the shell and the radula and the distribution pattern, was used to turn the primary species hypotheses into secondary species hypotheses. Allopatric primary species hypotheses defined with the COI gene were interpreted to correspond to intraspecific structure. Most of the species are found sympatrically in the Philippines, and only one is confidently identified as a new species and described as *lotyrris conotaxis* n. sp. The results obtained demonstrate the efficiency of the combined monolocus/multilocus approach to delimit species.

## KEYWORDS

DNA barcoding, integrative taxonomy, RAD-seq, species delimitation, Turridae

## 1 | INTRODUCTION

The last decade has seen a burst of methods available to propose species hypotheses based on molecular data (Carstens, Pelletier, Reid, & Satler, 2013; Ence & Carstens, 2011; Fujisawa & Barraclough, 2013; Leaché, Fujita, Minin, & Bouckaert, 2014; Leavitt, Moreau, & Lumbsch, 2015; Puillandre, Lambert, Brouillet, & Achaz, 2012; Yang & Rannala, 2014; Zhang, Kapli, Pavlidis, & Stamatakis, 2013). Genetic data collection for species delimitation can be separated into two

general approaches: (a) DNA barcoding, in which a high number of species/specimens are sequenced for one or a few markers, and (b) deep analyses of species complexes, where a limited number of species/specimens are analysed with larger genetic data sets (ranging from 5–6 markers to several thousand markers). These two approaches lend themselves to particular classes of methods. For example, species delimitation methods based on monolocus data, such as GMYC (General Mixed Yule Coalescent method; Fujisawa & Barraclough, 2013; Monaghan et al., 2009; Pons et al., 2006),

PTP (PoissonTree Process; Zhang et al., 2013) or ABGD (Automatic Barcode Gap Discovery; Puillandre, Lambert, et al., 2012) are typically applied to taxa with no or limited genomic data available (i.e., understudied and/or hyperdiverse groups). For such groups, sequencing many markers for many specimens and species can be problematic. For example, there may be financial constraints that limit genetic sequencing (but see e.g., Coissac, Hollingsworth, Lavergne, & Taberlet, 2016), but also because the lack of genomic information makes the identification of suitable markers for the species level difficult (but see e.g., Rutschmann, Detering, Simon, Fredslund, & Monaghan, 2017). Here, DNA-barcoding-like data can be used as a starting point to propose primary species hypotheses, in groups where such hypotheses do not exist or are highly questionable. Conversely, species delimitation methods based on multilocus data, such as SPEDESTEM (SPeCies Delimitation Using the Species Tree Estimation Method), BPP (Bayesian Phylogenetics and Phylogeography) or BFD (Bayes Factor Delimitation) (Ence & Carstens, 2011; Leaché et al., 2014; Yang & Rannala, 2014), are used to test alternative models of species delimitation. These methods are, generally, used for recently diverged lineages for which the speciation process may or may not be completely finalized (the “grey zone”—De Queiroz, 2007). Finally, the dichotomy between genetic sampling approaches can also be thought of in divergence times. For example, DNA-barcoding approaches are better suited for systems containing old lineages, where genotypes and phenotypes are distinct and have become fixed between lineages. In contrast, more recent lineages will suffer from contradictions between gene trees and species trees because not enough time has passed for lineages to completely sort. These systems will require finer methods based on population genetics concepts (Pante, Puillandre, et al., 2015).

RAD sequencing (restriction site-associated DNA markers) (Baird et al., 2008; Miller, Dunham, Amores, Cresko, & Johnson, 2006) is one relatively new approach that can be used to overcome the problem of collecting large multilocus genetic data sets in groups with traditionally poor genetic resources (Boucher, Casazza, Szövényi, & Conti, 2016; Herrera & Shank, 2016; Pante, Abdelkrim, et al., 2015). Here, we use a combination of DNA-barcoding-like approaches and RAD-seq data to delimit species. We initially apply DNA barcoding to define primary hypotheses of species delimitation and use RAD-seq data to verify and test the species hypotheses, in a poorly known group of marine gastropods, the *Xenuroturrus/lotyrris* complex. These species belong to the family Turridae, superfamily Conoidea, a hyperdiverse group of marine gastropods which developed a powerful venom apparatus to produce highly potent toxins used to capture their prey. The “turrids” are famous among malacologists for their highly variable shells, and this variability does not always coincide with species boundaries estimated using genetic data (e.g., Fedosov, Stahlschmidt, Puillandre, Aznar-Cormano, & Bouchet, 2017; Puillandre, Cruaud, & Kantor, 2010; Puillandre, Fedosov, Zaharias, Aznar-Cormano, & Kantor, 2017; Puillandre, Sysoev, Olivera, Couloux, & Bouchet, 2010). Several hypotheses (retention of ancestral polymorphism, convergence, phenotypic plasticity) have

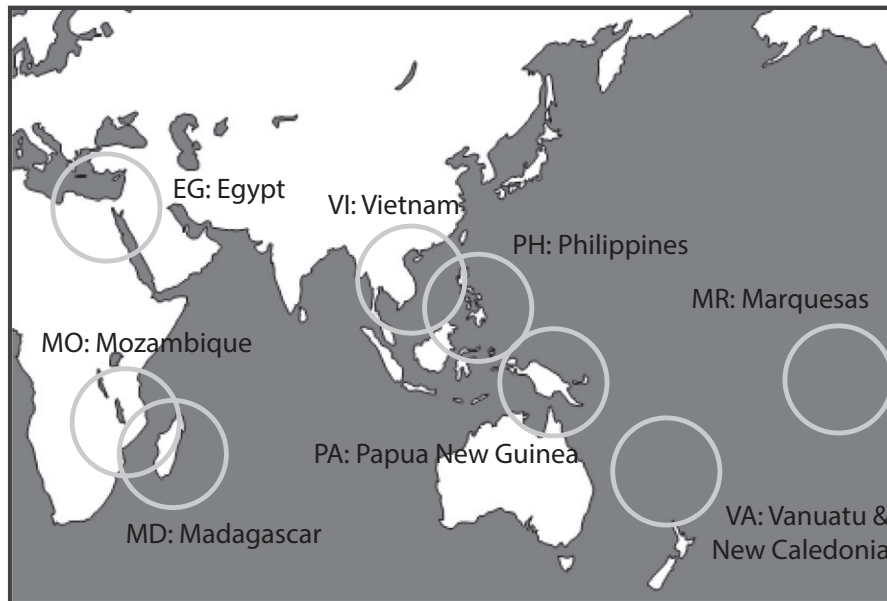
been proposed to explain how highly similar shells can actually correspond to species that diverged more than 60MYA or, conversely, how morphological variation among populations within species can exceed variation estimated at the species level (Duda, Bolin, Meyer, & Kohn, 2008; Puillandre, Baylac, Boisselier, Cruaud, & Samadi, 2009; Puillandre et al., 2011). In the *Xenuroturrus/lotyrris* complex, Kantor, Puillandre, Olivera, and Bouchet (2008) found that species with almost identical shells were easily distinguished with two genetic markers (COI—Cytochrome c Oxidase subunit I—and 28S) and by two very distinct radulae (i.e., teeth structure in gastropods). However, up until this study, only eight specimens spanning a small fraction of the known geographic area of this group (Vanuatu) were available.

The objective of the study is to revise the species delimitation in the *Xenuroturrus/lotyrris* complex using an increased number of specimens from a large geographic range, combining (a) the barcode fragment of the COI gene analysed using several species delimitation methods (ABGD, GMYC, mPTP) in order to propose primary species hypotheses (PSH); (b) a genomewide RAD sequencing approach, a method adapted to nonmodel organisms (Kess, Gross, Harper, & Boulding, 2015), analysed through tree reconstruction (using IQ-tree) and using the BFD species delimitation methods to test the alternative partitions of PSH proposed with the COI gene; and (c) morphological, anatomical and geographic data to propose secondary species hypotheses (SSH) in an integrative taxonomy context, where species, defined as definitely diverging lineages, are considered as hypotheses engaged in a process of validation or modification (Barberousse & Samadi, 2010; De Queiroz, 2007).

## 2 | MATERIALS AND METHODS

### 2.1 | Material

Specimens belonging to the *Xenuroturrus/lotyrris* species complex were collected during several field expeditions organized by the MNHN (Muséum National d'Histoire Naturelle): “Santo 2006” in Vanuatu, “Terrasses” in New Caledonia, “Inhaca 2011” and “Mainbaza” in Mozambique, “Atimo Vatae” in Madagascar, “Papua Niugini” in Papua New Guinea and “Pakaihi I Te Moana” in the Marquesas Islands; by joined Russian-Vietnamese tropical centre (Vietnam); by the University of Utah in collaboration with the University of the Philippines (Philippines); and by P. Stahlschmidt in Egypt (Figure 1). Specimens were routinely collected with use of SCUBA equipment, almost exclusively during night dives. Specimens were removed from the shell by using an isotonic solution of the magnesium chloride until relaxed (before 2012) or processed with the use of microwave (Galindo, Puillandre, Strong, & Bouchet, 2014). As part of another project, some shells were broken to access the venom glands; in these cases, the shells were photographed first. In all cases, a tissue-clip was preserved in ethanol. Voucher shells and the body of the molluscs were kept for future morphological studies; all vouchers were deposited in the MNHN collections and specimen data and COI sequences were submitted to BOLD and



**FIGURE 1** Map of the collecting sites

GenBank (Table 1). Because the monophyly of the species complex was demonstrated previously (see e.g., Puillandre, Modica, et al., 2012), only one closely related outgroup was included (*Lophiotoma jickelli*, Turridae).

A total of 95 samples were analysed; up to 12 samples were collected per locality and per morphospecies. However, most species in this complex are quite rare, and for some morphospecies and/or locality, only one sample was available. Furthermore, the RAD sequencing was not successful for several samples (see below), thus reducing the number of samples per morphospecies and locality, but still retaining at least one sample per locality and per morphospecies (Table 1).

## 2.2 | Sanger sequencing and primary species hypotheses

DNA was extracted using the Epmotion 5075 robot (Eppendorf), following the manufacturers' recommendations. A fragment of the COI gene was amplified using universal primers LCO1490/HCO2198 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994). PCRs were performed in 25  $\mu$ l, containing 3 ng of DNA, 1 $\times$  reaction buffer, 2.5 mM MgCl<sub>2</sub>, 0.26 mM dNTP, 0.3 mM of each primer, 5% DMSO, and 1.5 units of Qbiogene Q-Bio Taq. Amplification consisted of an initial denaturation step at 94°C for 4 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 50°C for COI, followed by extension at 72°C for 1 min. The final extension was at 72°C for 5 min. PCR products were purified and sequenced by the Eurofins sequencing facility (France).

Chromatograms were edited using CODONCODE ALIGNER V. 3.7.1.1 (CodonCode Corporation, Dedham, MA; [www.codoncode.com](http://www.codoncode.com)). Sequences were aligned using MUSCLE (Edgar, 2004), and the accuracy of the alignment was checked by eye. Maximum-likelihood (ML)

analyses were performed using RAXML 7.0.4 (Stamatakis, 2006), with a GAMMA model applied independently to each codon position. Accuracy of the results was assessed by bootstrapping (1,000 replicates). Bayesian analyses (BA) were performed running two parallel analyses in MRBAYES (Huelsenbeck, Ronquist, & Hall, 2001), consisting each of eight Markov chains of 200,000,000 generations each with a sampling frequency of one tree each 10,000 generations. The number of swaps was set to five, and the chain temperature at 0.02. Similar to the ML approach, unlinked models (each with six substitution categories, a gamma-distributed rate variation across sites approximated in four discrete categories and a proportion of invariable sites) were applied for each partition. For both MRBAYES and BEAST (see below) analyses, convergence was evaluated using TRACER 1.6 (Rambaut & Drummond, 2014), to confirm that the ESS values were > 200. A consensus tree was then calculated after omitting first 25% trees as burn-in. RAXML, MRBAYES and BEAST (see below) analyses were performed on the Cipres Science Gateway (<http://www.phylo.org/portal2/>) using the RAXML-HPC2 on TG, MRBAYES on XSEDE (3.2.6) and BEAST on XSEDE (1.8.2) tools, respectively.

Three methods of species delimitation were applied to propose PSH: (a) ABGD (Automatic Barcode Gap Discovery; Puillandre, Lambert, et al., 2012), which automatically detects a gap in the distribution of pairwise genetic distances, making the assumption that it corresponds to a threshold between intra- and interspecific distances; (b) GMYC (General Mixed Yule Coalescent model; Fujisawa & Barraclough, 2013; Monaghan et al., 2009; Pons et al., 2006), which tests whether branching rates along an ultrametric tree fits better with a speciation model or a coalescent model, using the transition point between speciation and coalescence to delimit species hypotheses; and (c) PTP (bayesian Poisson Tree Processes; Zhang et al., 2013), which also compares speciation and coalescent models but relies on substitution rates calculated for

**TABLE 1** Specimen list, with locality data, species names, BOLD and GenBank accession number (for the COI gene) and total number of reads, number of reads passing quality checks, total number of loci identified and number of loci present in the final data set, shared with at least one other sample (for the RAD-seq data set)

No MNHN	Country	Coordinates; depth	Genus	Species
MNHN-IM-2007-17686	Vanuatu	15°33,6'S, 167°16,6'E; 8–9 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-17246	Madagascar	25°02.6'S, 47°01.2'E; 45–49 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-24927	Mozambique	25°59.7'S, 32°54.5'E; –1 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-24928	Mozambique	25°59.0'S, 32°54.3'E; 3–5 m	<i>lotyrris</i>	<i>devoizei</i>
MNHN-IM-2009-24929	Mozambique	25°59.7'S, 32°54.5'E; –1 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-24930	Mozambique	25°59.7'S, 32°54.5'E; 2–5 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-24936	Mozambique	25°59.7'S, 32°54.5'E; 2–5 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-24939	Mozambique	26°10.9'S, 32°57.2'E; –15 m	<i>lotyrris</i>	<i>devoizei</i>
MNHN-IM-2009-24940	Mozambique	25°59.7'S, 32°54.5'E; 2–5 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-24949	Mozambique	25°58.6'S, 32°54.1'E; –15 m	<i>lotyrris</i>	<i>devoizei</i>
MNHN-IM-2009-24959	New Caledonia	23°19'S, 168°16'E; 180–220 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-24987	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-24988	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-24990	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-24992	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-24993	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-24997	Philippines	10°17'35"S, 123°55'31"E; 15–25 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-26289	Vanuatu	15°36,1'S, 166°58,5'E; 16 m	<i>lotyrris</i>	<i>devoizei</i>
MNHN-IM-2009-29709	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-29714	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>lotyrris</i>	<i>conotaxis n. sp.</i>
MNHN-IM-2009-29715	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-29719	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-29720	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>lotyrris</i>	<i>notata</i>
MNHN-IM-2009-29726	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>lotyrris</i>	<i>notata</i>
MNHN-IM-2009-33530	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33531	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33533	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33534	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33536	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33537	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33538	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33539	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>conotaxis n. sp.</i>
MNHN-IM-2009-33540	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33542	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-33544	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-33545	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-33546	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33548	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>conotaxis n. sp.</i>
MNHN-IM-2009-33549	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>
MNHN-IM-2009-33550	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-33553	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33554	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>kingae</i>
MNHN-IM-2009-33555	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturrus</i>	<i>legitima</i>

BOLD ID	Genbank # for COI	Total reads	Reads passed	Total loci	Final loci
CONO514-08	EU127880	694,199	630,308	3,617	2,311
TEMPO044-18	MH917863	3,600,990	3,374,050	17,905	12,615
TEMPO050-18	MH917812	1,293,596	1,223,835	7,515	4,754
TEMPO051-18	MH917824	3,259,241	3,039,143	14,371	9,983
TEMPO052-18	MH917809	1,981,423	1,866,477	11,077	7,354
TEMPO053-18	MH917868	3,740,707	3,504,351	17,562	12,488
TEMPO054-18	MH917867	2,707,631	2,537,846	15,716	8,064
TEMPO103-18	MH917821	1,487,639	1,425,098	7,175	4,774
TEMPO055-18	MH917866	2,115,504	1,999,266	9,503	6,312
TEMPO056-18	MH917823	2,673,240	2,487,408	11,010	7,162
TEMPO057-18	MH917843	1,147,331	1,073,660	6,083	3,242
TEMPO058-18	MH917848	621,317	590,364	3,927	2,687
TEMPO059-18	MH917842	2,547,012	2,401,207	13,361	9,616
TEMPO060-18	MH917847	4,790,989	4,454,342	22,688	16,716
TEMPO061-18	MH917846	2,904,170	2,709,792	15,444	11,259
TEMPO062-18	MH917864	3,125,200	2,957,616	15,195	10,822
TEMPO063-18	MH917838	2,007,246	1,895,232	11,599	8,662
CONO512-08	EU127879	1,253,969	1,143,822	6,011	3,804
TEMPO064-18	MH917815	3,884,108	3,659,529	19,469	13,259
TEMPO065-18	MH917853	616,563	583,635	3,011	1,718
TEMPO066-18	MH917837	2,800,010	2,680,824	15,992	12,044
TEMPO067-18	MH917862	1,188,373	1,098,161	4,985	3,181
TEMPO068-18	MH917840	9,135,778	8,614,984	32,432	21,909
TEMPO069-18	MH917839	2,968,648	2,785,558	13,687	9,024
TEMPO070-18	MH917861	1,534,833	1,450,158	8,924	6,261
TEMPO071-18	MH917860	286,377	2,484,120	1,1885	8,193
TEMPO072-18	MH917859	8,601,257	8,172,565	35,030	2,4127
TEMPO073-18	MH917858	2,988,491	2,765,920	1,3189	8,959
TEMPO074-18	MH917826	3,377,363	3,242,368	16,184	12,077
TEMPO075-18	MH917836	6,041,344	5,631,785	29,218	19,312
TEMPO076-18	MH917856	5,710,015	5,371,853	25,099	17,868
TEMPO077-18	MH917849	3,935,930	3,702,928	21,354	15,466
TEMPO078-18	MH917835	2,165,240	1,989,705	10,988	7,999
TEMPO079-18	MH917817	2,718,353	2,551,693	12,902	8,822
TEMPO080-18	MH917820	2,713,372	2,543,124	15,457	1,1012
TEMPO081-18	MH917819	2,260,769	2,101,961	12,753	9,094
TEMPO082-18	MH917834	2,615,219	2,453,002	13,097	9,545
TEMPO083-18	MH917852	2,499,420	2,357,567	15,697	11,702
TEMPO084-18	MH917857	739,942	698,742	3,750	2,132
TEMPO085-18	MH917841	1,967,315	1,823,804	8,611	5,820
TEMPO086-18	MH917833	4,597,154	4,372,658	22,254	16,663
TEMPO087-18	MH917825	5,770,810	5,492,875	26,157	1,8567
TEMPO088-18	MH917865	1,800,651	1,703,896	9,132	6,468

(Continues)



TABLE 1 (Continued)

No MNHN	Country	Coordinates; depth	Genus	Species
MNHN-IM-2009-33561	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-33562	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>olangoensis</i>
MNHN-IM-2009-33563	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-33564	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>conotaxis</i> n. sp.
MNHN-IM-2009-33568	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33569	Philippines	10°17' S, 123°55' E; 15–20 m	<i>Xenuroturris</i>	<i>legitima</i>
MNHN-IM-2009-33570	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>conotaxis</i> n. sp.
MNHN-IM-2009-33572	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33576	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33577	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33578	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-33579	Philippines	10°17' S, 123°55' E; 15–20 m	<i>lotyrris</i>	<i>musivum</i>
MNHN-IM-2009-6456	Mozambique	26°12' S, 35°03' E; 87–90 m	<i>lotyrris</i>	<i>conotaxis</i> n. sp.
MNHN-IM-2009-7022	Mozambique	25°59.0'S, 32°54.5'E	<i>Xenuroturris</i>	<i>legitima</i>
MNHN-IM-2009-7023	Mozambique	25°59.0'S, 32°54.5'E	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-7024	Mozambique	25°59.0'S, 32°54.5'E	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-7025	Mozambique	25°59.0'S, 32°54.5'E	<i>Xenuroturris</i>	<i>legitima</i>
MNHN-IM-2009-7081	Mozambique	25°59.0'S, 32°54.5'E	<i>Xenuroturris</i>	<i>legitima</i>
MNHN-IM-2013-14888	Papua New Guinea	05°11' S, 145°49.5'E; 2–10 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2013-40060	Marquesas Islands	09°45,67'S, 138°50,69'W; 10–25 m	<i>lotyrris</i>	<i>devozei</i>
MNHN-IM-2013-52076	Egypt	26°48'48"N, 33°56'54"E; 1–2 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2013-52078	Egypt	26°48'48"N, 33°56'54"E; 1–2 m	<i>lotyrris</i>	<i>cingulifera</i>
MNHN-IM-2009-29713	Vietnam	12°10,443'N, 109°16,298'E; 15–22 m	<i>Lophiotoma</i>	<i>jickelii</i>

each nodes instead of branching rates. The web server available at <http://www.wabi.snv.jussieu.fr/public/abgd/> (version of March 2017) was used for ABGD, with the default parameters. The distance matrix was computed by ABGD, using the Jukes–Cantor substitution model. BEAST 1.8.1 (Drummond & Rambaut, 2007) was used to obtain a relative-rates ultrametric tree for the GMYC and PTP analyses, with a relaxed lognormal clock and a coalescent prior, determined as the best-fitting parameters to be used with the GMYC model (Monaghan et al., 2009). A GTR+I+G substitution model was applied, and the Metropolis coupled Markov chains (MCMC) were run for 100,000,000 generations. GMYC (both single and multiple versions), PTP and mPTP (Kapli et al., 2016; Monaghan et al., 2009; Pons et al., 2006; Zhang et al., 2013) were run using the web servers at [species.h-its.org](http://species.h-its.org) and [mptp.h-its.org](http://mptp.h-its.org) (versions of March 2017), respectively, using default parameters.

### 2.3 | RAD sequencing

Single digest RAD sequencing (Baird et al., 2008) was conducted using the restriction enzyme SbfI on 95 samples. Barcoded Illumina library preparation and sequencing was subcontracted to Eurofins. Classical libraries were constructed and single-end sequenced on two lanes of Illumina HiSeq 2000. The sequencing resulted in a total of 293 million reads. Reads were demultiplexed according to the 10 base long barcodes

with the allowance of one mismatch using fastx barcode splitter from the FASTX-Toolkit suite ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). Seven million reads remained unassigned. The number of reads per sample varied from 6,700 to 23 million with a mean of 2.9 million.

Reads quality was checked using FastQC (<http://www.bioinformatics.babraham.ac.uk>) for each sample individually. Since part of the samples indicated quality drops after 70–80 bases, the ends of reads were cleaned-up following a sliding window approach using Fastq quality trimmer also from the FASTX-Toolkit suite. Reads were cleaned from the 3' end using a window size of five, a step size of one and an average minimum score within the window of 20. Following clean-up, reads were checked again using FastQC to make sure quality profiles were satisfying.

Usable loci were produced from raw reads using pyRAD 3.01 (Eaton, 2014). The choice of this pipeline was made according to the fact that both (a) the presence of indels (which was anticipated between more distant samples) and (b) the trimming of reads result in variable reads length which is not allowed in the more commonly used pipeline STACKS (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). Several combinations of parameters were used and quickly revealed that samples with <1 million reads resulted in low numbers of usable loci leading to long branches artefacts. Such samples were discarded to obtain the final data set with the 66 remaining samples, including one outgroup (Table 1). Once discarding these

BOLD ID	Genbank # for COI	Total reads	Reads passed	Total loci	Final loci
TEMPO089-18	MH917844	3,110,611	2,931,793	17,218	12,766
TEMPO090-18	MH917845	4,898,433	4,678,365	22,680	1,6345
TEMPO091-18	MH917814	4,990,679	4,612,622	2,1240	14,610
TEMPO104-18	MH917854	6,835,600	6,491,429	29,470	22,148
TEMPO092-18	MH917832	3383,055	3,228,317	18,860	13,367
TEMPO093-18	MH917869	4,451,716	4,094,310	20,965	14,700
TEMPO094-18	MH917851	3,232,203	3,038,981	15,976	1,1442
TEMPO095-18	MH917831	4,993,521	4,816,966	23,634	18,023
TEMPO096-18	MH917830	3,326,764	3,195,354	18,611	13,750
TEMPO097-18	MH917829	3,392,736	3,076,448	17,071	12,471
TEMPO098-18	MH917828	2,72,092	26,22,579	15,766	11,940
TEMPO099-18	MH917827	1,978,175	1,865,570	1,1715	8,576
TEMPO043-18	MH917850	612,170	548,856	1,571	779
TEMPO045-18	MH917855	6,486,587	6,148,897	33,000	19,236
TEMPO046-18	MH917813	3,943,003	3,682,059	21,778	12,034
TEMPO047-18	MH917811	2,651,450	2,500,060	16,263	6,246
TEMPO048-18	MH917871	8,045,196	7,655,973	40,943	14,177
TEMPO049-18	MH917870	11,296,176	10,605,392	52,463	19,298
TEMPO100-18	MH917810	1,841,688	1,720,104	11,264	7,242
TEMPO101-18	MH917822	9,016,668	8,525,416	33,530	23,998
TEMPO102-18	MH917818	4,929,978	4,693,937	30,097	16,529
TEMPO105-18	MH917816	1,894,537	1,730,192	10,657	5,192
CONO1878-17	KY570852	23,237,420	22,117,193	46,894	3,348

samples, the outputs of pyRAD were relatively stable. Among the 66 samples treated here, 46 were micro-waved to remove the mollusc from its shell. The average number of reads obtained per sample was slightly higher for the micro-waved samples (3,971,187) than for the non-micro-waved ones (3,290,728), confirming that the micro-waves do not damage the DNA. The higher number of reads for the micro-waved samples could be explained by the fact that the micro-waved samples were also more recently collected.

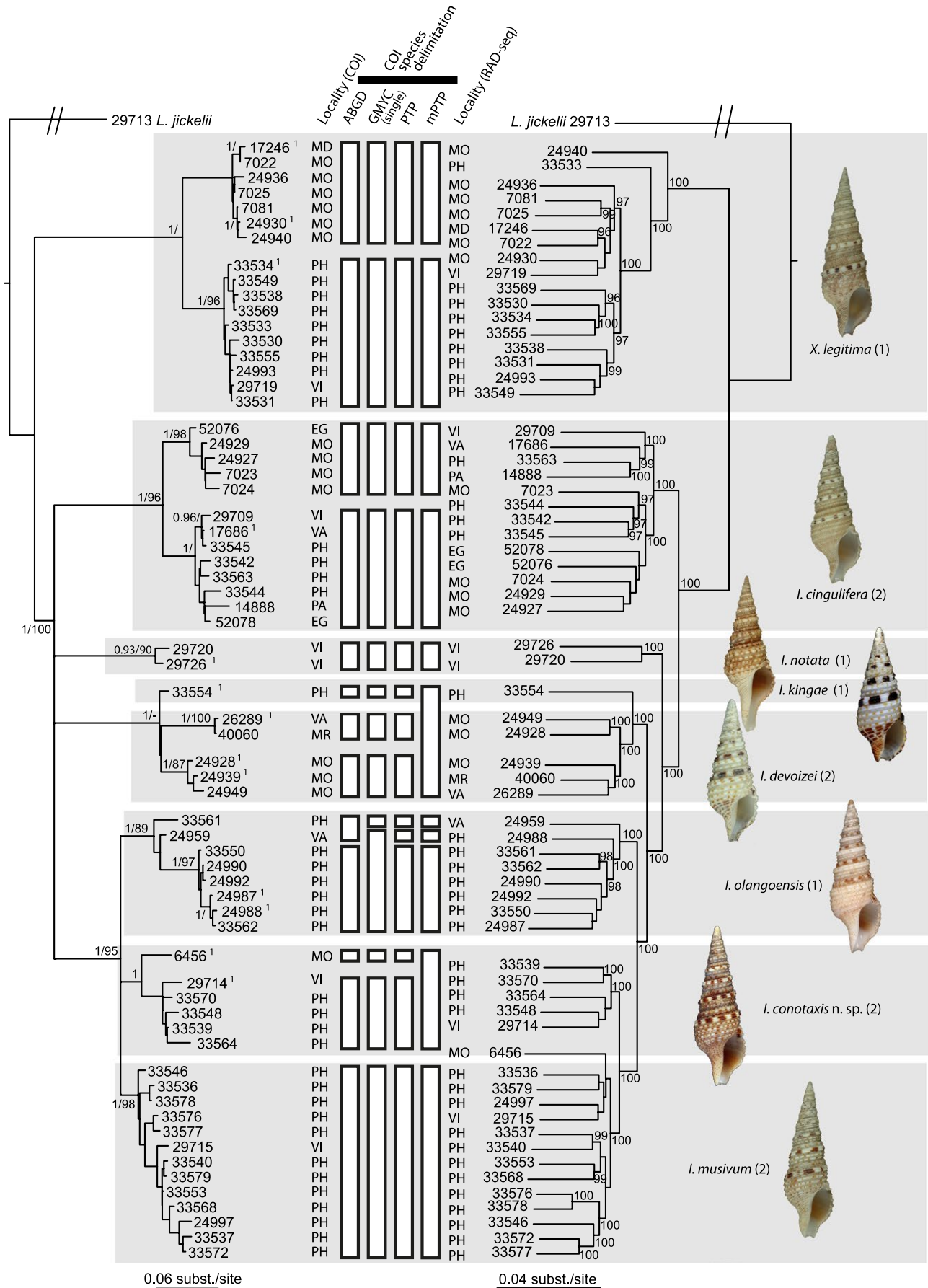
A minimum coverage (Mindepth) of five reads was used, as well as a clustering threshold (Wclust) of 0.89, and a minimum of two samples shared by any locus (MinCov). A maximum-likelihood tree was produced using IQ-tree (Nguyen, Schmidt, von Haeseler, & Minh, 2014). We estimated the best substitution model for each locus with ModelFinder (Kalyaanamoorthy, Minh, Wong, von Haeseler, & Jermini, 2017) following the BIC criterion. We then applied 1000 ultrafast bootstrap (UFBoot) (Hoang, Chernomor, von Haeseler, Minh, & Le, 2017) on each data set to obtain branch support. The same tree reconstruction was also conducted only keeping one SNP per locus and resulted in a similar topology. In order to test the effect of the parameter MinCov, which has a direct impact on the level of missing data in the final outputs, we also generated data sets corresponding to data missing for up to 30, 40, 50, 60, 70 and 80% of the samples (i.e., the maximum percentage of samples not having information for a specific locus to be kept, corresponding to MinCov

values of 46, 40, 33, 26, 20 and 13, respectively). For each data set, the total number of loci, the total number of SNP and the percentage of nodes in the IQ-tree showing a bootstrap value above 75 and 95 were calculated.

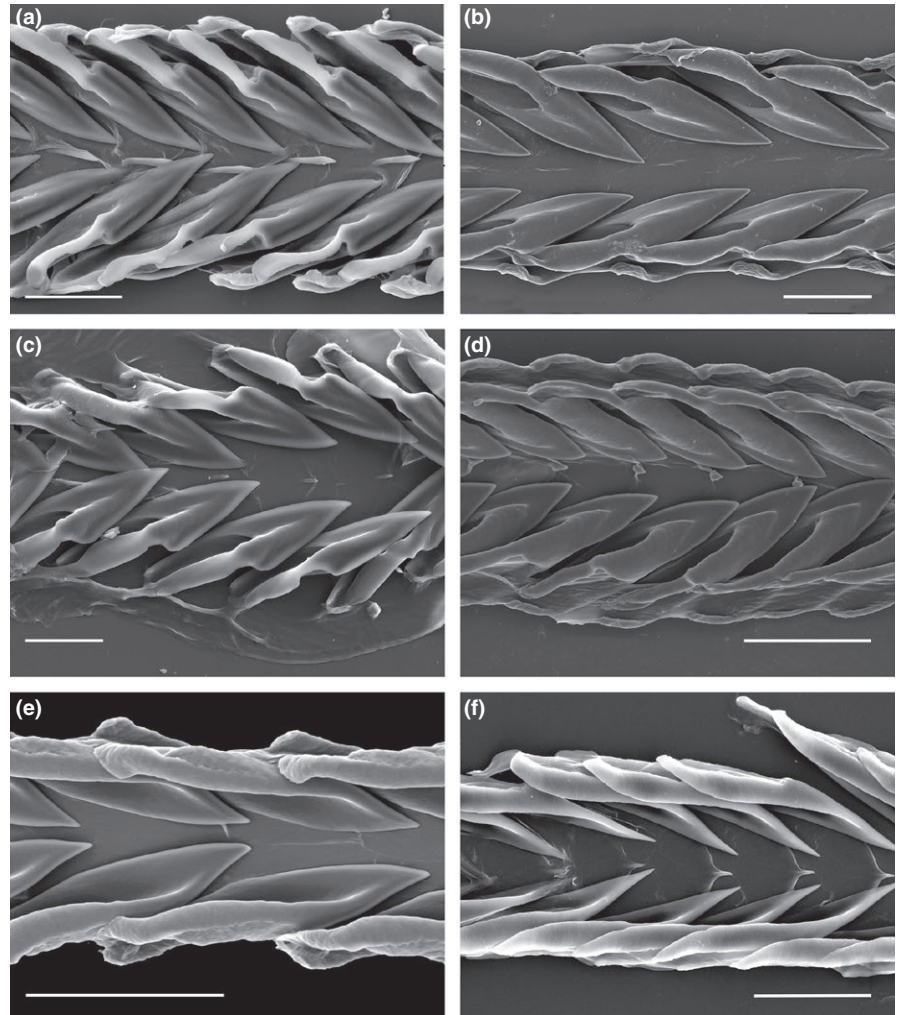
The RAD-seq data set was also analysed with BFD to test alternative partition of species (see below) proposed by ABGD, GMYC or PTP with the COI data set. MLE (Marginal Likelihood Estimates) for each partition of species were obtained using the implementation of BFD\* in the SNAPP (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) plug-in for BEAST v2.5 (Bouckaert et al., 2014). Given the high number of loci and the high level of missing data, we kept only the 10% of loci with the lowest level of missing data (corresponding to 470 loci) and performed multiple runs with various number of steps (20, 50, 100 and 200) and chain length (100,000 and 500,000) for the path sampling, with a preburnin of 50,000. Bayes factors (BF) were calculated from the MLE for each model pair. We followed Grummer, Bryson, and Reeder (2013) in recognizing a 2lnBf >10 as “decisive” support in distinguishing between competing species delimitation hypotheses.

## 2.4 | Secondary species hypotheses

The PSH proposed with ABGD, GMYC and PTP were compared with the results obtained with the RAD-seq data. Because specimens



**FIGURE 2** Bayesian tree (Mr Bayes) obtained with the COI data set (left) and ML tree obtained with the RAD-seq data set with 80% of missing data (right). Posterior Probabilities and bootstrap values are shown for each node in the COI tree for supported nodes only (Posterior Probabilities >0.95 and/or bootstrap values > 80); bootstrap values (above 95) are shown for each node in the RAD-seq tree. For clarity, we removed the prefix of each sample number (IM-20XX); full samples numbers are provided in Table 1. Species hypotheses proposed by ABGD, GMYC, PTP and mPTP are shown as vertical bars, as well as the locality for each specimen (MO = Mozambique, MD = Madagascar, EG = Egypt, VI = Vietnam, PH = Philippines, PA = Papua New Guinea, VA = Vanuatu and New Caledonia, MR = Marquesas). Grey boxes represent the final species hypotheses, with the species name, followed by the radula type (type 1 or type 2, shown on Figure 3) and one illustrated shell. <sup>1</sup>: specimens for which the radula has been observed [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 3** Radula type 1 (a–d): duplex marginal teeth. (a) *Xenuroturrus legitima* Iredale, 1929, MNHN-IM-2009-24931, Mozambique. (b) *Lotyrris notata* (Sowerby, 1889), MNHN-IM-2009-29726, Vietnam. (c) *Lotyrris olangoensis* (Olivera, 2002), MNHN-IM-2009-33561, Philippines. (d) *Lotyrris kingae* (Powell, 1964), MNHN-IM-2009-33554, Philippines. Radula type 2 (e, f): semi-enrolled marginal teeth. (e) *Lotyrris cingulifera* (Lamarck, 1822), MNHN-IM-2007-17686, Santo. (f) *Lotyrris devoizei* Kantor et al., 2008, MNHN-IM-2009-24939, Mozambique. Scale bars – 50  $\mu$ m

from the same species are supposed to recombine on independent loci, contrary to specimens from different species, intraspecific relationships inferred from these two data sets are expected to be different; conversely, interspecific relationships are expected to be more congruent in both COI and RAD-seq trees (Pante, Puillandre, et al., 2015). Based on this property, we looked for PSH defined with the COI gene that corresponded to clades in the phylogenetic tree obtained with the RAD-seq data (i.e., all the specimens of a COI PSH cluster together in a single clade in the RAD-seq tree), a pattern in support of the hypothesis that these PSH are actually different SSHs. We applied the integrative taxonomy flowchart described in (Puillandre, Modica, et al., 2012) to add arguments in favour of one or two SSHs, when two PSHs are compared. In particular, we

analysed species distribution patterns: an overlapping distribution between two sister-PSHs would support the hypothesis that they correspond to two different SSHs; an allopatric distribution is less informative, since although it demonstrates limited gene flow between populations, it does not allow judgments on whether the distant populations are two reproductively isolated species. All the analysed specimens exhibit a multispiral protoconch, indicative of planktotrophic development (Jablonski & Lutz, 1980) and suggestive of high dispersal capabilities.

A previous study (Kantor et al., 2008) demonstrated that in *Xenuroturrus/lotyrris* complex, the shell shape cannot be considered as a reliable character for delimitation of the species, with intraspecific variation exceeding the interspecific one. On the contrary, the



spiral sculpture as well as coloration of teleoconch appeared to be important diagnostic characters. Therefore, we paid special attention to the sculpture of the subsutural zone (subsutural ramp) and the zone of anal sinus (sinus cords) as well as colour pattern of studied species. The teleoconch characters were compared to the type specimens of known species in the group to link the species hypotheses to available species names. Because the radula has been shown to be variable within this complex, several specimens per PSHs (when possible, selected from the sequenced material) were dissected to identify the radula type. Radulae were prepared by standard methods (Kantor & Puillandre, 2012) and examined by scanning electron microscope TeScan TS5130MM in the Institute of Ecology and Evolution of Russian Academy of Sciences (IEE RAS). Terminology used for radula description follows Kantor (2006).

### 3 | RESULTS

For clarity, the species names used from here onwards are attributed to either *Xenuroturris* or *Iotyrris* following the results of the phylogenetic analysis: *X. legitima* Iredale, 1929, *I. cingulifera* (Lamarck, 1822), *I. notata* (Sowerby, 1889), comb. nov., *I. kingae* (Powell, 1964), comb. nov. *I. devoizei* Kantor et al., 2008; *I. olangoensis* (Olivera, 2002), comb. nov., *I. conotaxis* n. sp., *I. musivum* Kantor et al., 2008.

#### 3.1 | Results of the exploratory approach

When ABGD is used with default parameters, two groups corresponding to the main lineages (identified as two different genera in Kantor et al., 2008) were detected. Therefore, we analysed each genus separately with ABGD. In the vicinity of the barcode gap (i.e., ~2%–6%), ABGD consistently defines two PSH within the first lineage (*Xenuroturris*) and 11 PSHs in the second (*Iotyrris*), with both the initial and recursive approaches (Figure 2). As shown in Figure 2, species names available in the literature were attributed to the PSH through the comparison of shell morphology to the known species, including the type specimens; one PSH, consistently defined by all the methods, could not be attributed to an available name and is thus described as a new species (*I. conotaxis* n. sp.—see below). Within *X. legitima*, *I. cingulifera*, *I. devoizei* and *I. conotaxis* n. sp., two allopatric lineages are recognized, one in the Indo-West Pacific Ocean (IWP—Vietnam, Philippines, Vanuatu, Chesterfield and Marquesas Islands),

and the second in the West-Indian Ocean (WIO—Mozambique and Madagascar). Similar partitions are also found with the single version of GMYC, the PTP and the mPTP methods (the multiple version of GMYC returns an unrealistic number of PSHs—21, a result reported previously—for example (Fujisawa & Barraclough, 2013; Kekkonen & Hebert, 2014; Talavera, Dincă, & Vila, 2013). The exceptions are the following: with mPTP, the two lineages of *I. devoizei* and *I. kingae* are grouped in a single PSH; within *I. olangoensis*, two specimens, the unique one from Vanuatu, and one of the Philippines specimens, are considered as one or two PSH, in addition to a PSH, consistently found by all the methods, grouping the other specimens of *I. olangoensis* from the Philippines. Finally, the mPTP method groups in a single PSH the two allopatric lineages within *I. conotaxis* n. sp. Remarkably, the two specimens from Egypt do not cluster in the same PSH: one clusters in the IWP PSH of *I. cingulifera*, the other in the WIO PSH of *I. cingulifera* with all methods.

#### 3.2 | RAD-seq and secondary species hypotheses

Several sets of parameters were tested using the pipeline pyRAD for the generation of the alignment used for tree reconstructions. Unlike most other parameters tested, the parameter MinCov, corresponding to the minimum number of samples with information needed to keep a locus in the final data set, had a strong effect on the number of loci and SNP recovered from our RAD-seq data. The number of loci recovered varied from 2 to 17,060 and the number of total SNP from 22 to more than 97,393 when MinCov was changed in order to maintain 30 to 80% of the samples having information as the cut-off to keep a locus (Table 2). Nevertheless, even though adding more loci directly increases the percentage of missing data in a dramatic way, the topologies obtained were congruent with each other, when the supported nodes (Bootstraps values above 95) were considered (Figure 2, Appendix S1). Moreover, those obtained with more loci and thus more missing data were better resolved: nodes with bootstrap values more than 75 and 95 increase from 8 to 90% and from 3 to 65%, when MinCov is reduced from 46 to 13, respectively. This increase in number of supported nodes followed a linear function, with a higher correlation ( $R^2 = 0.95$ ) for nodes with bootstraps values above 95 than for nodes with bootstraps values above 75 ( $R^2 = 0.79$ ). This likely occurred because the number of nodes with bootstraps values above 75 reached a maximum with the data set containing 70% of missing data (92%), then decreasing to 90% with the data set

% of missing data	MinCov	# loci	# SNP	% nodes with bootstraps >75	% nodes with bootstraps >95
30	46	2	22	7.93	3.17
40	40	39	251	28.57	12.7
50	33	474	3,415	80.95	44.44
60	26	2,639	17,248	87.30	50.79
70	20	7,013	43,828	92.06	61.9
80	13	17,060	97,393	90.48	65.08

**TABLE 2** Variation of the level of missing data (expressed as the maximum percentage of samples not having information for a specific locus to be kept) and its effect on the number of loci and SNP and percentage of nodes with bootstrap values >50

containing 80% of missing data. With a MinCov value of 2, the data set included 103,060 loci after paralog removal.

The clades recovered with the RAD-seq data are congruent with the PSHs, or groups of PSHs, as defined with the COI (Figure 2). The PSHs *X. legitima*, *I. cingulifera*, *I. notata*, *I. kingae*, *I. devoizei* and *I. olangoensis* are found as independent lineages in the RAD-seq tree. In the more supported IQ-trees (obtained with 50 to 80% of missing data), these PSH are always recovered with high support, but the relationships within each of these PSH change from one tree to another. One exception is the clade that unites *I. musivum* and *I. conotaxis* n. sp.: these two PSHs are not always recovered because the position of the unique sample from the WIO, IM-2009-6456, constant moves from one PSH to the other. This pattern is probably artefactual, since this sample is the one with the lowest number of reads and loci. When this sample is removed, the pattern for *I. musivum* and *I. conotaxis* n. sp. is similar to the other species, with both PSH being always recovered as a fully supported clades in the IQ-trees, but with internal relationships changing from one tree to another.

The allopatric lineages within *X. legitima*, *I. cingulifera* and *I. devoizei* are not monophyletic in the RAD-seq tree. Furthermore, the monophyly of the WIO clade of *X. legitima* was not supported with the COI gene and the monophyly of the IWP clade of *I. cingulifera* was supported only in the Bayesian tree of the COI gene. The two specimens (MNHN-IM-2009-33561 and MNHN-IM-2009-24959) considered as separate PSH by some methods with the COI gene within *I. olangoensis* are embedded within the other *I. olangoensis* PSH in the RAD-seq tree.

For the two PSH that include allopatric lineages and sufficient number of samples, *X. legitima* and *I. cingulifera*, the BFD method was used to test the two alternative scenarios: one species with an Indo-Pacific distribution, or two allopatric (WIO vs. PO) species each. To reduce computation time, each data set was analysed separately, together with one sample (IM-2013-40060, the most complete sample in terms of reads and loci recovered) used as an outgroup. Whatever the number of steps and the chain length used for the path sampling, the MLE values recovered for the partition with only one species is always lower than for the partition with two allopatric species, with 2lnBf values all >100 (Table 3).

Given all these data, the PSHs were turned into eight SSHs, represented as grey boxes in Figure 2, as follows:

1. the two allopatric lineages within *X. legitima* are not found in the RAD-seq data set and not supported as the best model in the BFD analysis, and they have the same radula type (see below) and similar shells: they are considered a single SSH, named *X. legitima*;
2. the two allopatric lineages within *I. cingulifera* are not found in the RAD-seq data set and not supported as the best model in the BFD analysis, and they have the same radula type and similar shells: they are considered a single SSH, named *I. cingulifera*;
3. *I. notata* is consistently recognized as an independent group with all data sets and methods, and the radula type and the shells are similar among studied specimens; it is considered as a SSH;

**TABLE 3** Results of the BFD analyses. For each species, two scenarios (one single species vs. two species in allopatry) were tested, for various numbers of steps and chain lengths. MLE values are reported for each analysis

Species	Number of species considered	Number of steps	Chain length	MLE
<i>lotyrris cingulifera</i>	1	20	100,000	1,188
<i>lotyrris cingulifera</i>	1	50	100,000	889
<i>lotyrris cingulifera</i>	1	100	100,000	559
<i>lotyrris cingulifera</i>	1	200	100,000	771
<i>lotyrris cingulifera</i>	1	200	500,000	1,038
<i>lotyrris cingulifera</i>	2	20	100,000	1,336
<i>lotyrris cingulifera</i>	2	50	100,000	Infinity
<i>lotyrris cingulifera</i>	2	100	100,000	1,334
<i>lotyrris cingulifera</i>	2	200	100,000	Infinity
<i>Xenuroturrus legitima</i>	1	20	100,000	1,142
<i>Xenuroturrus legitima</i>	1	50	100,000	1,109
<i>Xenuroturrus legitima</i>	1	100	100,000	867
<i>Xenuroturrus legitima</i>	1	200	100,000	535
<i>Xenuroturrus legitima</i>	1	200	500,000	238
<i>Xenuroturrus legitima</i>	2	20	100,000	1,498
<i>Xenuroturrus legitima</i>	2	50	100,000	1,497
<i>Xenuroturrus legitima</i>	2	100	100,000	1,402
<i>Xenuroturrus legitima</i>	2	200	100,000	1,330



**TABLE 4** Diagnostic character states for each studied species of the *Xenuroturris/lotyrris* complex

Characters/species	<i>X. legitima</i>	<i>I. cingulifera</i>	<i>I. notata</i>	<i>I. kingae</i>	<i>I. devoizei</i>	<i>I. olangoensis</i>	<i>I. musivum</i>	<i>I. conotaxi</i>
Subsutural ramp sculpture	3–5 subequal cords	3–5 subequal cords	3 cords, central most prominent	3 bulging cords, central very prominent	3 bulging cords, central very prominent, rarely 2 cords	3 cords, central most prominent	3 cords, central most prominent	3 cords, central most prominent
Subsutural ramp coloration	Background colour with speckles on cords	Background colour with speckles on cords	Background colour with speckles on cords	Large subrectangular brown spots	Large subrectangular brown spots	Large subrectangular brown spots	Background colour with speckles on cords	Speckles on cords, strong to weak brown spots
Sinus (=peripheral) cords morphology	Paired subequal, upper and lower ones subdivided in two in large specimens	Paired subequal, only upper one subdivided in two in large specimens	Paired subequal, sharp on top, not subdivided	Paired subequal rather weak	Paired subequal rather weak	Paired subequal rather weak, upper can be subdivided	Paired subequal, prominent, upper can be subdivided	Paired subequal, subequal, prominent, upper can be subdivided
Sinus (=peripheral) cords coloration	Large subrectangular brown spots and speckles on cords	Large subrectangular brown spots and very few speckles on cords (usually absent)	Large speckles on cords	Large aligned speckles on cords	Large aligned speckles on cords	Speckles on cords	Large subrectangular brown spots	Large subrectangular brown spots
Radular central formation	Spear-headed shaped central cusp	Absent	Absent	Short blunt central cusp and indistinct lateral flaps	Broad with distinct lateral flaps and small central cusp	Narrow central cusp	Short central cusp, sometimes indistinct lateral flaps	Broad with distinct lateral flaps and small central cusp
Radular marginal teeth	Duplex	Semi-enrolled	Duplex	Duplex	Semi-enrolled	Duplex	Semi-enrolled	Semi-enrolled

4. *I. kingae*, although grouped with *I. devoizei* with mPTP, is morphologically (both shell and radula) and genetically different from *I. devoizei*; therefore, it is considered as an independent SSH;
5. the two allopatric lineages within *I. devoizei* are not found in the RAD-seq data set, and they have the same radula type and similar shells: they are considered a single SSH, named *I. devoizei*;
6. the two specimens (MNHN-IM-2009-33561 and MNHN-IM-2009-24959) that fall outside the main clade of *I. olangoensis* in the COI tree are not isolated with the RAD-seq data. The shell and radula are consistent among all the *I. olangoensis* specimens, and it is thus considered a single SSH;
7. *I. conotaxis* n.sp. is consistently found separate from the other lineages, with both COI and RAD-seq data. Although radula type in specimens from this lineage is the same as in *I. musivum*, both species are genetically and morphologically (see species description below) distinct, and *I. conotaxis* n. sp. is considered a single SSH. The placement of the single specimen from the WIO remains uncertain, either as sister to *I. conotaxis* n. sp. or to *I. musivum*. The shell morphology of this single specimen is not conclusive as it is a subadult from deeper waters (80–90 m) compared to the IWP samples, and its morphology slightly differs from the adult shells of both *I. conotaxis* n. sp. and *I. musivum*. More specimens of this clade from the WIO are necessary to conclude.
8. *I. musivum* is consistently found as a single PSH and is considered a single SSH.

Shell and radular morphology is congruent with the eight defined SSH. The examination of numerous specimens allowed concluding that the most informative characters of the shell are the spiral sculpture and colour pattern of the subsutural ramp and sinus cords. The subsutural ramp is the zone below the suture, which in *Xenuroturris* and *lotyrris* is delimited by narrow but distinct abapical groove. It bears three to five cords, varying interspecifically in width and coloration: cords can have numerous dark speckles limited to cords or large subrectangular spots, that extend to the interspaces between cords (*I. devoizei*, *I. kingae*, *I. olangoensis*). Sinus cords are distinct elements of the spiral sculpture, originating at the posterior part of the anal sinus (Figure 4b). Originally paired, they can be subdivided by narrow longitudinal grooves. The pattern of this subdivision as well as the coloration of the cords is also important diagnostic characters. Although there is no single character that allows recognition of all studied species, each defined SSH has its own combination of characters states. The summary of diagnostically important shell characters is provided in Table 4.

The radula in studied species of *Xenuroturris/lotyrris* complex is rather variable; two major radular types can be recognized, differing in the morphology of the marginal teeth (Figure 3). In the first type, the marginal teeth have duplex shape, typical for Turridae and many other Conoidea. The anterior part of the tooth (closer to the tip), up to half of total tooth length, is solid, while the posterior part has two thickened edges, a thinner dorsal one (the accessory limb) and a thicker ventral one (the major limb) attached to the radular membrane (Figure 3a–d). The second radular type is characterized by “semi-enrolled” marginal teeth. Both edges of the marginal teeth are elevated and equally developed along the entire tooth length,

and they delimit an intervening trough. The anterior solid part of the tooth is absent (Figures 3e,f; 5). This second type is unique in the family Turridae and is confined only to the genus *lotyrris*. Nevertheless, we found that the radular types were not correlated with the phylogenetic relationships: *X. legitima*, *I. notata*, *I. kingae* and *I. olangoensis* have the duplex marginal radular teeth, while *I. cingulifera*, *I. devoizei*, *I. conotaxis* n. sp. and *I. musivum* have the semi-enrolled marginal teeth.

### 3.3 | Species description

Superfamily CONOIDEA Fleming, 1822

Family TURRIDAE H. & A. Adams, 1853 (1838)

Genus *lotyrris* Medinskaya & Sysoev, 2001

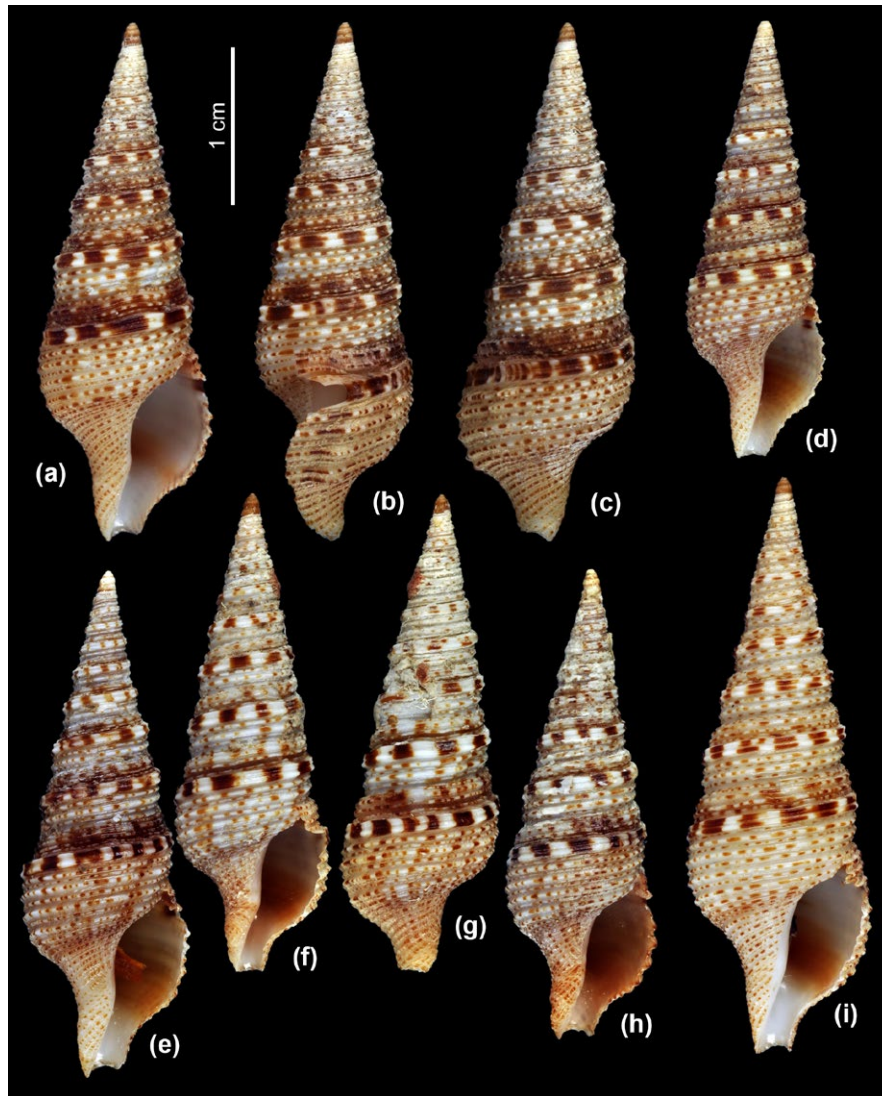
*lotyrris conotaxis* n. sp. (Figures 4 and 5)

ZooBank registration: urn:lsid:zoobank.org:act:4FB76AA0-1519-4361-A3D9-4E47F002ABAA

Holotype: MNHN-IM-2009-33570. Paratypes, MNHN-IM-2009-33539, MNHN-IM-2009-33564, both from type locality.

Type locality: Philippines, Olango Island, off Cow-Oy, 15–20 m, collected during night dive.

DESCRIPTION (HOLOTYPE). Shell conical, consisting of 9.75 evenly convex teleoconch whorls, with high spire; shell diameter to shell height 0.34, aperture height (without canal) to shell height 0.28, spire height to shell height 0.49. Protoconch brown, partially eroded, remaining part of 4.25 convex whorls, sculptured with closely spaced axial arcuate threads, slightly prosocline, nearly orthocline on most whorls, but turning to strongly opisthocline on posteriormost part of last protoconch whorl. 26 threads on last protoconch whorl. 4–5 early teleoconch whorls nearly flat, slightly angulated at shoulder. Late teleoconch whorls convex in outline, particularly the last whorl. Last whorl sharply narrowing towards attenuated but short nearly straight siphonal canal. Suture deeply impressed, slightly wavy, nearly canaliculate due to raised subsutural cord. Subsutural ramp narrow, on last whorl with three distinct broadly spaced thin nodulose cords, with interspaces twice broader than cords. Abapical edge of subsutural ramp with very narrow but deep groove, sometimes obscured by nodulose edges of minor cords. This groove clearly seen on apertural lip, bordered by narrow cord. Two strongly raised subequal in width sinus cords, triangular in section abut subsutural ramp. Whorls portion below sinus cords sculptured with subequal narrow cords, from one on uppermost teleoconch whorls to four on penultimate whorl. On last whorl, seven subequal spiral cords below sinus before transition to canal, and 15 weaker cords on shell base and canal with narrow thread between some of them. Axial sculpture absent, except for inconspicuous growth lines. Aperture narrowly oval. Outer lip thin, evenly rounded. Anal sinus deep, U-shaped, situated on the shoulder. Inner lip slightly convex, columellar part straight, callus very narrow, not extending onto the parietal wall. Canal delimited from aperture by inconspicuous fold. Colour creamy white, with regularly spaced, light-brown spots covering cords of the entire shell surface and much more pronounced, strong, darker



**FIGURE 4** Shells of *lotyrris conotaxis* n. sp. (a–c) Holotype, MNHN-IM-2009-33570, SL 33.2 mm. (d) MNHN-IM-2009-29714, SL 27.9 mm, Vietnam, Nha-Trang Bay. (e) Paratype, MNHN-IM-2009-33539, SL 32.1 mm. (f, g) MNHN-IM-2009-33548, SL 30.7 mm. (h) Paratype, MNHN-IM-2009-33564, SL 30.2 mm. (i) MNHN-IM-2009-24996, SL 36.7 mm. All specimens except D from the type locality. All shells at the same scale [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

subrectangular spots on sinus cords. Subsutural ramp with light-brown irregularly shaped blurring spots, in some parts of shells merged, producing uniform brown subsutural ramp.

Shell height 32.8 mm, shell diameter 11.1 mm, last whorl height 16.8 mm, aperture height (without canal) 9.2 mm.

Radula was examined in three specimens: MNHN-IM-2009-29714 (Vietnam), MNHN-IM-2009-33539 and MNHN-IM-2009-33548 (Philippines) (Figure 5). It is very similar in all specimens, formed by semi-enrolled unbarbed marginal teeth, edges of the marginal teeth are elevated and equally developed along entire tooth length, and delimit a trough. The anterior solid part of the tooth is absent. The marginal teeth on both sides of the radular membrane are interlocked, so that tooth of one row is lying within the trough of a tooth of the subsequent row. Central formation is formed by a small and short central cusp and more or less developed lateral flaps.

**Remarks.** The species is represented in our material by six specimens, five from Olango Island, in the Philippines and one from Vietnam. Although most specimens were collected by hookah divers in the Philippines and the exact depth is not known, the usual operation depths during collecting in Olango Island is 15–20 m, while the specimen from Vietnam was collected by SCUBA at similar depths (15–22 m). The largest specimen reaches 36.7 mm in length.

All available specimens are rather similar in shell shape and coloration, most variable is the degree of development of light-brown spots on subsutural ramp, sometimes absent on most of whorl, but always present at least on some parts.

There is not a single pure diagnostic character for *I. conotaxis* n. sp. in the COI alignment.

The species is most similar to the closely related *lotyrris musivum*; the two species sympatric and probably syntopic in Olango Island and

also occur syntopically in Nha-Trang Bay in Vietnam (Figures 6 and 7). Despite the strong similarity, the new species can be reliably distinguished from *I. musivum* by the presence of brown spots on, or sometimes nearly completely brown subsutural ramp, which is creamy in *I. musivum*. Both species have similar radulae with semi-enrolled marginal teeth. There is some similarity between the new species and *I. olangoensis*, also sympatric and syntopic in the Philippines and Vietnam, but *I. conotaxis* n. sp. can be readily distinguished by the presence of distinct brown spots on the sinus cords, which are only speckled in *I. olangoensis*. Besides the two species differ in the radular morphology as *I. olangoensis* has duplex-type marginal teeth.

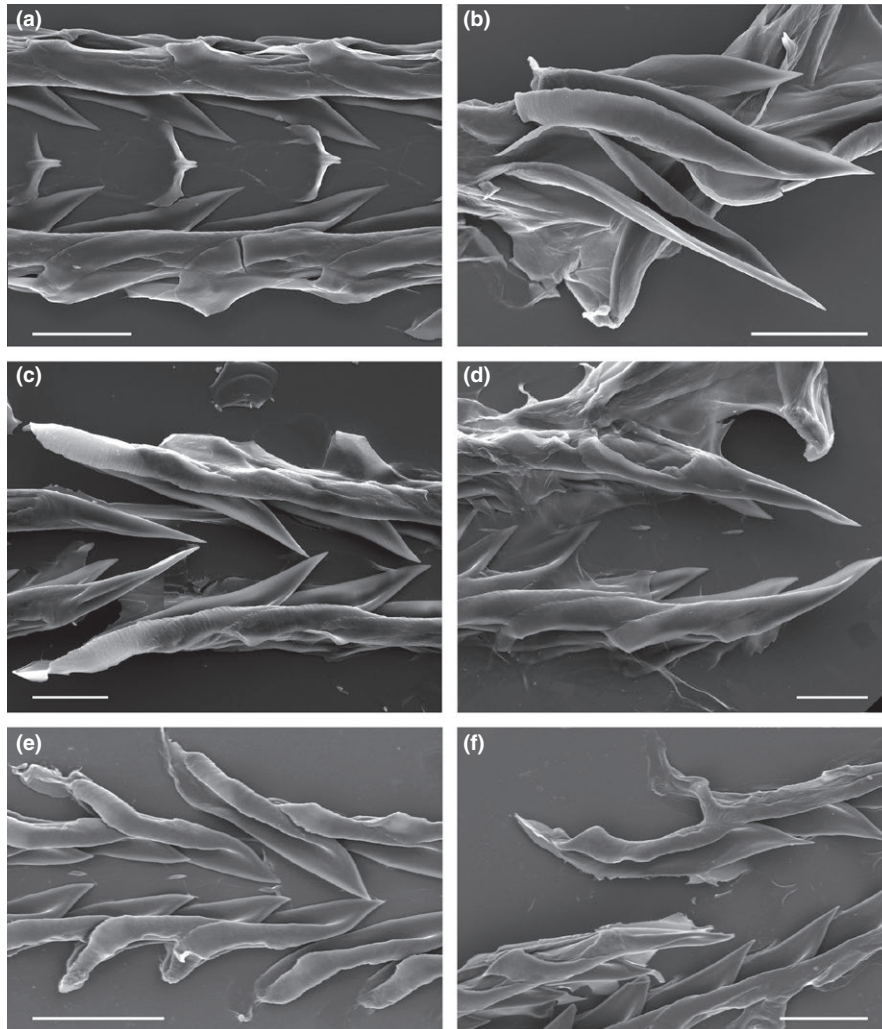
## 4 | DISCUSSION

### 4.1 | Diversity in the *Xenuroturrus/lotyrris* complex

The combination of both exploratory (ABGD, GMYC and PTP) and hypothesis-testing (BFD) methods allowed us to delimit eight SSH in

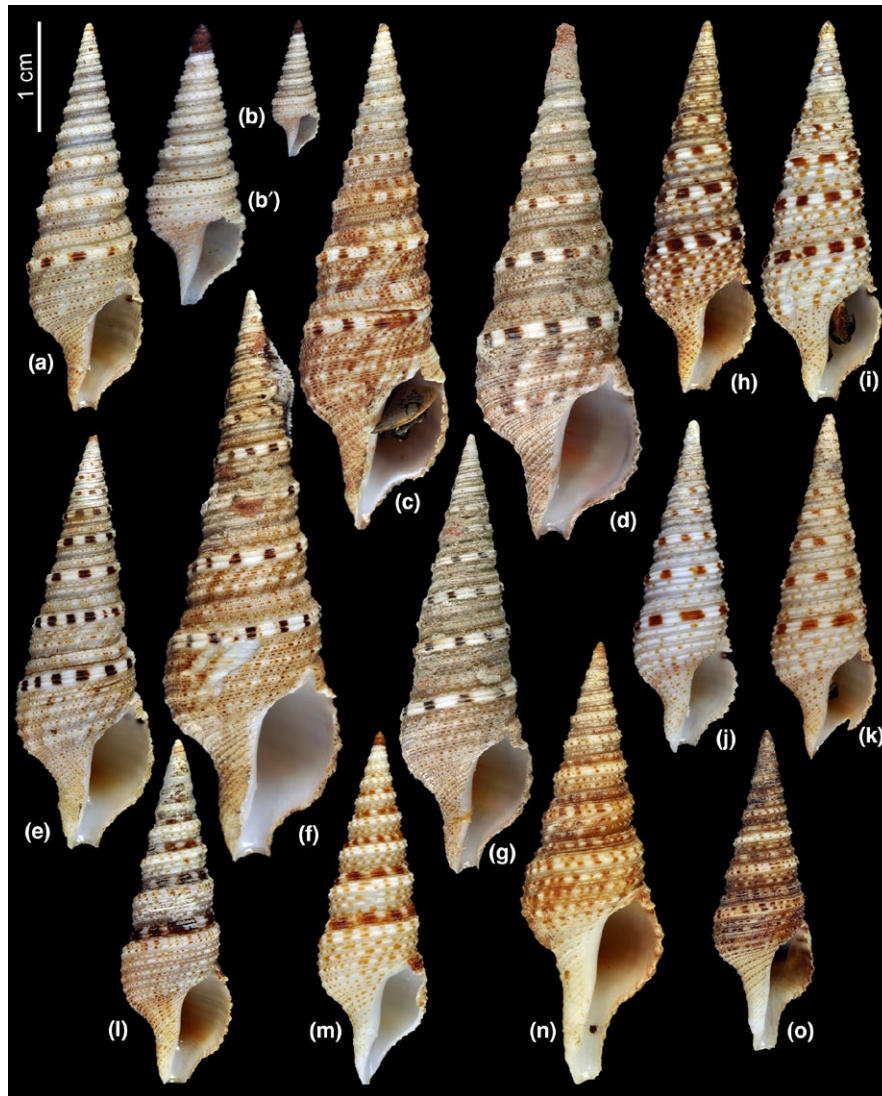
the *Xenuroturrus/lotyrris* complex, one of which being described as a new species. All these SSH are supported by both the mitochondrial and nuclear data sets, but also by diagnostic features of the shells and the radula. Although only two main types of radula morphology are recognized in the species complex, closely related species can possess different radula types (e.g., *I. olangoensis* vs. *I. conotaxis* n. sp. and *I. musivum*).

Indeed, the distribution of the two radula types in the phylogenetic tree, with species exhibiting the same radula type not clustering together, raises taxonomic and evolutionary questions. Previously the radular type (semi-enrolled vs. duplex marginal teeth) was considered as a reliable character to differentiate the genera *lotyrris* Medinskaya & Sysoev, 2001 (type species *I. marquesensis* Sysoev, 2002) with semi-enrolled teeth and *Xenuroturrus* Iredale, 1929 with the duplex ones (Kantor et al., 2008). The originally monotypical *lotyrris* Medinskaya & Sysoev, 2001, was described on the presence of a semi-enrolled marginal radular teeth, unique—at that time—for the Turridae. The type species, *I. marquesensis*, was not



**FIGURE 5** Radulae of *lotyrris conotaxis* n. sp. (a–d) and *I. musivum* (e, f). *lotyrris conotaxis* n. sp.: (a, b) MNHN-IM-2009-29714, SL 27.9 mm, Vietnam, Nha-Trang Bay (shell on Figure 4d). (c) MNHN-IM-2009-33548, SL 30.7 mm, Philippines (shell on Figure 4f,g). (d) MNHN-IM-2009-33539, SL 32.1 mm, paratype, Philippines (shell on Figure 4e). *lotyrris musivum*: (e) holotype, MNHN-IM-2009-26290, Vanuatu. (f) Loyalty Islands, Atelier LIFOU 2000, st. 1421. Scale bars – 50  $\mu$ m



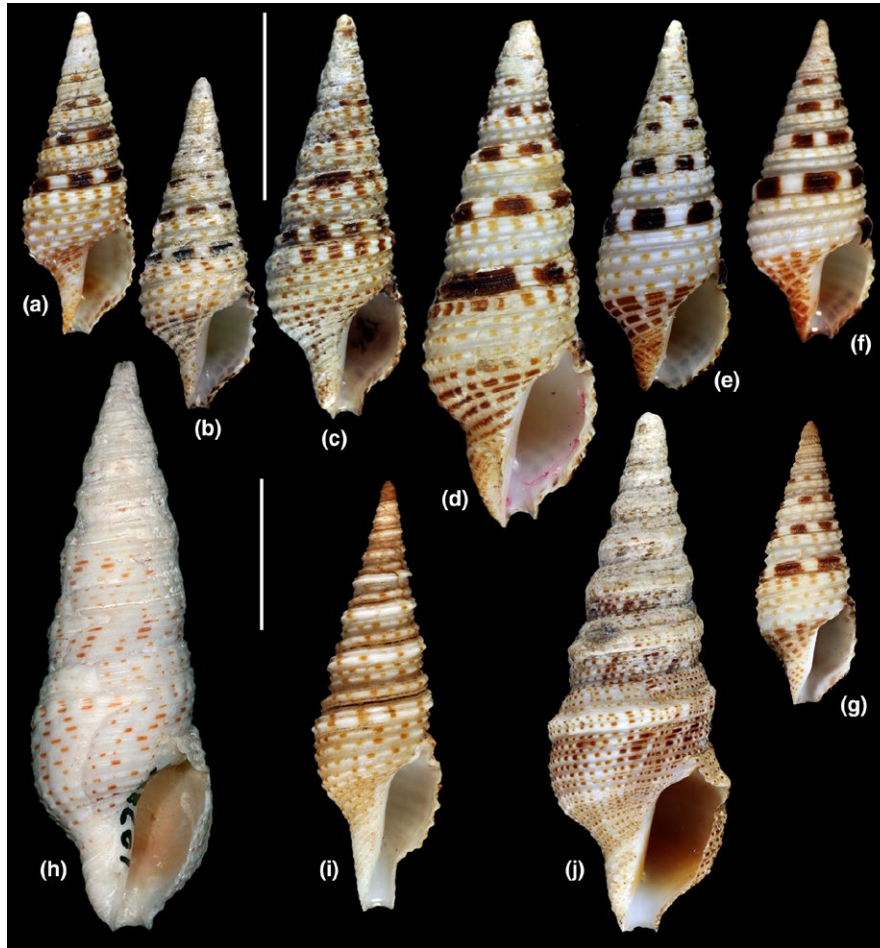


**FIGURE 6** Shells of *Xenuroturris/lotyrris* complex. (a–d). *Xenuroturris legitima*. (a) MNHN-IM-2009-24936, Mozambique, SL 35.5 mm. (b, b'). MNHN-IM-2009-17246, Madagascar, SL 12.3 mm. (c) MNHN-IM-2009-29710, Vietnam, SL 46.7 mm. (d) MNHN-IM-2009-33532, Philippines, SL 47.3 mm. (e–g) *lotyrris cingulifera*. (e) MNHN-IM-2009-24942, Mozambique, SL 37.9 mm. (f) MNHN-IM-2009-29709, Vietnam, SL 53.9 mm. (g) MNHN-IM-2009-33545, Philippines, SL 39.9 mm. (h–k) *lotyrris musivum*. (h) MNHN-IM-2009-29715, Vietnam, SL 33.9 mm. (i) MNHN-IM-2009-24995, Philippines, SL 34.3 mm. (j) MNHN-IM-2009-33557, Philippines, SL 30.1 mm. (k) MNHN-IM-2009-24997, Philippines, SL 31.2 mm. (l, m) *lotyrris olangoensis*. (l) MNHN-IM-2009-33561, Philippines, SL 31.4 mm. (m) MNHN-IM-2009-24959, New Caledonia, SL 31.7 mm. (n, o) *lotyrris notata*. (n) MNHN-IM-2009-29726, Vietnam, SL 40.2 mm. (o) MNHN-IM-2009-29720, Vietnam, SL 28.9 mm. All specimens sequenced. Shells at the same scale. The shells of the same species are arranged geographically in general direction from west to east. The colour background of the letters corresponds to the colours, used in the map on Figure 1 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and is still not sequenced. With the enlarged data set, the situation became more complex: three species with duplex marginal radular teeth were confidently placed into *lotyrris* clade, namely *lotyrris olangoensis*, *I. notata* and *I. kingae*, and the supposed congruence between the phylogeny and the radula type is thus no longer true. It may also appear that *I. marquesensis* does not belong to this clade and therefore the use of the name *lotyrris* may not be appropriate, or the name itself may appear to be a synonym. Furthermore, within the same type (semi-enrolled or duplex teeth), the radular morphology can be quite different. Thus in *I. marquesensis*, the semi-enrolled marginal teeth have a distinct barb on the dorsal limb (Sysoev, 2002):

Figure 2c,d), which is not pronounced in any species here assigned to *lotyrris*. The central formation is absent in *I. marquesensis* and *I. cingulifera*, while being rather distinct in *I. devozei* and *I. conotaxis* n. sp. Similarly the shape of duplex radular teeth in the first radular type is quite variable, as well as the degree of development of the central formation (Figure 3a–d).

Contrary to what Kantor et al. (2008) have concluded (“radular type is indeed reliable for revealing relationships”), the evolution of the radulae types does not follow the phylogeny, and multiple convergence and/or reversion would be needed to explain the observed pattern. Remarkable is the difference in the radula between



**FIGURE 7** Shells of *Xenuroturrus/lotyrris* complex. (a–c). *lotyrris devoizei*. (a) MNHN-IM-2009-24941, Mozambique, SL 17.1 mm. (b) Holotype, MNHN-IM-2000-20014, Vanuatu, SL 17.8 mm. (c) MNHN-IM-2013-40058, Marquesas Islands, SL 21.6 mm. (d) New Caledonia, Expédition MONTRouZIER, st. 1271, SL 27.9 mm. (e–g) *lotyrris kingae*. (e) MNHN-IM-2009-33554, Philippines, SL 19.9 mm. (f) New Caledonia, Expédition MONTRouZIER, st. 1319, SL 17.3 mm. (g) Holotype, Bishop Museum, BPBM 9970, SL 15 mm. (h) ?*Xenuroturrus cerithiformis*, holotype, USNM 338601, SL 38.5 mm. (i) *lotyrris marquesensis*, holotype, MNHN-IM-2000-3074, SL 28.8 mm. (j) *lotyrris millepunctata*, MNHN-IM-2013-57331, New Caledonia, SL 34.6 mm. (a–g)—at the same scale. (h–j)—at the same scale. Scale bars 1 cm. (a, b, c, e, j)—sequenced specimens. Figure (g)—courtesy of Bishop Museum, Honolulu, Hawaii, USA [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

sister species (nearly identical in shell morphology and coloration) *I. kingae* and *I. devoizei* (Figure 3d,f, correspondingly for radula and Figure 7 for shells). One hypothesis is that the radula is more labile than is commonly assumed and its morphology in Turridae may be related to diet, as it has been shown in the related family Conidae (Kohn, Nishi, & Pernet, 1999; Tucker & Tenorio, 2009, p. 200). This hypothesis would also suggest that diet itself is labile, with closely related species being characterized by different diets, a hypothesis already proposed for cone snails (Kohn & Orians, 1962). Identifying the preys would therefore be necessary to test a potential correlation between the radula type and the food habit, thus explaining the high lability of this anatomical structure.

There remains several nonsequenced species probably belonging to *Xenuroturrus/lotyrris* complex, some with known radular type (based on published and our unpublished data) – *I. marquesensis* (semi-enrolled) (Figure 7i), *X. millepunctata* (Sowerby, 1909) (duplex) (Figure 7j), *X. gemmuloides* Powell, 1964 (duplex), *X. cerithiformis*

Powell, 1964 (duplex) (Figure 7h), as well as the more enigmatic *X.? castanella* Powell, 1964 and *X.? emmae* Bozetti, 1994. Their generic position can be clarified only based on molecular studies.

#### 4.2 | Species delimitation methodology

As emphasized in the introduction, DNA-based species delimitation methods are generally designed either for less-known taxa, using monolocus data, to propose *de novo* species hypotheses, or for more difficult to tackle species complexes, using multilocus data, to test predefined competing partitions of species. The two strategies, that we can term as “exploratory” and “hypothesis-testing,” respectively, are rarely combined. When predefined partitions of species are already available, and disagreements among taxonomists on species boundaries exist, exploratory methods will be mostly useless, and hypothesis-testing methods will be favoured to identify the most likely species partition. On the contrary, when dealing with a largely



unknown group, exploratory methods will be favoured to propose PSH. Here, we combined both strategies, applying a monolocus, COI barcode, data set, and a multilocus, RAD-seq, data set, being used to test alternative partitions proposed by the former.

The methods applied to the COI data set, now widely accepted as robust and congruent (see e.g., Kekkonen, Mutanen, Kaila, Nieminen, & Hebert, 2015; Schwarzfeld & Sperling, 2015), all proposed similar partitions in the case of the *Xenuroturrus/lotyrris* complex. However, each method has their own limitations (Carstens et al., 2013; Miralles & Vences, 2013; Reid & Carstens, 2012) (e.g., linked to the substitution model used to calculate the genetic distances or to the quality of the input phylogenetic tree, sensitivities to the quality of the sampling (Ahrens et al., 2016; Hamilton, Hendrixson, Brewer, & Bond, 2014)), justifying the use of several methods.

More importantly, most of the PSH proposed with the COI gene were confirmed by the analysis of the RAD-seq data set. Interestingly, the phylogenetic reconstructions based on RAD-seq data sets exhibiting varying levels of missing data were quite congruent. Trees based on more loci, but consequently on higher levels of missing data, were better resolved, even for deeper relationships. This counter-intuitive result has been observed and discussed recently by several studies (see Eaton, Spriggs, Park, & Donoghue, 2017; Tripp, Tsai, Zhuang, & Dexter, 2017 for recent reviews). Consensus on the position to deal with missing data in RAD-seq analysis has not been reached and several studies strongly advocate for a case to case approach since in some situations, as selecting loci to decrease the proportion of missing data can possibly generate a strong bias (see for example Huang & Knowles, 2014).

In addition to the phylogenetic approach applied to the RAD-seq data set, the BFD method clearly identified, for both *X. legitima* and *I. cingulifera*, the hypothesis considering all Indo-Pacific samples as only one species as the more likely. This was in comparison with a partition in two species each, one in the WIO and the other in the IWP. Here again, the level of missing data probably led to unstable MLE when various number of steps and chain length for the path sampling are used, but the important difference between the MLE of the two competing partitions convinced us that the BF was decisive.

Thus, the successive use of exploratory and hypothesis-testing methods was necessary to reject the hypothesis that allopatric populations within *X. legitima* and *I. cingulifera* constitute different species. Indeed, analysing the COI data set only would have led to recognize these allopatric populations as species, as supported by all the exploratory methods (Figure 2). As exemplified here, the tree topology obtained with two distant populations within species can easily mimic the topology obtained with two sister species, even when considering the genetic distances: the genetic distances between the two allopatric pairs within *X. legitima* and *I. cingulifera* (K2P distances = 1.7%–3.6%) are actually similar to the genetic distances between *I. devoizei* and *I. kingae* (1.5%–2.6%) and *I. conotaxis n. sp.*, *I. musivum* and *I. olangoensis* (1.9%–3%). Applying various methods of species delimitation remains the best strategy to counter-balance the limitations of each, as discussed before, but more important is the joint analysis of independent genetic markers, since gene trees,

taken independently, are not necessarily congruent with the species tree (Degnan & Rosenberg, 2009).

Consequently, and except the case of *I. devoizei* and *I. notata*, all the species are present in sympatry (in the Philippines). However, a scenario in which allopatric speciation with subsequent changes in the distribution areas cannot be ruled out (Berlocher, 1998; Chesser & Zink, 1994). Following the criteria of Coyne and Orr (2004), determining the age of the speciation events could help to test these two scenarios, but in the absence of calibration points (either biogeographic events or fossils), we were not able to reconstruct a dated tree. Similarly, exploring whether, if *Xenuroturrus* species may have diverged in sympatry (e.g., by niche partitioning linked to the apparition of new toxins and prey shifts (Duda & Lee, 2009; Fedosov, Tiunov, Kiyashko, & Kantor, 2014; Nicolas Puillandre et al., 2014)) or not would require sequencing the transcriptomes of the venom glands and identifying the preys of different species, and in particular, in the only pair of sympatric sister species, *I. musivum* and *I. conotaxis n. sp.*

## 5 | CONCLUSION

The combined use of classical barcoding, next generation sequencing and morphological observations enabled us to give new insight into the evolution of this species complex. Most defined SSH were linked to already described species and one new species is described. This study demonstrates the utility of combining both exploratory and hypothesis-testing methods. In the absence of primary species hypotheses, or, as it is the case here, when morphology-based species hypotheses are doubtful, analysing monolocus data with exploratory methods such as ABGD, GMYC and PTP rapidly produce PSH. However, such PSH had to be taken cautiously, since high levels of divergence can also be observed between populations within species (and conversely, low genetic distances can result from a lack of variability between species for a given marker). Evaluating these PSHs with independent markers using a hypothesis-testing method constitutes a desirable strategy to tell apart SSH from, for example, populations within species.

## ACKNOWLEDGEMENTS

A large part of the molecular material in this paper originates from various shore-based expeditions and deep sea cruises, conducted respectively by MNHN (Inhaca 2011); by MNHN, Pro-Natura International (PNI) and Institut de Recherche pour le Développement (IRD) as part of the Our Planet Reviewed programme (Santo 2006, Atimo Vatae, Papua Niugini); and by MNHN and Institut de Recherche pour le Développement (IRD) as part of the Tropical Deep-Sea Benthos programme (Mainbaza, Terrasses, Pakaihi I Te Moana). In-country partners include the Maritime College, Luganville; Universidade Eduardo Mondlane, Maputo and University of Papua New Guinea, Port Moresby. Funders and sponsors include the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, Richard Lounsbury

Foundation, Vinci Entrepouse Contracting, Fondation EDF, the French Ministry of Foreign Affairs, the French Fonds Pacifique and the Government of New Caledonia. The specimens from the Philippines used in this study were obtained by AF in conjunction with a collection trip supported in part by the “Conus-Turrid” project (principal investigator B. M. Olivera, University of Utah, USA). Collection of material in Vietnam was supported by the Russian-Vietnamese Tropical Center. We are thankful to the staff of the Tropical Center for assistance in organization of the field sampling and loan of some laboratory equipment. Peter Stahlschmidt collected the specimens from Egypt, with a permit from the Hurghada Environmental Protection and Conservation Association (Hurghada, Egypt). 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 also thank Virginie Héros, Julien Brisset, Philippe Maestrati and Manuel Caballer Gutierrez for their help in curating specimens, and Mark Phuong, Sarah Samadi and Guillaume Achaz for constructive comments on the manuscript. We are grateful to Dr. Norine Yeung from Bishop Museum, Honolulu Hawaii for providing the photos of the types of *Xenuroturrus kingae*. This project was partly supported by the Service de Systématique Moléculaire (UMS 2700 CNRS-MNHN) and by the project CONOTAX, funded by the French ANR (grant number ANR-13-JSV7-0013-01). The contribution of Yu.I. Kantor and A.E. Fedosov was supported by the grant No. 16-14-10118 from the Russian Science Foundation (principal investigator Yu.I. Kantor). The scanning electron microscopy was conducted using Joint Usage Center “Instrumental methods in ecology” at the IEE RAS.

#### AUTHOR CONTRIBUTION

J.A. and L.A.C. performed the molecular experiments; B.B. vouchered and registered the samples; J.A., P.Z. and N.P. analysed the data; A.F. and Y.K. performed the morphological analyses and described the new species; N.P., B.B., A.F. and Y.K. participated in the field sampling; all authors participated in the research design and wrote the manuscript.

#### DATA ACCESSIBILITY

All samples are vouchered in the MNHN collection. Sample data and COI sequences will be uploaded in BOLD. COI sequences will be uploaded in GenBank. RAD-seq data and COI and RAD-seq trees have been uploaded in Dryad: <https://doi.org/10.5061/dryad.k2q42>

#### ORCID

Nicolas Puillandre  <http://orcid.org/0000-0002-9797-0892>

#### REFERENCES

Ahrens, D., Fujisawa, T., Krammer, H.-J., Eberle, J., Fabrizi, S., & Vogler, A. P. (2016). Rarity and incomplete sampling in DNA-based

- species delimitation. *Systematic Biology*, 65(3), 478–494. <https://doi.org/10.1093/sysbio/syw002>
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, 3(10), e3376. <https://doi.org/10.1371/journal.pone.0003376>
- Barberousse, A., & Samadi, S. (2010). Species from Darwin onward. *Integrative Zoology*, 5, 187–197. <https://doi.org/10.1111/j.1749-4877.2010.00204.x>
- Berlacher, S. H. (1998). Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence. In *Endless forms: Species and speciation* (pp. 99–113). Oxford, UK: Oxford University Press.
- Boucher, F. C., Casazza, G., Szövényi, P., & Conti, E. (2016). Sequence capture using RAD probes clarifies phylogenetic relationships and species boundaries in *Primula* sect. *Auricula*. *Molecular Phylogenetics and Evolution*, 104, 60–72. <https://doi.org/10.1016/j.ympev.2016.08.003>
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., ... Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10(4), e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>
- Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29(8), 1917–1932. <https://doi.org/10.1093/molbev/mss086>
- Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, 22, 4369–4383. <https://doi.org/10.1111/mec.12413>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chesser, R. T., & Zink, R. M. (1994). Modes of speciation in birds: A test of Lynch's method. *Evolution*, 48(2), 490–497. <https://doi.org/10.1111/j.1558-5646.1994.tb01326.x>
- Coissac, E., Hollingsworth, P. M., Laverigne, S., & Taberlet, P. (2016). From barcodes to genomes: Extending the concept of DNA barcoding. *Molecular Ecology*, 25(7), 1423–1428. <https://doi.org/10.1111/mec.13549>
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.
- De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56, 879–886. <https://doi.org/10.1080/10635150701701083>
- Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, 24(6), 332–340. <https://doi.org/10.1016/j.tree.2009.01.009>
- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214. <https://doi.org/10.1186/1471-2148-7-214>
- Duda, T. F. Jr, Bolin, M. B., Meyer, C., & Kohn, A. J. (2008). Hidden diversity in a hyperdiverse gastropod genus: Discovery of previously unidentified members of a *Conus* species complex. *Molecular Phylogenetics and Evolution*, 49, 867–876.
- Duda, T. F. Jr, & Lee, T. (2009). Ecological release and venom evolution of a predatory marine Snail at Easter Island. *PLoS ONE*, 4, e5558. <https://doi.org/10.1371/journal.pone.0005558>
- Eaton, D. A. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, 30(13), 1844–1849. <https://doi.org/10.1093/bioinformatics/btu121>
- Eaton, D. A., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, 66(3), 399–412.

- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797.
- Ence, D. D., & Carstens, B. C. (2011). SpedeSTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources*, 11, 473–480. <https://doi.org/10.1111/j.1755-0998.2010.02947.x>
- Fedosov, A. E., Stahlschmidt, P., Puillandre, N., Aznar-Cormano, L., & Bouchet, P. (2017). Not all spotted cats are leopards: Evidence for a *Hemilienardia ocellata* species complex (Gastropoda: Conoidea: Raphitomidae). *European Journal of Taxonomy*, 268, 1–20. <https://doi.org/10.5852/ejt.2017.268>
- Fedosov, A., Tiunov, A., Kiyashko, S., & Kantor, Y. I. (2014). Trophic diversification in the evolution of predatory marine gastropods of the family Terebridae as inferred from stable isotope data. *Marine Ecology Progress Series*, 497, 143–156. <https://doi.org/10.3354/meps10585>
- Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.
- Fujisawa, T., & Barraclough, T. G. (2013). Delimiting Species Using Single-locus Data and the Generalized Mixed Yule Coalescent (GMYC) Approach: A Revised Method and Evaluation on Simulated Datasets. *Systematic Biology*. Retrieved from <http://sysbio.oxfordjournals.org/content/early/2013/05/16/sysbio.syt033.abstract>
- Galindo, L. A., Puillandre, N., Strong, E. E., & Bouchet, P. (2014). Using microwaves to prepare gastropods for DNA barcoding. *Molecular Ecology Resources*, 14(4), 700–705. <https://doi.org/10.1111/1755-0998.12231>
- Grummer, J. A., Bryson, R. W. Jr, & Reeder, T. W. (2013). Species delimitation using Bayes factors: Simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology*, 63(2), 119–133.
- Hamilton, C. A., Hendrixson, B. E., Brewer, M. S., & Bond, J. E. (2014). An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: A case study of the North American *Tarantula* genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution*, 71, 79–93. <https://doi.org/10.1016/j.ympev.2013.11.007>
- Herrera, S., & Shank, T. M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution*, 100, 70–79. <https://doi.org/10.1016/j.ympev.2016.03.010>
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Le, S. V. (2017). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.
- Huang, H., & Knowles, L. L. (2014). Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of RAD sequences. *Systematic Biology*, 65, 357–365.
- Huelsenbeck, J. P., Ronquist, F., & Hall, B. (2001). MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, 17, 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
- Jablonski, D., & Lutz, R. A. (1980). Molluscan larval shell morphology - ecological and paleontological applications. In D. C. Rhoads, & R. A. Lutz (Eds.), *Skeletal growth of aquatic organisms* (pp. 323–377). New York, NY: Plenum Press. <https://doi.org/10.1007/978-1-4899-4995-0>
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587. <https://doi.org/10.1038/nmeth.4285>
- Kantor, Y. I. (2006). On the morphology and homology of the “central tooth” in the radula of Turrinae (Conoidea: Turridae). *Ruthenica*, 16(1–2), 47–52.
- Kantor, Y. I., & Puillandre, N. (2012). Evolution of the radular apparatus in Conoidea (Gastropoda: Neogastropoda) as inferred from a molecular phylogeny. *Malacologia*, 55, 55–90. <https://doi.org/10.4002/040.055.0105>
- Kantor, Y. I., Puillandre, N., Olivera, B. M., & Bouchet, P. (2008). Morphological proxies for taxonomic decision in turrids (Mollusca, Neogastropoda): A test of the value of shell and radula characters using molecular data. *Zoological Science*, 25, 1156–1170. <https://doi.org/10.2108/zsj.25.1156>
- Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2016). Multi-rate Poisson Tree Processes for single-locus species delimitation under Maximum Likelihood and Markov Chain Monte Carlo. *BioRxiv*, 063875.
- Kekkonen, M., & Hebert, P. D. N. (2014). DNA barcode-based delimitation of putative species: Efficient start for taxonomic workflows. *Molecular Ecology Resources*, 14(4), 706–715. <https://doi.org/10.1111/1755-0998.12233>
- Kekkonen, M., Mutanen, M., Kaila, L., Nieminen, M., & Hebert, P. D. N. (2015). Delineating species with DNA barcodes: A case of taxon dependent method performance in moths. *PLoS ONE*, 10(4), e0122481. <https://doi.org/10.1371/journal.pone.0122481>
- Kess, T., Gross, J., Harper, F., & Boulding, E. G. (2015). Low-cost ddRAD method of SNP discovery and genotyping applied to the periwinkle *Littorina saxatilis*. *Journal of Molluscan Studies*, 82(1), eyv042. <https://doi.org/10.1093/mollus/eyv042>
- Kohn, A. J., Nishi, M., & Pernet, B. (1999). Snail spears and scimitars: A character analysis of *Conus* radular teeth. *Journal of Molluscan Studies*, 65(4), 461–481. <https://doi.org/10.1093/mollus/65.4.461>
- Kohn, A. J., & Orians, G. H. (1962). Ecological data in the classification of closely related species. *Systematic Zoology*, 11, 119–127. <https://doi.org/10.2307/2411874>
- Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, 63(4), 534–542. <https://doi.org/10.1093/sysbio/syu018>
- Leavitt, S. D., Moreau, C. S., & Lumbsch, H. T. (2015). The dynamic discipline of species delimitation: Progress toward effectively recognizing species boundaries in natural populations. In D. K. Upreti, P. K. Divakar, V. Shukla, & R. Bajpai (Eds.), *Recent advances in lichenology* (pp. 11–44). New Delhi: Springer. <https://doi.org/10.1007/978-81-322-2235-4>
- Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2006). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, 17(2), 240–248.
- Miralles, A., & Vences, M. (2013). New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS ONE*, 8, e68242. <https://doi.org/10.1371/journal.pone.0068242>
- Monaghan, M. T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D. J. G., ... Vogler, A. P. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, 58, 298–311. <https://doi.org/10.1093/sysbio/syp027>
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2014). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274.
- Pante, E., Abdelkrim, J., Viricel, A., Gey, D., France, S., Boisselier, M.-C., & Samadi, S. (2015). Use of RAD sequencing for delimiting species. *Heredity*, 114(5), 450–459. <https://doi.org/10.1038/hdy.2014.105>
- Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., ... Samadi, S. (2015). Species are hypotheses: Avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, 24(3), 525–544. <https://doi.org/10.1111/mec.13048>
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, 55, 595–609. <https://doi.org/10.1080/10635150600852011>

- Puillandre, N., Baylac, M., Boisselier, M. C., Cruaud, C., & Samadi, S. (2009). An integrative approach of species delimitation in the genus *Benthomangelia* (Mollusca: Conoidea). *Biological Journal of the Linnean Society*, 96, 696–708. <https://doi.org/10.1111/j.1095-8312.2008.01143.x>
- Puillandre, N., Cruaud, C., & Kantor, Y. I. (2010). Cryptic species in *Gemmuloborsonia* (Gastropoda: Conoidea). *Journal of Molluscan Studies*, 73, 11–23. <https://doi.org/10.1093/mollus/eyp042>
- Puillandre, N., Fedosov, A. E., Zaharias, P., Aznar-Cormano, L., & Kantor, Y. I. (2017). A quest for the lost types of *Lophiotoma* (Gastropoda: Conoidea: Turridae): Integrative taxonomy in a nomenclatural mess. *Zoological Journal of the Linnean Society*, 181(2), 243–271. <https://doi.org/10.1093/zoolinnean/zlx012>
- Puillandre, N., Kantor, Y., Sysoev, A., Couloux, A., Meyer, C., Rawlings, T., ... Bouchet, P. (2011). The dragon tamed? A molecular phylogeny of the Conoidea (Mollusca, Gastropoda). *Journal of Molluscan Studies*, 77, 259–272. <https://doi.org/10.1093/mollus/eyr015>
- Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 21, 1864–1877. <https://doi.org/10.1111/j.1365-294X.2011.05239.x>
- Puillandre, N., Modica, M. V., Zhang, Y., Sirovitch, L., Boisselier, M.-C., Cruaud, C., ... Samadi, S. (2012). Large scale species delimitation method for hyperdiverse groups. *Molecular Ecology*, 21, 2671–2691. <https://doi.org/10.1111/j.1365-294X.2012.05559.x>
- Puillandre, N., Stöcklin, R., Favreau, P., Bianchi, E., Perret, F., Rivasseau, A., ... Bouchet, P. (2014). When everything converges: Integrative taxonomy with shell, DNA and venom data reveals *Conus conco*, a new species of cone snails (Gastropoda: Conoidea). *Molecular Phylogenetics and Evolution*, 80, 186–192. <https://doi.org/10.1016/j.ympev.2014.06.024>
- Puillandre, N., Sysoev, A., Olivera, B. M., Couloux, A., & Bouchet, P. (2010). Loss of planktotrophy and speciation: Geographical fragmentation in the deep-water gastropod genus *Bathytoma* (Gastropoda, Conoidea) in the western Pacific. *Systematics and Biodiversity*, 8, 371–394. <https://doi.org/10.1080/14772001003748709>
- Rambaut, A., & Drummond, A. J. (2014). *Tracer v1.6*. Retrieved from <http://beast.bio.ed.ac.uk/Tracer>.
- Reid, N. M., & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: A Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology*, 12(1), 196. <https://doi.org/10.1186/1471-2148-12-196>
- Rutschmann, S., Detering, H., Simon, S., Fredslund, J., & Monaghan, M. T. (2017). DISCOMARK: Nuclear marker discovery from orthologous sequences using draft genome data. *Molecular Ecology Resources*, 17(2), 257–266. <https://doi.org/10.1111/1755-0998.12576>
- Schwarzfeld, M. D., & Sperling, F. A. (2015). Comparison of five methods for delimitating species in *Ophion* Fabricius, a diverse genus of parasitoid wasps (Hymenoptera, Ichneumonidae). *Molecular Phylogenetics and Evolution*, 93, 234–248. <https://doi.org/10.1016/j.ympev.2015.08.003>
- Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>
- Sysoev, A. (2002). On the type species of *Iotyrris* Mediskinya et Sysoev, 2001 (Gastropoda, Turridae). *Ruthenica*, 12, 169–171.
- Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: Insights from a butterfly survey. *Methods in Ecology and Evolution*, 4(12), 1101–1110. <https://doi.org/10.1111/2041-210X.12107>
- Tripp, E. A., Tsai, Y. E., Zhuang, Y., & Dexter, K. G. (2017). RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. *Ecology and Evolution*, 7(19), 7920–7936. <https://doi.org/10.1002/ece3.3274>
- Tucker, J. K., & Tenorio, M. J. (2009). *Systematic classification of recent and fossil conoidean gastropods*. Hackenheim, Germany: Conchbooks.
- Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, 31, 3125–3235. <https://doi.org/10.1093/molbev/msu279>
- Zhang, J., Kapli, R., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29, 2869–2876. <https://doi.org/10.1093/bioinformatics/btt499>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Abdelkrim J, Aznar-Cormano L, Buge B, et al. Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Mol Ecol*. 2018;27:4591–4611. <https://doi.org/10.1111/mec.14882>







## Taxonomie intégrative, phylogénie et diversification des Turridae (Gastropoda, Conoidea)

### Résumé

Les Turridae constituent un groupe de gastéropodes marins prédateurs venimeux, présents dans presque toute la zone pantropicale et des côtes à plus de 3000 m de profondeur. Étant donné leur faible diversité (environ 200 espèces décrites), relativement à leurs taxons frères au sein des Conoidea, et l'absence apparente d'innovation-clé spécifique à ce groupe, la question de la thèse est la suivante : quelles sont les dynamiques de diversification et les moteurs de cette diversification chez les Turridae ? Toutefois, il n'est possible d'y répondre qu'à condition d'avoir un cadre systématique robuste. Les approches traditionnelles de systématique, basées sur des considérations morpho-anatomiques, sont confrontées à des problèmes de plasticité de la coquille et des organes, que ce soit dans des temps évolutifs courts ou longs. L'outil moléculaire a donc été privilégié pour répondre à cette question.

Dans une première partie, un pipeline de taxonomie intégrative a été défini et utilisé pour délimiter des hypothèses d'espèces chez les Turridae. Ce pipeline a résulté en environ 200 hypothèses d'espèces, dont il n'est pas toujours facile de savoir si elles ont déjà été décrites ou non. J'explore cette problématique avec une analyse poussée du genre *Cryptogemma*. L'acquisition semi-automatisée de données morphométriques a permis d'attribuer avec confiance 23 spécimens porte-noms à 8 espèces délimitées, illustrant la confusion taxonomique qui existe chez les Turridae.

Dans une seconde partie, je construis une phylogénie des Turridae en utilisant une approche de « génome-réduit ». Ces approches ont permis de séquencer environ 4000 loci pour 110 hypothèses d'espèce. Les topologies retrouvées mettent en évidence de façon robuste la polyphylie du genre *Gemmula*, au sein duquel 13 genres potentiellement doivent être décrits.

Enfin, j'utilise les résultats combinés de la taxonomie intégrative et la phylogénie, ainsi que des méthodes récentes pour étudier la dynamique de diversification de groupe. La famille des Turridae apparaît comme un groupe dont les taux de diversification passés seraient plus importants qu'actuellement. Des liens ont été trouvés entre la diversification et certains traits comme la protoconque, la forme et la taille de la coquille et la bathymétrie.

La stratégie de systématique utilisée pendant cette thèse s'est révélée efficace pour permettre à répondre à des questions-clés liées à la diversification des gastéropodes marins, proposant des résultats inédits.

## Integrative taxonomy, phylogeny and diversification of the Turridae (Gastropoda, Conoidea)

### Abstract

The Turridae are a group of predatory venomous marine gastropods, found in almost all the pantropical zone, from shallow-water to more than 3000m deep. Considering their low diversity (about 200 described species), relatively to their sister taxa within the Conoidea, and the apparent absence of specific key innovation, the question of this PhD is: what is the diversification dynamic and the drivers of diversification within the Turridae? However, it is possible to answer to this question only with a robust systematic framework. Traditional approaches of systematics, based on morpho-anatomical considerations only, are confronted to problems of shell and organ plasticity, whether it is on a short or long evolutionary timeframe. Hence, molecular tools were privileged to answer this question.

In the first part, a taxonomic pipeline was defined and used to delimit species hypotheses within the Turridae. This pipeline resulted in about 200 species hypotheses, for which it is not always easy to know if they have been previously described or not. I explore this problematic with an analysis of *Cryptogemma*. The semi-automatized acquisition of morphometric data allowed attributing 23 name-bearing specimens to the 8 delimited species with more confidence, illustrating the great taxonomic confusion in the Turridae.

In the second part, I constructed a phylogeny of Turridae using a "reduced-genome" approach. These approaches allowed me to sequence about 4000 loci for 110 species hypotheses. The retrieved topologies showed robust evidence for the paraphyly of *Gemmula*, for which 13 new genera can potentially be described.

Finally, I combined the results from integrative taxonomy and phylogeny, as well as recent methods to study diversification dynamics. The Turridae family appears as a group with higher past diversification rates than present ones. Furthermore, links have been found between diversification rates and some traits like protoconch, shell shape and size and bathymetry.

The systematic strategy applied in this PhD show great efficiency and allowed answering key evolutionary questions linked to diversification in marine gastropods, with unprecedented results.