Taxonomie intégrative des Turridae : phylogénie, délimitation d'espèces et barcoding

Nicolas Puillandre

MUSEUM NATIONAL D'HISTOIRE NATURELLE

Ecole doctorale SCIENCES DE LA NATURE ET DE L'HOMME (ED 227)

Thèse pour obtenir le grade de
DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE

Discipline : Systématique Evolutive

Présentée et soutenue publiquement par

NICOLAS PUILLANDRE

Le 19 septembre 2008

Taxonomie intégrative des Turridae :

phylogénie, délimitation d'espèces et barcoding

Directeurs de thèse : Marie-Catherine Boisselier et Philippe Bouchet

Jury

Frédéric AUSTERLITZ (CNRS, Orsay)	Rapporteur
Thierry BACKELJAU (Université d'Anvers & IRScNB, Bruxelles)	Rapporteur
Marie-Catherine BOISSELIER (CNRS, Paris)	Directeur
Philippe BOUCHET (MNHN, Paris)	Directeur
Patrice DAVID (CNRS, Montpellier)	Examinateur
Paul HEBERT (Université de Guelph, Canada)	Examinateur
Guillaume LECOINTRE (MNHN, Paris)	Examinateur
Sarah SAMADI (IRD, Paris)	Examinateur

MUSEUM NATIONAL D'HISTOIRE NATURELLE

Ecole doctorale SCIENCES DE LA NATURE ET DE L'HOMME (ED 227)

Thèse pour obtenir le grade de
DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE

Discipline : Systématique Evolutive

Présentée et soutenue publiquement par

NICOLAS PUILLANDRE

Le 19 septembre 2008

Taxonomie intégrative des Turridae :

phylogénie, délimitation d'espèces et barcoding

Directeurs de thèse : Marie-Catherine Boisselier et Philippe Bouchet

Jury

Frédéric AUSTERLITZ (CNRS, Orsay)	Rapporteur
Thierry BACKELJAU (Université d'Anvers & IRScNB, Bruxelles)	Rapporteur
Marie-Catherine BOISSELIER (CNRS, Paris)	Directeur
Philippe BOUCHET (MNHN, Paris)	Directeur
Patrice DAVID (CNRS, Montpellier)	Examinateur
Paul HEBERT (Université de Guelph, Canada)	Examinateur
Guillaume LECOINTRE (MNHN, Paris)	Examinateur
Sarah SAMADI (IRD, Paris)	Examinateur

A mes parents, pour tout ce que je leur dois.

Remerciements

Au cours de ces trois années, j'ai pu lire dans l'œil de chaque malacologiste à qui j'étais en train d'expliquer que je travaillais sur les Turridae une expression de surprise, aussitôt suivie d'un regard de compassion. Il pensait presque à voix haute : « bon courage, mon gars ». A croire qu'il n'aurait pris ma place pour rien au monde. Je peux donc affirmer que ma complète ignorance de ce taxon "à s'en arracher les cheveux" et de la tâche qui m'attendait m'ont permis de m'y plonger sans *a priori* et de découvrir un groupe fascinant où chaque nouvelle mission sur le terrain apporte la quasi certitude de découvrir de nouvelles espèces.

Pour m'avoir offert l'opportunité de travailler sur ce sujet, merci tout d'abord à mes trois encadrants de thèse, Marie-Catherine Boisselier, Sarah Samadi et Philippe Bouchet. C'est l'avantage d'un co-encadrement, j'ai pu bénéficier de l'expertise de trois personnes très différentes mais totalement complémentaires. Soutien matériel, intellectuel et moral, rien n'a manqué. Merci surtout pour la confiance que vous m'avez témoignée, souvent en me laissant libre de mes choix et de mes décisions, mais sans jamais me laisser dériver trop longtemps.

Merci aux membres du jury, Guillaume Lecointre, Frédéric Austerlitz, Thierry Backeljau, Patrice David et Paul Hebert d'avoir accepté d'évaluer mon travail. Merci également à Jean-Yves Rasplus et à Jean François Silvain, pour les conseils qu'ils ont pu m'apporter en acceptant de faire partie de mon comité de thèse. Merci tout spécialement à Jean-François de m'avoir accueilli il y a quatre ans maintenant dans son laboratoire et d'avoir su me communiquer son enthousiasme et ainsi me conforter dans mes objectifs.

Merci à Hervé Le Guyader et à Simon Tillier pour leur accueil au sein de l'UMR7138 et de l'équipe « Espèce et spéciation ». Merci à toute l'équipe du Service de Systématique Moléculaire : Eric Pasquet, Annie Tillier, Josie Lambourdière, Céline Bonillo. Je pense qu'il va falloir que je fréquente d'autres laboratoires pour me rendre compte de la qualité de l'encadrement et du soutien dont les utilisateurs du SSM bénéficient !

Merci à toute l'équipe de malacologistes pour leur accueil et pour leur expertise indispensable sur laquelle je me suis appuyée au cours de ma thèse : Virginie Héros, Philippe Maestrati, Pierre Lozouet, Delphine Brabant, Barbara Buge, Jean-Michel Bichain, Joëlle Rameau, Rudo Von Cosel, Didier Molin, Olivier Gargomini, Benoît Fontaine, Danielle Placais.

Merci également à tous ceux dont j'ai croisé la route ici ou ailleurs, et qui m'ont chacun apporté leur aide : Ellen Strong (how to extract a Turridae from its shell?), Bertrand Richer de Forges, Anders Warén, Laurent Albenga, Regis Cleva, Jason Biggs, Marco Oliverio, Franck Dorkeld, Serge Gofas, Jacques Pelorce, André Hoareau, John Taylor, Alan Kohn, Chris Meyer, Tim Barrackough, Philippe Gaubert, Nicolas Vidal, Philippe Lopez. Pour leur collaboration sur différents projets, merci plus particulièrement à Yuri Kantor, Alexander Sysoev, Yves Terryn, Baldomero Olivera et Mandë Holford. Merci tout spécialement à Yuri pour son soutien et son aide précieuse.

Merci également aux membres de l'équipe morphométrie pour l'aide réellement indispensable qu'ils m'ont apportée : Michel Baylac, Vincent Debat, Raphaël Cornette, Allowen Evin. Merci à Raphaël Leblois et à Guillaume Achaz pour leurs explications des tenants et aboutissants de la statistique.

J'ai également pu réaliser dans des conditions idéales trois années de monitorat au sein du MNHN. Je remercie pour leur aide Thérèse Garestier, le CIES Jussieu, mais surtout Sophie Pons, Alain Pothet et tous les collègues moniteurs (Mickaël, Lucie, Juliette, Guillaume, Mathieu…)

Merci plus particulièrement à Magalie et Julien pour les âpres discussions sur la signification du monde et la place de la taxonomie dans le contexte économico-politico-éthico-scientifique actuel.

Merci enfin et surtout à tous ceux qui ont eu le courage de prendre le risque de trinquer avec un buveur de limonade (voir de lait-fraise). C'est aussi grâce à vous finalement que cette thèse a pu être menée à bout : Clara, Julien, Mag, Jean-François, Lucie, Marie-Lilith, Juliette, Aude, Jérôme, Philou, Charlène, Julien, Bob, Thomas, Guillaume, Lucio, Mandë, Toto C., Jean-Michel, Andrea, Olivier, Benoît, Lucas, Mathieu, Damien, Alan, Bruno, Stéphane,…

Je n'oublie pas les seuls qui me soutiennent depuis le début, et la route est longue ! Merci Jérôme, merci Papa, merci Maman.

Merci à Emilie. Ce qui est fait n'est plus à faire.

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

SOMMAIRE

Articles réalisés pendant la thèse ou en cours de réalisation
(* : inclus dans la thèse ; IF : Impact Factor 2007)

Publié :

*** Puillandre N.**, Samadi S., Boisselier M.-C., Sysoev A.V., Kantor, Y.I., Cruaud C., Couloux A., Bouchet P. 2008. Starting to disentangle the toxoglossan knot: molecular phylogeny of the "turrids" (Neogastropoda: Conoidea). Molecular Phylogenetics and Evolution 47: 1122-1134. IF = 3.994

In press :

*** Puillandre N.**, Baylac M., Boisselier M.-C., Cruaud C., Samadi S. An integrative approach to species delimitation in *Benthomangelia* (Mollusca: Conoidea). Biological Journal of the Linnean Society. IF = 2.368

* Kantor, Y.I., **Puillandre N.**, Olivera B., Bouchet P. Morphological proxies for taxonomic decision in turrids (Mollusca, Neogastropoda): a test of the value of shell and radula characters using molecular data. Zoological Science. IF = 1.125

* Holford M., **Puillandre N.**, Terryn Y., Olivera B., Cruaud C., Bouchet P. Evolution of the Toxoglossa venom apparatus as inferred by molecular phylogeny of the Terebridae. Molecular Biology and Evolution. IF = 6.438

Soumis :

Puillandre N., Samadi S., Boisselier M.-C., Bouchet P. New insights on the phylogeny of the Conoidea (Neogastropoda) using molecular data. Soumis à Nautilus. IF = 0.475

* Samadi S., **Puillandre N.**, McPherson E., Lambourdière J., Cruaud C., Couloux A., Boisselier M.-C. DNA barcode, type specimens and species delimitation in the genus *Eumunida*. Soumis à Biological Journal of the Linnean Society. IF = 2.368

En préparation :

*** Puillandre N.,** Couloux A., Samadi S. Identification of neogastropod eggs using DNA barcodes. Pour soumission à Molecular Ecology Resources. IF = 1.257

* **Puillandre N.**, Sysoev A.V., Olivera B., Couloux A., Bouchet P. Loss of planktotrophy, fragmentation and speciation: the deep-water gastropod *Bathytoma* (Gastropoda: Conoidea) in the West Pacific. Pour soumission à Zoologica Scripta. IF = 2.364

*** Puillandre N.**, Boisselier M.-C., Cruaud C., Samadi S. An integrative framework for species delimitation: a case studies in a hyperdiverse group of marine molluscs, the Turrinae (Gastropoda, Conoidea). Pour soumission à Systematic Biology. IF = 8.802

*** Puillandre N.**, Samadi S., Bouchet P. Diversity and distribution of the Turrinae (Gastropoda; Conoidea) in South-West Pacific. Pour soumission à Malacologia. IF = 2.115

Samadi S., Dayrat B., **Puillandre N.** Proceedings of the Second Barcode Conference, Taipei.

Puillandre N., Kantor Y. I., Bouchet, P. A new species in the genus *Gemmuloborsonia*.

Autres projets :

- Evolution of venom apparatus in Raphitominae. En collaboration avec Alexander Fedosov.
- Phylogeny, classification and evolution of the Conoidea. En collaboration avec Philippe Bouchet, Yuri Kantor et Alexander Sysoev et Chris Meyer.
- Statistical recognition of the Barcode gap. En collaboration avec Guillaume Achaz.

INTRODUCTION

Our generation is the first to fully comprehend the threat of the biodiversity crisis and the last with the opportunity to explore and document the species diversity of our planet.

Wheeler 2004

The family Turridae is more perplexing than any other molluscan family.

Hedley 1922

Dans le contexte actuel des changements climatiques globaux et des bouleversements écologiques que la Terre est en train de subir, la crise de la biodiversité est certainement le seul point sur lequel tous les acteurs politiques, sociologiques ou scientifiques sont d'accord. La liste des espèces menacées s'allonge en effet chaque jour. A ce problème, à propos duquel certains n'hésitent pas à parler de 6ème extinction (Thomas 2007), s'ajoute la pauvreté de nos connaissances sur la biodiversité : selon les estimations, seul 1 à 10 % de la biodiversité est actuellement connu (Blaxter 2004 ; Savolainen *et al*. 2005). Le taux d'extinction est si rapide, et nos connaissances si fragmentaires, que certaines espèces disparaissent certainement avant d'être découvertes (Pimm *et al*. 2006).

Une bonne connaissance taxonomique de la biodiversité est cruciale pour les différents acteurs impliqués dans sa préservation et sa conservation (Mace 2004 ; Raven 2004) : nous ne pouvons protéger que ce que nous connaissons. Certainement un des parents pauvres de la biologie, la taxonomie, et plus particulièrement l'α-taxonomie, est en passe de retrouver l'attrait scientifique et médiatique qu'elle avait perdu ces dernières décennies (Godfray 2002 ; Hogg & Hebert 2004 ; May 2004 ; Schander & Willassen 2005 ; Agnarsson & Kuntner 2007). Un retour indispensable, quand on connaît l'importance de la taxonomie pour les sciences biologiques (Agapow *et al*. 2004 ; Gotelli 2004) : autant essayer d'écrire dans une langue qui n'a ni dictionnaire, ni règle de grammaire.

Au rythme actuel de description d'espèces, plusieurs centaines ou milliers d'années seront nécessaires pour nommer l'ensemble de la biodiversité. Pour pallier ce problème, et également pour répondre à la demande croissante en matière de sauvegarde de la biodiversité, plusieurs projets récents ont souligné le besoin urgent de développer de nouvelles méthodes permettant l'accélération du rythme de description des espèces. Le succès grandissant des projets barcoding a permis ainsi de focaliser l'attention sur cette problématique, entraînant ainsi une nouvelle dynamique dans la communauté des systématiciens.

Cette thèse s'inscrit dans cette dynamique, avec l'objectif d'adapter les nouveaux outils, concepts et méthodes de la taxonomie à l'analyse de la diversité d'un groupe de gastéropodes marins, les Turridae. Ils constituent à la fois le groupe de mollusques marins le plus diversifié et l'un des moins connus. Les missions organisées depuis 30 ans dans le Pacifique par le MNHN (Muséum National d'Histoire Naturelle) et l'IRD (Institut de Recherche pour le Développement) ont révélé une diversité pour ce groupe largement supérieure à celle décrite dans la littérature (Bouchet *et al*. 2002). Cependant, la taxonomie traditionnelle ne parvient pas à proposer des hypothèses de délimitations robustes. Ce constat a été le point de départ de ma thèse : il était en effet nécessaire d'appliquer de nouveaux outils de la taxonomie, basés par exemple sur l'utilisation de l'ADN, pour accélérer le rythme de description d'espèces tout en proposant des hypothèses taxonomiques robustes.

A. Les Turridae : présentation et problématique

1. Présentation générale

Les Turridae constituent la famille de mollusques marins la plus diversifiée, avec plus de 4000 espèces décrites (10000 en incluant les fossiles) (Bouchet 1990 ; Tucker 2004). Ils sont présents dans toutes les mers du monde, des tropiques aux pôles, du littoral aux abysses, et constituent dans les zones bathyales et abyssales l'un des groupes de gastropodes dominants (Bouchet & Warén 1980). Ce sont des prédateurs de polychètes, d'autres mollusques et également de poissons. La plupart d'entre eux (voir chapitre 1) possède un appareil venimeux hautement spécialisé, composé d'une glande à venin et d'un système d'injection (harpon). Ce mécanisme est largement modifié dans certains groupes, et l'appareil venimeux est parfois absent (Miller 1970 & 1971 ; Kantor & Sysoev 1989 ; Kantor & Taylor 2002, Fedosov & Kantor 2008). Les Turridae sont inclus dans la super-famille des Conoidea (Caenogastropoda, Neogastropoda), qui comprend également les Cônes (genre *Conus*) et les Terebridae, dont la diversité est mieux décrite. Les mécanismes de prédations ont d'ailleurs été mieux étudiés pour ces groupes, et montrent une grande variété à la fois dans les stratégies adoptées et les proies chassées (Miller 1970 ; Letourneux 2004).

La plupart des efforts de la communauté des malacologistes, chercheurs ou amateurs, qui travaillent sur ce groupe, sont concentrés sur le genre *Conus*. La variabilité des formes et des couleurs observables chez les cônes, la facilité de se procurer des spécimens et la fascination provoquée par le danger que représente certaines espèces mortelles pour l'homme ont focalisé les attentions sur ce groupe, bien qu'il ne représente que moins de 7 % de la diversité spécifique connue (600 espèces décrites – Duda & Kohn 2005). La majeure partie de la diversité des Turridae est donc largement inexplorée : les relations phylogénétiques au sein des Conoidea sont pour la plupart encore inconnues, et la majorité de la diversité spécifique n'est pas encore décrite.

2. Classification

La grande variabilité des classifications proposées pour les Conoidea permet de se rendre compte des problèmes posés par ce groupe. En 1922, Hedley considère que la famille des Turridae « *is more perplexing than any other molluscan family* ». Les différences entre les classifications sont la conséquence de l'utilisation de différents caractères par les taxonomistes : ce qui est considéré comme un caractère essentiel au niveau familial ou sous-familial par un taxonomiste pourra être négligé par un autre (Chapitre 1 page 16).

J'ai résumé sur la figure 1 l'évolution entre trois classifications proposées pour ce groupe. De part leurs morphologies caractéristiques, le genre *Conus*, (spireconique), alors unique représentant de la famille des Conidae, et la famille des Terebridae (spire particulièrement allongée) ont été séparés des autres Conoidea. Par conséquent, tous les autres membres des Conoidea se sont retrouvés regroupés au sein de la famille des Turridae, sans autre raison que de n'être ni un *Conus*, ni un Terebridae. Sans remettre en cause cette organisation générale, la plupart des auteurs suivants ont tenté de définir des sous-familles au sein des Turridae (Powell 1942, 1966 ; Morrison 1965 ; McLean 1971 ; Kilburn 1983 à 1995). En revanche, Taylor *et al*. (1993) proposent une classification totalement remaniée du groupe, divisée en six familles, avec notamment la famille des Conidae élargie pour inclure cinq sous-familles placées précédemment au sein des Turridae. Les noms Conidae et Turridae ont donc été utilisés pour désigner des ensembles différents d'organismes aux cours des différentes

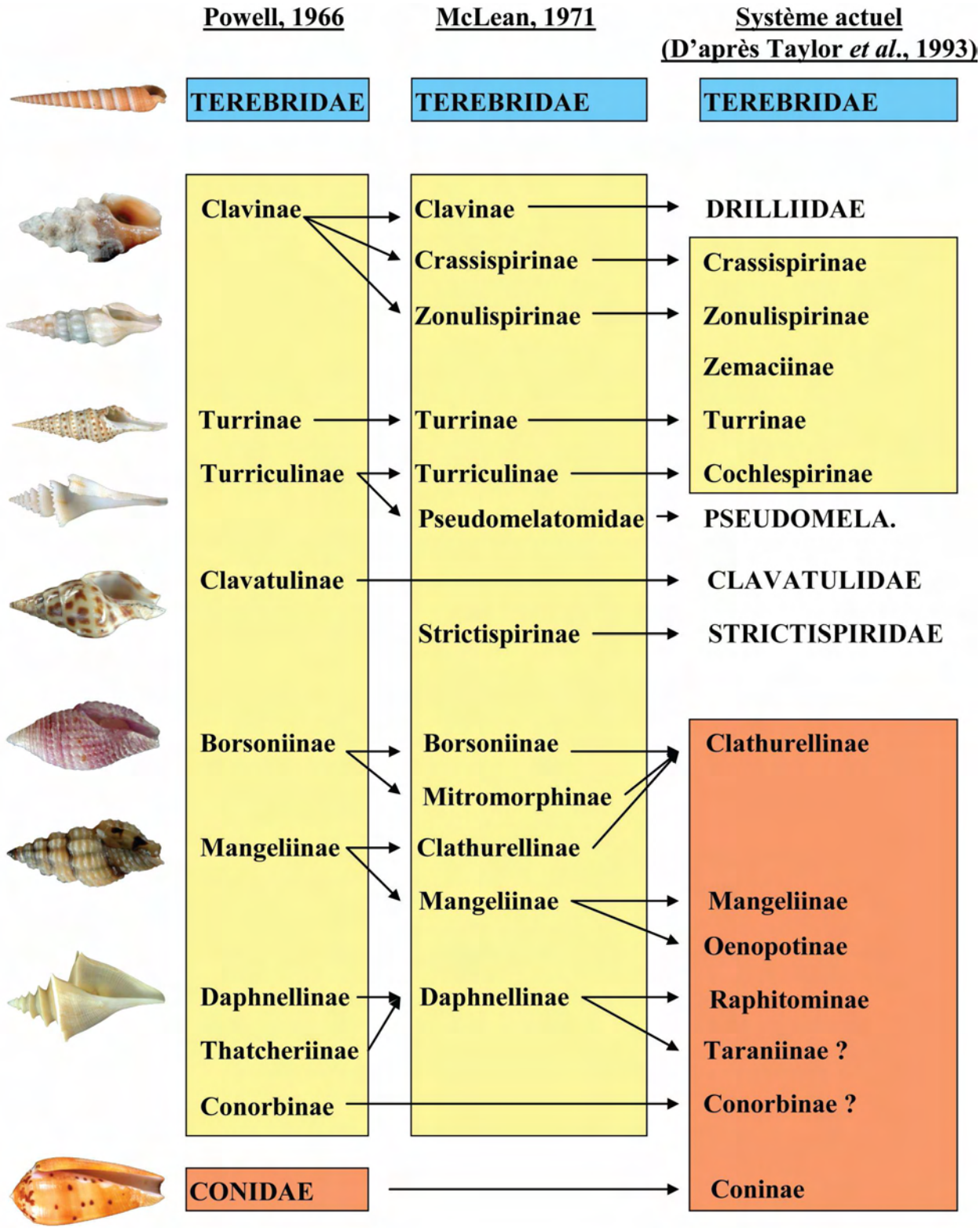


Figure 1 : Evolution de la classification des Conoidea, avec les principales correspondances entre les classifications de Powell (1966), McLean (1971) et Taylor *et al.* (1993). Les trois principales familles sont figurées en couleur (Terebridae en bleu, Turridae en jaune, Conidae en orange).

classifications. Le nom Conidae sera utilisé ici dans son acception la plus récente (Taylor *et al*. 1993), pour désigner non seulement le genre *Conus*, mais également plusieurs autres sous-familles (Figure 1). De la même façon, le nom Turridae sera utilisé tel que définit par Taylor *et al*. (1993), c'est-à-dire n'incluant que cinq sous-familles. Dans ce cas, il sera suivi de « sensu stricto » (Turridae *s.s.*). Généralement, c'est l'acception plus ancienne du nom Turridae qui sera également utilisée : « Turridae » désignera alors l'ensemble des Conoidea sauf le genre *Conus* et les Terebridae. Dans la communauté des malacologistes, les termes Turridae et Conidae sont encore couramment employés dans leurs anciennes acceptions (avant Taylor *et al*. 1993).

3. A-taxonomie

L'α-taxonomie du groupe est également sujette à de nombreuses controverses parmi les spécialistes, pour principalement cinq raisons (Figure 2) :
1) comme chez la plupart des mollusques, les espèces de Turridae sont principalement décrites sur des caractères de la coquille (Encadré 1). Les diagnoses, et plus particulièrement dans les descriptions les plus anciennes, sont parfois brèves, et applicables ainsi à différents taxa. De plus, il peut exister dans la littérature plusieurs noms mis en synonymie, ce qui ajoute à la confusion.
2) il est souvent difficile d'identifier des caractères morphologiques discrets qui permettraient de proposer des hypothèses consensuelles de délimitation d'espèces. En effet, la variabilité morphologique de la coquille peut correspondre à un continuum et il est donc souvent très difficile de placer les limites entre espèces le long de ce continuum.
3) par conséquent, les interprétations que font les taxonomistes de cette variabilité morphologique peuvent varier d'un taxonomiste à l'autre. La distinction entre les taxonomistes « splitters », qui ont tendance à séparer les espèces même sur la base de faibles différences morphologiques, et les « lumpers », qui au contraire interprètent plus facilement cette même variabilité comme intraspécifique, est particulièrement vraie pour les Turridae. Les « lumpers » auront donc tendance à regrouper des individus dans une même espèce, alors qu'ils auraient été séparés en plusieurs espèces par les « splitters ». Les études réalisées par différentes personnes sont donc parfois difficilement comparables.
4) la diversité des Turridae est encore largement inconnue, et chaque mission sur le terrain rapporte à la surface son lot de formes nouvelles qui amènent souvent les taxonomistes à modifier les hypothèses de délimitation d'espèces qu'ils avaient proposées précédemment. Par exemple, ce qui pouvait apparaître comme deux morphes très distincts, et qui auront donc été décrits comme deux espèces différentes, peuvent simplement correspondre à deux

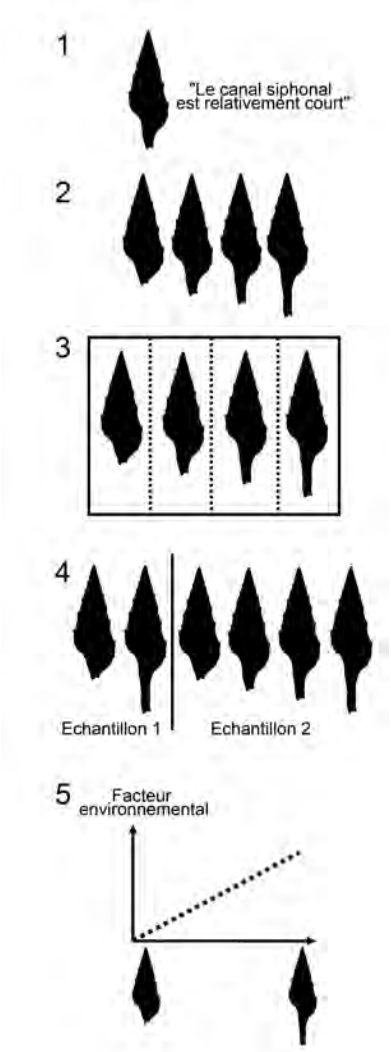


Figure 2 : modèle caricatural d'une espèce décrite brièvement (1), et qui présente sur un caractère (longueur du canal siphonal) une variabilité continue (2). Cette variabilité sera interprétée différemment par un taxonomiste lumper (trait plein) ou splitter (pointillés) (3). Cette interprétation changera selon l'échantillonage disponible (4). De plus, cette variabilité peut-être liée à l'environnement (5).

formes extrêmes d'une même espèce. La découverte de formes intermédiaires peut permettre de détecter ce genre de problème et modifie donc les délimitations d'espèces.
5) le déterminisme des caractères morphologiques utilisés dans la plupart des descriptions d'espèces n'est pas connu. Les différences au niveau de la morphologie de la coquille peuvent ainsi refléter non pas des différences entre espèces, mais des différences liées à l'environnement (plasticité phénotypique ou adaptation locale - Hollander *et al*. 2006 ; Brookes & Rochette, 2007).

Ces difficultés avaient d'ailleurs déjà été identifiées par les premiers taxonomistes ayant étudiés les Turridae, que ce soit pour délimiter des espèces ou pour regrouper ces espèces en genres, sous-familles et familles :

> *"In no other group of mollusks is it so difficult to make a satisfactory classification as in the Pleurotomidae* (= Turridae). *The forms are exceedingly numerous, and known in many species to be very variable in their characters, whilst the material for the recognition of most of those described is generally scanty. Of the figured species, a very large proportion were described from single or few specimens, and most cabinets, however large, do not possess shells which can be certainly identified with these: then there is an unusually large proportion (amounting to hundreds) of unfigured species, the recognition of which is simply impossible" (Tryon, 1884).*

ENCADRE 1

Morphologie de la coquille et intérêt en taxonomie

Le groupe des Turridae est caractérisé par un attribut particulier de la coquille, la présence d'un canal siphonal sur la partie supérieure de l'ouverture, comme illustrée sur la Figure 3. Cette coquille représente une forme classique pour un Turridae (*Turris babylonia*, Linnaeus 1758). Elle se compose d'une protoconque, ou coquille larvaire, formée selon les taxa d'un ou plusieurs tours. On distingue deux types de protoconques, paucispirales (généralement moins de 2 tours) et multispirales (généralement plus de 2 tours) (Shuto 1974). La différence dans le nombre de tours de la protoconque est liée au temps que la larve passe à se nourrir dans la colonne d'eau : une coquille larvaire multispirale indique que la larve est planctotrophe, une coquille paucispirale qu'elle est non planctortrophe. La planctotrophie est généralement considérée comme un indicateur des capacités de dispersion du spécimen : une larve planctotrophe qui reste dans la colonne d'eau disperse plus (Article 6 page 128).
La sculpture de cette protoconque est utilisée en taxonomie (Powell 1942 ; Bouchet & Waren 1980). Certains types de sculpture apparaissent en effet diagnostic d'un groupe (treillissée chez les Raphitominae par exemple). Cependant, d'autres semblent partagés par plusieurs groupes (carénée chez les Clathurellinae et chez certains Crassispirinae par exemple). Après la protoconque vient la téléoconque, ou coquille adulte. Les principaux caractères utilisés concernent la forme générale, la taille de la coquille, la forme de l'ouverture et l'ornementation (côtes, cordons, sutures entre les tours, épines,…). La présence d'un opercule, ainsi que sa forme, sont également des caractères utilisés en taxonomie. En revanche, les couleurs de la coquille, des spots et des flammes, souvent plastiques, sont en général peu utilisées. Bien que ne faisant pas partie de la coquille, la radula est également riche en caractères morphologiques. Plusieurs types existent, différents selon le nombre de rangées de dents et leurs formes, et couramment utilisés à tous les niveaux taxonomiques, de la famille à l'espèce (Powell 1966 ; Kohn com. pers.). J'y reviendrais à la fin du chapitre 3.

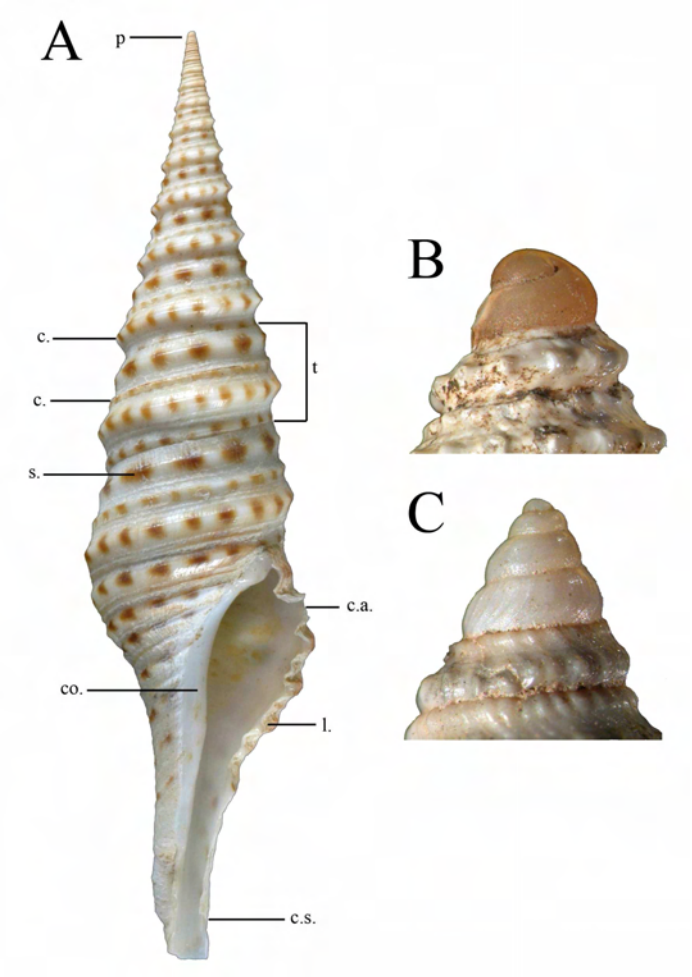


Figure 3 : A. Coquille de *Turris babylonia*. c. = cordon ; c.a. = canal anal ; c.s. = canal siphonal ; co. = columelle ; l. = labre ; p. = protoconque ; s. = spot ; t. = 1 tour de spire. B. Protoconque paucispirale de *Bathytoma* (actuel). C. Protoconque multispirale de *Bathytoma* (fossile).

B. Objectifs et plan de la thèse

Caractérisé par leur grande diversité et par les difficultés que rencontre la taxonomie traditionnelle pour définir une classification et une α-taxonomie robustes, les Turridae constituent donc un défi intéressant dans le contexte actuel de crise de la biodiversité. L'analyse taxonomique des Turridae proposée dans cette thèse inclut une analyse phylogénétique, afin de définir les relations de parentés entre les organismes et de proposer ainsi une nouvelle classification, et une analyse α-taxonomique, qui a pour but de regrouper ces organismes en espèces (Wheeler & Meier 2000). Ces deux aspects seront explorés en s'appuyant principalement sur l'analyse de l'ADN.

1. Phylogénie et α-taxonomie

La phylogénie moléculaire a connu un essor considérable dans les 20 dernières années. Elle a notamment permis de mettre à jour de nombreuses relations phylogénétiques inédites entre les grands groupes d'organismes, modifiant profondément la classification du monde vivant (Lecointre & Le Guyader 2001). L'ADN présente notamment l'avantage d'être un caractère héritable, ce qui n'est pas nécessairement le cas pour les caractères morphologiques. Il sera utilisé pour inférer les relations phylogénétiques au sein du groupe étudié, les Turridae. Toutefois, d'après la dernière classification des Conoidea (Taylor *et al*. 1993), les Turridae ne forment pas un groupe monophylétique. Pour tester cette hypothèse, l'ensemble des Conoidea (Turridae, *Conus* et Terebridae) sera analysé afin de proposer une nouvelle classification du groupe.

L'analyse phylogénétique permettra également de définir des groupes monophylétiques au sein desquels les limites entre espèces pourront être analysées. Restreindre l'analyse α-taxonomique à des groupes monophylétiques permet de comparer des espèces qui ont une histoire commune récente et ainsi de mieux prendre en compte les contraintes phylogénétiques, pour par exemple une meilleure gestion de l'homoplasie (Chapitre 3). Comme pour la phylogénie, l'ADN est un caractère de plus en plus utilisé en α-taxonomie. Hebert *et al*. (2003a) proposent notamment d'utiliser un fragment d'ADN comme un code-barre moléculaire (barcode ADN – Encadré 2) permettant l'identification quasi automatique des taxa terminaux déjà reconnus et éventuellement la mise en évidence de nouveaux taxa à décrire et nommer. C'est cette seconde facette de l'approche barcoding qui sera utilisée pour proposer des hypothèses de délimitation d'espèces chez les Turridae. Cependant, il sera nécessaire de confronter ces hypothèses à d'autres caractères, notamment morphologiques, dans une approche intégrative.

2. Plan de la thèse

Dans le premier chapitre, j'expose les résultats obtenus concernant la phylogénie et la classification des Conoidea. Après avoir rappelé les problèmes posés par les caractères traditionnellement utilisés dans les classifications précédentes, je présente la phylogénie moléculaire des Conoidea obtenue (Article 1, publié dans Molecular Phylogenetics and Evolution). Cette première phylogénie a été complétée grâce à un jeu de données élargi qui permet à la fois de résoudre de nouvelles relations phylogénétiques au sein de ce groupe, mais également de proposer une nouvelle classification. Cette seconde analyse est actuellement en cours et devrait faire l'objet prochainement d'une publication. De plus, je présente dans ce chapitre la phylogénie d'un groupe au sein des Conoidea, les Terebridae. Cette phylogénie a

ENCADRE 2

DNA barcodes et barcoding

Il existe dans la littérature liée au barcoding une large confusion sur l'utilisation des termes (Köhler 2007 ; Vogler & Monaghan 2007). Pour certains, le barcoding n'est qu'un outil d'identification (Chapitre 2 page 57). D'autres auteurs regroupent sous le terme barcoding toutes les approches récentes de taxonomie qui font appel à un séquençage massif du gène COI et à son application en α-taxonomie. Cela peut notamment inclure :

- les approches des DNA-taxonomie, où l'ADN est utilisé pour délimiter des espèces quand les autres caractères sont difficiles à analyser (Monaghan *et al.* 2006). C'est par exemple le cas des nématodes, dont la diversité insoupçonnée a été révélée après plusieurs études « environnementales » (Blaxter *et al.* 2005, De Ley *et al.* 2005).
- L'α-taxonomie intégrative, encore une fois présente dans les sciences systématiques depuis quelques décennies, et qui s'appuie sur une combinaison de différents caractères (morphologique, moléculaire, écologique,…), notamment le gène COI utilisé dans les approches barcoding, pour proposer des hypothèses de délimitation d'espèces.

Décider quelles pratiques doivent être regroupées sous le terme barcoding constitue en fait le principal débat entre pro et anti-barcoding (Wheeler 2005 ; Will *et al.* 2005), la majorité d'entre eux étant d'accord sur les limites liées à l'utilisation d'un seul marqueur et sur l'intérêt d'utiliser une analyse intégrative. La position officielle du CBOL (Consortium for the Barcode Of Life) consiste à restreindre le terme barcoding à un outil d'identification de spécimen : dès qu'il s'agit de (re)définir des limites entre espèces, on ne parle plus de barcoding mais d'α-taxonomie (intégrative, moléculaire,…).

Pour compliquer un peu le problème, plusieurs termes existent pour désigner à la fois le fragment du gène COI utilisé et/ou les applications que l'on peut en faire : barcode, barcoding, barcode ADN, code-barres ADN, barcode of life.

J'utiliserai le mot barcode pour désigner le fragment du gène COI choisi et le mot barcoding pour en désigner les applications. Le mot barcoding pourra être utilisé pour désigner son utilisation en tant qu'outil d'identification taxonomique (on pourrait parler de barcoding *sensu stricto*) ou pour désigner son utilisation dans un contexte intégratif de délimitation d'espèces (barcoding *sensu lato*). Le contexte d'utilisation sera à chaque fois précisé dans le texte.

notamment permis d'analyser l'évolution du caractère présence/absence de l'appareil venimeux au sein du groupe (Article 2, soumis à Molecular Biology and Evolution).

Dans le second chapitre, je développe tout d'abord le contexte conceptuel et méthodologique dans lequel s'inscrit l'approche α-taxonomique suivie. Le point essentiel réside dans la distinction entre le concept d'espèce utilisé (Lineage Species Concept), développé notamment par De Queiroz (1998) et Samadi & Barberousse (2006), et les propriétés des espèces (interfécondité, ressemblance, monophylie des lignées) qui sont autant de critères de délimitation d'espèces. Ces critères peuvent être appliqués sur différents types de caractères, notamment moléculaires. Je détaille ensuite le principe du barcoding, où un fragment d'ADN (en l'occurrence une partie du gène COI) peut être utilisé pour diverses applications : (i) expertise taxonomique, qui peut conduire à modifier des hypothèses taxonomiques déjà existantes et (ii) mise en place de nouvelles hypothèses de délimitation

d’espèces. La première application sera illustrée à la fin de ce chapitre (Article 3, soumis au Biological journal of the Linnean Society ; Article 4, en préparation, pour soumission à Molecular Ecology Resources).

Dans le troisième et dernier chapitre, je développe la seconde application du barcoding. Il peut en effet être utilisé pour proposer de nouvelles hypothèses de délimitation d’espèces qui sont ensuite testées à l’aide d’autres caractères dans une approche de α-taxonomie intégrative. Après une présentation des méthodes d’échantillonnage utilisées, je détaille quatre exemples au sein des Turridae : la sous-famille des Turrinae et les genres *Bathytoma*, *Benthomangelia*, *Xenuroturris* (Article 5, en préparation, pour soumission à Systematic Biology ; Article 6, en préparation, pour soumission à Zoologica Scripta ; Article 7, accepté dans le Biological journal of the Linnean Society ; Article 8, accepté avec modifications mineures dans Zoological Science). J’insisterai notamment sur les méthodes qui ont été mises au point et appliquées pour proposer des hypothèses taxonomiques robustes au sein de ces groupes. Je dresse en conclusion le bilan des résultats obtenus et je développe notamment ce qu’ils nous permettent de savoir sur la diversité et la distribution des Turridae (Article 9, en préparation, pour soumission à Malacologia).

CHAPITRE I

PHYLOGÉNIE MOLÉCULAIRE

The family is commonly called a "taxonomic nightmare": if you meet Turrid taxonomists and they appear sane, they are probably faking it.

http://www.manandmollusc.net/

A. Phylogénie et classification

1. Problèmes liés aux caractères morphologiques

Les différentes classifications proposées pour la superfamille des Conoidea ont connu principalement deux évolutions majeures. Premièrement, les caractères utilisés ne sont pas les mêmes d'une classification à une autre : les premières classifications sont basées sur la présence/absence d'un opercule et sur le type de radula (toxoglossate/non-toxoglossate) chez Thiele (1929) par exemple, puis sur la morphologie de la protoconque et de la radula chez Powell (1942) et McLean (1971), et enfin sur les caractères anatomiques chez Taylor *et al*. (1993). Deuxièmement, leur complexité augmente, depuis les premières classifications en trois groupes (*Conus*, Terebridae et Turridae), jusqu'à la plus récente, proposée par Taylor *et al*. (1993) et qui définit six familles et plus de 10 sous-familles au sein des Conoidea (Figure 1 page 8). Cette évolution est particulièrement bien illustrée par le changement de statut du genre *Conus*. Dans les premières classifications, il est le seul représentant de la famille des Conidae, tous les autres Conoidea (les Terebridae exceptés) étant regroupés dans la famille des Turridae (Figure 4). Dans la classification de Taylor *et al*. (1993), plusieurs groupes inclus auparavant dans la famille des Turridae sont placés dans la famille des Conidae avec le genre *Conus*.

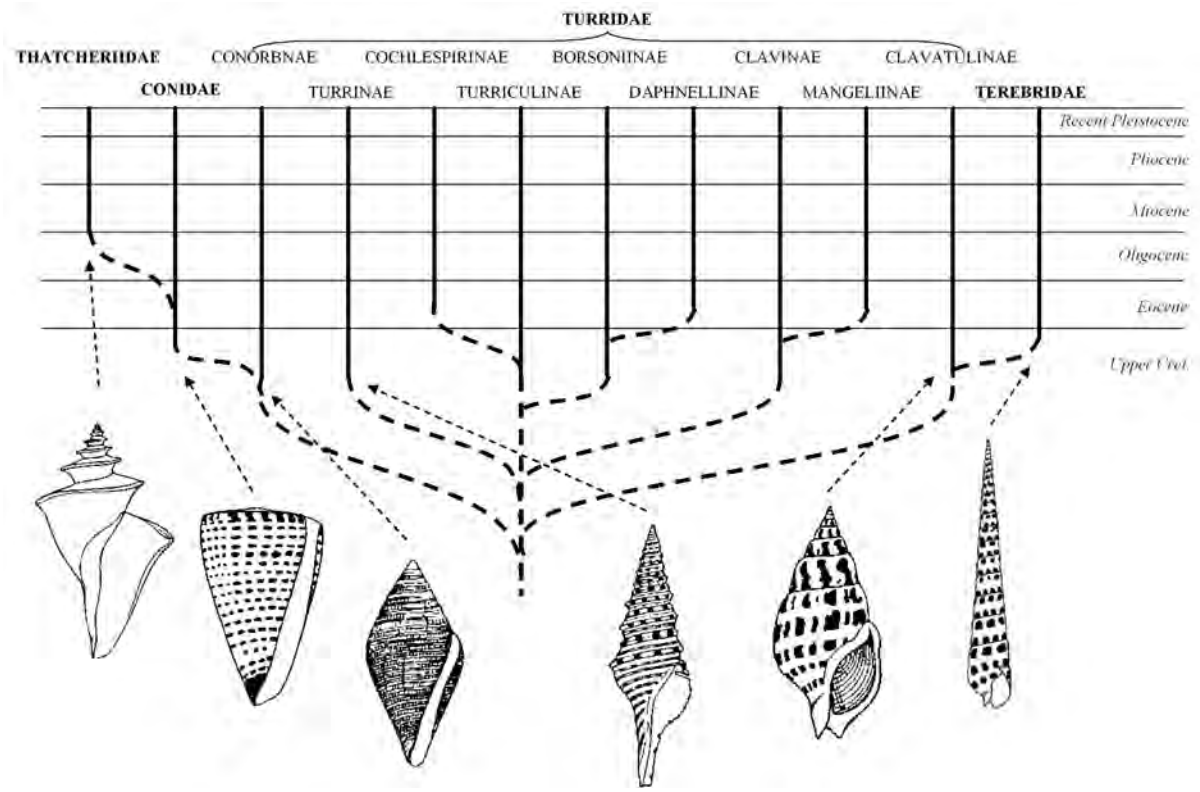


Figure 4 : Phylogénie des Toxoglossa (= Conoidea), d'après Powell, 1942. Les traits verticaux pleins représentent les distributions connues et les traits pointillés les possibles relations entre les groupes. Espèces illustrées, de gauche à droite : *Thatcheria mirabilis*, *Conus literatus*, *Conorbis dormitor*, *Turris babilonius*, *Pusionella nifat* et *Terebra subulata*. L'auteur précise que l'espèce *P. nifat* est un Turridae qui ressemble à un Terebridae : même si l'intuition n'est pas réellement justifiée, il semble que les Terebridae soient inclus dans les Turridae *s.s.* d'après les résultats moléculaires. Il est intéressant de constater également que la famille des Turridae telle que représentée sur cette figure n'est pas monophylétique.

Ces deux évolutions sont étroitement liées : l'utilisation de caractères nouveaux a non seulement enrichi les classifications passées mais aussi montré que certains états de caractères précédemment utilisés étaient homoplasiques. Les classifications anciennes ne permettaient donc pas de retracer les relations phylogénétiques au sein des Conoidea. C'est notamment le cas de l'absence de l'opercule (voir Tryon 1884 pour une discussion de ce caractère), ou encore de certains types de protoconques. L'analyse de la phylogénie moléculaire permettra notamment d'étayer ces hypothèses sur l'invalidité des caractères « présence/absence de l'opercule » et « type de protoconque ». Par exemple, la sous-famille des Oenopotinae (Bogdanov 1987), caractérisée par la présence d'un opercule, a été séparée de la sous-famille des Mangeliinae (absence d'opercule). Mais les analyses moléculaires indiquent que la sous-famille des Oenopotinae est incluse dans la sous-famille des Mangeliinae (Figure 5 page 23). De plus, des groupes caractérisés par une morphologie de la coquille qui divergeait de façon frappante ont parfois été élevé à un rang taxonomique sous-familial ou familial. Ainsi, à cause de sa morphologie caractéristique (Figure 4), Powell (1966) isole le genre *Thatcheria* de la sous-famille des Raphitominae pour en faire une famille à part (Thatcheriidae). En fait, les analyses moléculaires indiquent que le genre *Thatcheria* appartient bien à la sous-famille des Raphitominae et qu'il n'est pas nécéssaire d'en faire une autre famille ou sous-famille (voir le clade « Raphitominae » de la figure 5 page 23).

2. Apport des caractères moléculaires

La principale difficulté à laquelle se sont confrontés les auteurs qui ont tenté d'établir une classification des Conoidea est donc l'identification de caractères homologues. La plupart de ces classifications ne sont basées que sur un nombre limité de caractères, et la présence parmi eux de caractères homoplasiques a conduit à des classifications qui s'avèrent maintenant erronées. De plus, face à l'importante diversité morphologique du groupe, certains auteurs ont limité géographiquement leur analyse (par exemple à l'Afrique du Sud et au Mozambique pour Kilburn 1983-1995), sous-estimant ainsi la diversité du groupe et aboutissant à une classification incomplète.

L'objectif de ce travail est donc d'utiliser un grand nombre de caractères pour proposer une nouvelle classification des Conoidea. Plusieurs gènes ont donc été séquencés (un mitochondrial, le gène COI, et trois nucléaires, les gènes 28S, 18S et H3), de façon à éviter les problèmes liés à l'utilisation d'un petit nombre de caractères. De plus, les méthodes phylogénétiques utilisées (de vraisemblance et bayésienne) permettent de gérer les problèmes liés à l'homoplasie. Enfin, même si la plupart des échantillons ont été collectés lors de missions organisées par le Muséum National d'Histoire Naturelle (MNHN) et l'Institut de Recherche pour le Développement (IRD) dans le Pacifique Sud-Ouest, plusieurs échantillons récoltés dans d'autres régions géographiques (Nouvelle-Zélande, Norvège, Afrique) ont permis de baser l'analyse sur un échantillonnage représentatif de la diversité connue (Annexe). Cent quatorze spécimens ont été séquencés, représentant 57 genres, huit sous-familles et cinq familles de Conoidea.

Ces résultats ont été publiés dans la revue Molecular Phylogenetics and Evolution (Article 1).

Cette étude a notamment permis de définir plusieurs clades regroupant chacun plusieurs genres. Certains de ces clades correspondent à des groupes (familles ou sous-familles) proposés notamment par Taylor *et al*. (1993) : c'est le cas des Drilliidae, des Terebridae ou des Raphitominae. Les autres sous-familles semblent polyphylétiques, avec par

exemple la sous-famille des Cochlespirinae représentée dans quatre clades différents, ou celle des Clathurellinae dans sept. En revanche, la plupart des relations entre ces clades ne sont pas soutenues, même si certaines d'entre elles permettent de retrouver d'autres groupes déjà proposés. C'est par exemple le cas du clade B, faiblement soutenu mais identique à la famille des Conidae telle que l'avait définie Taylor *et al*. (1993).

De nouvelles hypothèses ont également été proposées. Au niveau générique, la polyphylie de certains genres est soutenue (*Leucosyrinx* et *Comitas*), et l'appartenance de certains genres à des sous-familles ou familles est modifiée (le genre *Otitoma* passe ainsi de la sous-famille des Mangeliinae (Conidae) à la sous-famille des Crassispirinae (Turridae *s.s.*)). A un niveau taxonomique supérieur, un clade correspondant à la sous-famille des Crassispirinae (Turridae) est plus proche de la famille des Drilliidae que des autres membres de sa famille.

Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 1122–1134
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Starting to unravel the toxoglossan knot: Molecular phylogeny of the "turrids" (Neogastropoda: Conoidea)

N. Puillandre [a,b,*], S. Samadi [a], M.-C. Boisselier [a], A.V. Sysoev [c], Y.I. Kantor [d], C. Cruaud [e], A. Couloux [e], P. Bouchet [b]

[a] UMR 7138, Systématique, adaptation, évolution, UPMC, IRD, MNHN, CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France
[b] USM 602/UMS CNRS 2700, Muséum National d'Histoire Naturelle, Département Systématique et Évolution, 55 rue Buffon, 75231 Paris Cedex 05, France
[c] Zoological Museum of Moscow State University, Bolshaya Nikitskaya str. 6, Moscow 103009, Russia
[d] A. N. Severtzov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninski prospect 33, Moscow 119071, Russia
[e] GENOSCOPE, Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry Cedex, France

Received 13 September 2007; revised 7 November 2007; accepted 15 November 2007
Available online 3 January 2008

Abstract

The superfamily Conoidea is one of the most speciose groups of marine mollusks, with estimates of about 340 recent valid genera and subgenera, and 4000 named living species. Previous classifications were based on shell and anatomical characters, and clades and phylogenetic relationships are far from well assessed. Based on a dataset of ca. 100 terminal taxa belonging to 57 genera, information provided by fragments of one mitochondrial (COI) and three nuclear (28S, 18S and H3) genes is used to infer the first molecular phylogeny of this group. Analyses are performed on each gene independently as well as for a data matrix where all genes are concatenated, using Maximum Likelihood, Maximum Parsimony and Bayesian approaches. Several well-supported clades are defined and are only partly identifiable to currently recognized families and subfamilies. The nested sampling used in our study allows a discussion of the classification at various taxonomical levels, and several genera, subfamilies and families are found polyphyletic.

Keywords: 18S rRNA; 28S rRNA; Classification; COI gene; Conoidea; Conidae; H3 gene; Molecular phylogeny; Toxoglossa; Turridae; West Pacific

1. Introduction

The superfamily Conoidea (= Toxoglossa) includes small to medium (3–50 mm on average) sized species of marine snails that are specialist predators on annelids, other mollusks, and even fishes, and occupy all marine habitats from the tropics to the poles, from shallow to deep water, and from hard to soft substrates. This is the most diverse groups of marine mollusks, with almost 700 recent and fossil nominal genera and 10,000 described species (Bouchet, 1990), and current estimates of about 340 recent valid genera and subgenera (Taylor et al., 1993) and 4000 named living species (Tucker, 2004). *Conus* alone includes over 500 valid species, making it the most speciose genus of marine animals (Kohn, 1990; Duda and Kohn, 2005). The monophyly of the Conoidea, characterized by a venom apparatus, is not questioned (Taylor et al., 1993), but subdivisions within Conoidea, and relationships between them are controversial, mostly because the extensive morphological and anatomical variation encountered is itself not well understood. In this context, molecular data can bring new characters, allowing to root the classification of Conoidea in an evolutionary perspective using a phylogenetic analysis.

* Corresponding author. Address: UMR 7138, Systématique, adaptation, évolution, UPMC, IRD, MNHN, CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.
E-mail address: puillandre@mnhn.fr (N. Puillandre).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.11.007

During most of the 19th and 20th century, classifications (e.g., Fischer, 1887; Cossmann, 1896; Hedley, 1922; Thiele, 1929; Wenz, 1938–1944) were based on characters of the shell and of the radula, and Powell (1942, 1966) later gave emphasis on characters of the protoconch (larval shell). All these authors traditionally recognized three families of recent Conoidea: (i) Conidae, only containing the genus *Conus*, (ii) Terebridae containing species with acuminate shells without a siphonal canal, and (iii) Turridae, including the rest, i.e., the vast majority of the group. Turridae was considered by Hedley (1922) to be "more perplexing than any other molluscan family". Powell's (1942, 1966) subdivision of the Turridae in nine subfamilies (see Table 1) was the basis for turrid classifications in the latter half of the 20th century. Subsequent authors diverged on the number of subfamilies they recognized, mostly splitting one subfamily into several: working mainly on East Pacific faunas, McLean (1971) thus recognized 15 subfamilies of Turridae; Kilburn (various papers from 1983 to 1995) recognized eight subfamilies in the South African fauna; while in their monograph of European deep-sea turrids, Bouchet and Warén (1980) criticized the use of "more or less randomly selected shell characters" and did not use subfamilies at all. Other shell- and radula-based classifications, mostly regional, include Morrison (1965), Shimek and Kohn (1981) and Chang (1995, 2001). A turning point in toxoglossate classification was the work of Taylor et al. (1993) who extensively used anatomical characters, in addition to radulae. Their entirely novel classification recognized six families (Conidae, Turridae, Terebridae, Drilliidae, Pseudomelatomidae and Strictispiridae), the salient point being that Conidae was by then enlarged beyond Coninae (*Conus*) to include five subfamilies previously placed in Turridae, and the newly restricted Turridae included a further five subfamilies. Bouchet and Rocroi's (2005) recent review of gastropod classification essentially kept Taylor's classification with updates based mainly on Rosenberg (1998) and Medinskaya and Sysoev (2003): Clavatulinae was raised to the family level; Taraninae was synonymized with Raphitominae; and the novel subfamily Zemaciinae was accepted in the Turridae. Thereafter, we use "Turridae *s.l.*" to designate all Conoidea except *Conus* and Terebridae (i.e., Turridae sensu Powell (1966) and most 20th century authors) and "Turridae *s.s.*" to designate the family as restricted by Taylor et al. (1993), while "Conidae" designates the expanded family after Taylor et al. (1993).

Since Taylor et al. (1993), several anatomical studies have highlighted the high level of homoplasy of the characters of the shell and the radula (e.g., Kantor and Taylor, 1994; Kantor et al., 1997; Taylor, 1994), but although *Conus* itself has been subjected to intensive molecular studies (e.g., Duda and Kohn, 2005), the phylogeny of the broader Conoidea has not yet been addressed based on molecular characters. The present paper is thus the first molecular phylogeny, based on fragments of one mitochondrial and three nuclear genes, of the crown clade of the Caenogastropoda. It provides insights at several taxonomic levels (generic, subfamilial and familial) and the adequacy of previous classifications is thus re-evaluated.

2. Materials and methods

2.1. Taxon sampling

Because of the instability of the taxonomy of the group, currently accepted synonymies cannot be taken for certain and must be re-evaluated. Ideally, a molecular sampling should thus include several representatives of all the nominal family group-names, including their type genera, whether they are currently regarded as taxonomically valid or not. In practice, this goal is difficult or impossible to reach because (a) a number of nominal (sub)families are based on fossil type genera (e.g., Borsoniinae, Conorbinae), and (b) a number of type genera are restricted in distribution and/or live in deep water and are difficult to obtain alive (e.g., Pseudomelatomidae, Thatcheriidae). To overcome these difficulties, our taxon sampling includes several genera for as many as possible of the subfamilies proposed in the literature (see detail in Table 1). Of the 114 specimens sequenced, few were replicates and the taxon sampling represents about 100 species in 57 valid genera.

2.2. Materials

The bulk of the material was obtained during expeditions carried out in the tropical western Pacific during research expeditions by the Muséum National d'Histoire Naturelle (MNHN) and the Institut de Recherche pour le Développement (IRD) (see Table 2). Living specimens were anesthetized using $MgCl_2$, a piece of tissue was cut from the head-foot, and fixed in 95% ethanol. This dataset was supplemented by specimens collected in West Africa by Serge Gofas in the mid 1980s. Shells were kept intact for identification. Identifications were carried to genus level using the classically admitted shell-based genus definitions, but, given the chaotic state of turrid systematics, no attempt was made to identify our material to species level; a number of species, especially from deep water, probably represent new species. Even so, eight specimens could not confidently be attributed to a genus and are denoted thereafter "*cf.* Genus". Conversely, specimens of Terebridae and *Conus* were identified to species level. A specimen of a species of *Nassaria* and a specimen of a species of *Cancellopollia*, both in the neogastropod family Buccinidae, closely related to Conoidea (Harasewych et al., 1997; Colgan et al., 2007), were used as outgroups. *Littorina littorea*, belonging in the non-neogastropod family Littorinidae, was used as a third outgroup, with sequences taken from GenBank (GenBank Accession Nos: AJ622946.1, Q279985.1, AJ488712.1 and DQ093507.1). Outgroups were chosen to form a non-monophyletic group, as recommended by Darlu and Tassy (1993). All vouchers are kept in MNHN.

Table 1
Evolution of Conoidea classification

	Powell, 1966	McLean, 1971	Taylor et al., 1993	Actual system*	
TURRIDAE *s.l.*	**Clavinae** *Ceritoturris* *Crassispira* *Horaiclavus ?* *Inquisitor* = *Ptychobela* *Microdrillia* *Splendrillia*	**Clavinae** *Splendrillia*	**DRILLIIDAE** *Clavus* *Horaiclavus* *Splendrillia*	**DRILLIIDAE** *Clavus* *Conopleura:* Tippet and Tucker, 1995 *Splendrillia*	TURRIDAE *s.s.*
		Crassispirinae *Inquisitor*	**Crassispirinae** *Ceritoturris* *Crassispira* *Funa* *Inquisitor* *Iwaoa* *Ptychobela* *Turridrupa*	**Crassispirinae** *Anacithara:* Kilburn, 2004 *Ceritoturris* *Crassispira* *Funa* *Horaiclavus:* Kantor, pers. com. *Inquisitor* *Iwaoa* *Ptychobela*	
		Zonulispirinae	**Zonulispirinae**	**Zonulispirinae**	
				Zemaciinae	
	Turrinae *Gemmula* *Lophiotoma* *Turridrupa* *Turris*	**Turrinae** *Gemmula* *Lophiotoma*	**Turrinae** *Gemmula* *Lophiotoma* *Turris*	**Turrinae** *Gemmula* *Gemmuloborsonia:* Sysoev and Bouchet, 1996 *Lophiotoma* *Turris* *Turridrupa:* Kantor et al., 1997	
	Turriculinae *Cochlespira* *Comitas* *Iwaoa* *Leucosyrinx*	**Turriculinae**	**Cochlespirinae** *Cochlespira* *Comitas* *Leucosyrinx*	**Cochlespirinae** *Cochlespira* *Comitas* *Leucosyrinx*	
		Pseudomelatomidae	**PSEUDOMELATOMIDAE**	**PSEUDOMELATOMIDAE**	
		Strictispirinae	**STRICTISPIRIDAE**	**STRICTISPIRIDAE**	
	TEREBRIDAE		**TEREBRIDAE**	**TEREBRIDAE** *Cinguloterebra:* Terryn, 2007 *Terebra*	
	Clavatulinae *Clavatula* *Perrona* *Pusionella*		**Clavatulinae** *Clavatula* *Perrona* *Pusionella*	**CLAVATULIDAE** *Clavatula* *Perrona* *Pusionella*	
	Borsoniinae *Bathytoma* *Borsonia* *Mitromorpha*	**Borsoniinae** *Borsonia*	**Clathurellinae** Borsoniid: *Borsonia* *Typhlomangelia* Mitromorphid: *Anarithma* *Mitromorpha*	**Clathurellinae** Borsoniid: *Borsonia* *Typhlomangelia* Mitromorphid: *Anarithma* *Mitromorpha*	CONIDAE
		Mitromorphinae *Mitromorpha*			
	Mangeliinae *Anacithara* *Benthomangelia* *Conopleura* *Etrema* *Eucithara* *Guraleus* *Leiocithara* *Lienardia* *Macteola* *Mangelia*	**Clathurellinae**	Bathytomid: *Bathytoma* Clathurellid: *Etrema ?* *Nannodiella* Tomopleurid: *Heteroturris* *Microdrillia* *Tomopleura*	Bathytomid: *Bathytoma* Clathurellid: *Etrema ?* *Nannodiella* Tomopleurid: *Heteroturris* *Microdrillia* *Tomopleura*	
TURRIDAE *s.l.*		**Mangeliinae**	**Mangeliinae** *Benthomangelia* *Eucithara* *Guraleus* *Leiocithara* *Lienardia* *Macteola* *Mangelia*	**Mangeliinae** *Benthomangelia* *Eucithara* *Guraleus* *Leiocithara* *Lienardia* *Macteola* *Mangelia* *Otitoma:* Kilburn, 2004 *Toxicochlespira:* Sysoev and Kantor, 1990	
			Oenopotinae	**Oenopotinae**	
	Daphnellinae *Daphnella* *Eucyclotoma* *Gymnobela* *Kermia* *Pleurotomella* *Raphitoma* *Rimosodaphnella* *Veprecula*	**Daphnellinae** *Daphnella* *Kermia* *Philbertia* *Veprecula*	**Daphnellinae** *Daphnella* *Rimosodaphnella* *Eucyclotoma* *Teretiopsis* *Gymnobela* *Thatcheria* *Kermia* *Tritonoturris* *Pleurotomella* *Veprecula* *Raphitoma*	**Raphitominae** *Daphnella* *Eucyclotoma* *Glyphostomoides:* Shuto, 1983 *Gymnobela* *Kermia* *Pleurotomella* *Raphitoma* *Rimosodaphnella* *Teretiopsis* *Thatcheria* *Tritonoturris* *Veprecula*	
	Thatcheriinae		**Taraniinae ?**		
	Conorbinae *Benthofascis*		**Conorbinae ?** *Benthofascis*	**Conorbinae** *Benthofascis*	
	CONIDAE *Conus*		**Coninae** *Conus*	**Coninae** *Conus*	

History of conoidean classification and position of the genera included in the present dataset in the classifications of Powell (1966), McLean (1971) and Taylor et al. (1993). Subfamilies are in bold, families in bold and capital. *Recent modifications proposed since the classification of Taylor et al. (1993) (details given for each genus), resulting in the actual system used as a basis for our discussion. (See above-mentioned references for further information).

Table 2
Specimens of Conoidea used in this study

ID	Cruise	Station ID	Coordinates, depth	Genus (or species) identification	COI	28S	18S	H3	Clades	
17700[a]	BOA 1	CP2462	16°37.5′S, 167°57.4′E, 618–641 m	*Bathytoma* Harris and Burrows, 1891	x	x	x	x	20	B
17701[a]	BOA 1	CP2432	14°59.7′S, 166°55.0′E, 630–705 m	*Leucosyrinx* Dall, 1889	x	x	x	x	9	A
17702	BOA 1	CP2432	14°59.7′S, 166°55.0′E, 630–705 m	*Leucosyrinx* Dall, 1889	x	x	x	x	9	A
17754[a]	Panglao 2004	R42	9°37.1′N, 123°52.6′E, 8–22 m	*Turris* Röding, 1798	x	x	x	x	5	A
17755[a]	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Crassispira* Swainson, 1840	x	x	x	x	2, C	A
17829	Angola	Ilha de Luanda	8°78′S, 13.23′E, 40–60 m	*Clavatula* Lamarck, 1801	x				22	A
17830	Angola	Cacuaco	10°51′S, 14°23′E, 5–10 m	*Pusionella* Gray, 1847	x				22	A
17831	Angola	Cacuaco	10°51′S, 14°23′E, 5–10 m	*Pusionella* Gray, 1847	x				22	A
17832	Cameroun	Victoria	3°54′N, 9°08′E, 34–37 m	*Pusionella* Gray, 1847	x				22	A
17833	Angola	Moçâmedes	15°14′S, 12°29′E, 50 m	*Perrona* Schumacher, 1817	x				22	A
17834	Gabon	Port-Gentil	1°17′S, 11°53′E	*Pusionella* Gray, 1847	x				22	A
17835[a]	BOA 1	CP2462	16°37.5′S, 167°57.4′E, 618–641 m	*Benthomangelia* Thiele, 1925	x	x	x	x	17	B
17836	BOA 1	CP2462	16°37.5′S, 167°57.4′E, 618–641 m	*Rimosodaphnella* Cossmann, 1915	x	x	x	x	10	B
17837	EBISCO	DW2547	21°06′S, 158°36′E, 356–438 m	*Inquisitor* Hedley, 1918	x	x	x	x	2, C	A
17838	EBISCO	DW2533	22°18′S, 159°28′E, 360–370 m	*Gemmula* Weinkauff, 1875	x	x	x	x	5	A
17839[a]	EBISCO	CP2557	21°07′S, 158°30′E, 800–923 m	*Borsonia* Bellardi, 1839	x	x	x	x	16	B
17840[a]	EBISCO	DW2631	21°03′S, 160°44′E, 372–404 m	*Horaiclavus* Oyama, 1954	x	x	x	x	7	A
17841	EBISCO	CP2648	21°32′S, 162°30′E, 750–458 m	*Gymnobela* Verrill, 1884	x	x	x	x	10	B
17842[a]	EBISCO	DW2553	21°03′S, 158°36′E, 352–370 m	*Cochlespira* Conrad, 1865	x	x	x		8	A
17843	EBISCO	DW2522	22°46′S, 159°21′E, 310–318 m	*Funa* Kilburn, 1988	x	x	x	x	2, C	A
17844	EBISCO	CP2645	20°58′S, 160°58′E, 641–652 m	*Gymnobela* Verrill, 1884	x	x	x	x	10	B
17845	EBISCO	CP2651	21°29′S, 162°36′E, 883–957 m	*Teretiopsis* Kantor and Sysoev, 1989	x	x	x	x	10	B
17846[a]	EBISCO	CP2600	19°38′S, 158°46′E, 603–630 m	*Leucosyrinx* Dall, 1889	x	x	x	x	3, C	A
17847[a]	EBISCO	DW2617	20°06′S, 160°22′E, 427–505 m	*Splendrillia* Hedley, 1922	x	x	x	x	1, C	A
17848	EBISCO	DW2625	20°05′S, 160°19′E, 627–741 m	*Pleurotomella* Verrill, 1873	x	x	x	x	10	B
17849[a]	EBISCO	DW2619	20°06′S, 160°23′E, 490–550 m	*cf. Gemmuloborsonia* Shuto, 1989	x	x	x	x		A
17850	EBISCO	DW2607	19°33′S, 158°40′E, 400–413 m	*Turridrupa* Hedley, 1922	x	x	x	x	5	A
17851	EBISCO	DW2625	20°05′S, 160°19′E, 627–741 m	*Inquisitor* Hedley, 1918	x	x	x	x	2, C	A
17852	EBISCO	DW2625	20°05′S, 160°19′E, 627–741 m	*Gemmula* Weinkauff, 1875	x	x	x	x	5	A
17853[a]	EBISCO	DW2629	21°06′S, 160°46′E, 569–583 m	*Heteroturris* Powell, 1967	x	x	x	x	18	B
17855[a]	Norfolk 2	DW2155	22°52′S, 167°13′E, 453–455 m	*Benthofascis* Iredale, 1936		x	x	x		B
17857	EBISCO	CP2551	21°06′S, 158°35′E, 637–650 m	*Bathytoma* Harris and Burrows, 1891	x	x	x	x	20	B
17858	Panglao 2004	S12	9°29.4′N, 123°56.0′E, 6–8 m	*Clavus* Monfort, 1810	x	x	x	x	1, C	A
17859	Panglao 2004	S12	9°29.4′N, 123°56.0′E, 6–8 m	*Turridrupa* Hedley, 1922	x	x	x	x	5	A
17860	Panglao 2004	R44	9°33.3′N, 123°43.9′E, 2 m	*Lophiotoma* Casey, 1904	x	x	x	x	5	B
17861	Panglao 2004	B14	9°38.5′N, 123°49.2′E, 2–4 m	*Kermia* Oliver, 1915	x	x	x	x	10	B
17862	Panglao 2004	T10	9°33.4′N, 123°49.6′E, 117–124 m	*Gemmula* Weinkauff, 1875	x	x	x	x	5	A
17863	Panglao 2004	B16	9°37.6′N, 123°47.3′E, 20 m	*Macteola* Hedley, 1918	x	x	x	x	11	B
17864	Panglao 2004	S18	9°35.7′N, 123°44.4′E, 0–2 m	*cf. Guraleus* Hedley, 1918	x	x	x	x	11	B
17865	Panglao 2004	P2	9°39′N, 123°44′E, 400 m	*Bathytoma* Harris and Burrows, 1891	x	x	x	x	20	B
17866[a]	Panglao 2004	S19	9°42.1′N, 123°51.4′E, 3–4 m	*Mangelia* Risso, 1826	x	x	x	x	11	B
17867	Panglao 2004	B19	9°29.4′N, 123°56.0′E, 17 m	*Borsonia* Bellardi, 1839	x	x	x	x	16	B
17868	Panglao 2004	B19	9°29.4′N, 123°56.0′E, 17 m	*Anacithara* Hedley, 1922	x	x	x	x	7	A
17869	Panglao 2004	S21	9°41.7′N, 123°50.9′E, 4–12 m	*Etrema* Hedley, 1918	x	x	x	x	12	B
17870	Panglao 2004	S25	9°41.5′N, 123°51.0′E, 21 m	*Otitoma* Jousseaume, 1898	x	x	x	x	2, C	A
17871	Panglao 2004	S26	9°41.5′N, 123°51.0′E, 21 m	*Kermia* Oliver, 1915	x	x	x	x	10	B
17872	Panglao 2004	S26	9°41.5′N, 123°51.0′E, 21 m	*Macteola* Hedley, 1918	x	x	x	x	11	B
17873	Panglao 2004	T26	9°43.3′N, 123°48.8′E, 123–135 m	*Guraleus* Hedley, 1918	x	x	x	x	11	B
17874	Panglao 2004	T26	9°43.3′N, 123°48.8′E, 123–135 m	*Guraleus* Hedley, 1918	x	x	x	x	11	B
17875[a]	Panglao 2004	T26	9°43.3′N, 123°48.8′E, 123–135 m	*Tomopleura* Casey, 1924	x	x	x	x	14	B
17876	Panglao 2004	B21	9°37.2′N, 123°46.4′E, 20–21 m	*Lienardia* Jousseaume, 1928	x	x	x	x	12	B
17877[a]	Panglao 2004	B21	9°37.2′N, 123°46.4′E, 20–21 m	*Mitromorpha* Carpenter, 1865	x	x	x	x	13	B
17878	Panglao 2004	B25	9°29.4′N, 123°56.1′E, 16 m	*Kermia* Oliver, 1915	x	x	x	x	10	B
17879	Panglao 2004	T32	9°36.4′N, 123°53.8′E, 60–62 m	*Inquisitor* Hedley, 1918	x	x	x	x	2, C	A
17880	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Kermia* Oliver, 1915	x	x	x	x	10	B
17881	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Daphnella* Hinds, 1844	x	x	x	x	10	B
17882[a]	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Raphitoma* Bellardi, 1848	x	x	x	x	10	B
17883	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Veprecula* Melvill, 1917	x	x	x	x	10	B
17884	Panglao 2004	L46	9°30.9′N, 123°41.2′E, 90–110 m	*Leiocithara* Hedley, 1922	x	x	x	x	11	B
17885	Panglao 2004	T36	9°29.3′N, 123°51.5′E, 95–128 m	*Ceritoturris* Dall, 1924	x	x	x	x	7	A
17886	Panglao 2004	T36	9°29.3′N, 123°51.5′E, 95–128 m	*Splendrillia* Hedley, 1922	x	x	x	x	1, C	A
17887	Panglao 2004	T36	9°29.3′N, 123°51.5′E, 95–128 m	*Microdrillia* Casey, 1903	x	x	x	x	18	B
17888	Panglao 2004	T36	9°29.3′N, 123°51.5′E, 95–128 m	*Ceritoturris* Dall, 1924	x	x	x	x	7	A

(*continued on next page*)

Table 2 (*continued*)

ID	Cruise	Station ID	Coordinates, depth	Genus (or species) identification	COI	28S	18S	H3	Clades	
17889	Panglao 2004	T41	9°29.7′N, 123°50.2′E, 110–112 m	*Conopleura* Hinds, 1844	x	x	x	x	1, C	A
17890	Panglao 2004	L49	9°36.5′N, 123°45.3′E, 90 m	*Raphitoma* Bellardi, 1848	x	x	x	x	10	B
17891	Panglao 2004	T39	9°30.1′N, 123°50.4′E, 100–138 m	*cf. Tritonoturris* Dall, 1924	x	x	x	x	10	B
17892	Panglao 2004	T39	9°30.1′N, 123°50.4′E, 100–138 m	*cf. Glyphostomoides* Shuto, 1983	x	x	x	x	10	B
17893	Panglao 2004	T41	9°29.7′N, 123°50.2′E, 110–112 m	*cf. Mitromorpha* Carpenter, 1865	x	x	x	x	13	B
17894	Panglao 2004	B7	9°35.9′N, 123°51.8′E, 4–30m	*Lienardia* Jousseaume, 1928	x	x	x	x	12	B
17895	Panglao 2004	D5	9°33.6′N, 123°43.5′E, 0–3 m	*Inquisitor* Hedley, 1918	x	x	x	x	2, C	A
17896	Panglao 2004	D5	9°33.6′N, 123°43.5′E, 0–3 m	*Eucithara* Fischer, 1883	x	x	x	x	11	B
17897[a]	Panglao 2004	B8	9°37.1′N, 123°46.1′E, 3 m	*Lienardia* Jousseaume, 1928	x	x	x	x	12	B
17898	Panglao 2004	B8	9°37.1′N, 123°46.1′E, 3 m	*Mitromorpha* Carpenter, 1865	x	x	x	x	13	B
17899	Panglao 2004	B8	9°37.1′N, 123°46.1′E, 3 m	*Eucithara* Fischer, 1883	x	x	x	x	11	B
17900	Panglao 2004	B8	9°37.1′N, 123°46.1′E, 3 m	*Eucithara* Fischer, 1883	x	x	x	x	11	B
17901	Panglao 2004	S5	9°37.1′N, 123°46.1′E, 2–4 m	*Anarithma* Iredale, 1916	x	x	x	x	13	B
17902	Panglao 2004	S6	9°38.5′N, 123°49.2′E, 1–4 m	*Clavus* Monfort, 1810	x	x	x	x	1, C	A
17903	Panglao 2004	S12	9°29.4′N, 123°56.0′E, 6–8 m	*Eucyclotoma* Boettger, 1895	x	x	x	x	10	B
17904	Panglao 2004	T9	9°33.5 N, 123°49.5′E, 97–120 m	*cf. Nannodiella* Dall, 1919	x	x	x	x	12	B
17905	Panglao 2005	CP2348	9°29.6′N, 123°52.5′E, 196–216 m	*Otitoma* Jousseaume, 1898	x	x	x	x	2, C	A
17906	Panglao 2005	CP2349	9°31.6′N, 123°55.7′E, 219–240 m	*Ptychobela* Thiele, 1925	x	x	x	x	2, C	A
17907	Panglao 2005	CP2349	9°31.6′N, 123°55.7′E, 219–240 m	*Gemmula* Weinkauff, 1875	x	x	x	x	5	A
17908	Panglao 2005	CP2332	9°38.8′N, 123°45.9′E, 396–418 m	*Iwaoa* Kuroda, 1953	x	x	x	x	7	A
17909	Panglao 2005	CP2343	9°27.4′N, 123°49.4′E, 273–356 m	*Cinguloterebra cf. fujitai* Kuroda and Habe, 1952	x	x	x	x	6	A
17910	Panglao 2005	CP2349	9°31.6′N, 123°55.7′E, 219–240 m	*Tomopleura* Casey, 1924	x	x	x	x	14	B
17911	Panglao 2005	CP2333	9°38.2′N, 123°43.5′E, 584–596 m	*cf. Heteroturris* Powell, 1967	x	x	x	x	18	B
17912	Panglao 2005	CP2377	8°40.6′N, 123°20.3′E, 85–88 m	*Conus praecellens* Adams, 1854	x	x	x	x	19	B
17913[a]	Panglao 2005	CP2377	8°40.6′N, 123°20.3′E, 85–88 m	*Conus sulcatus* Hwass in Bruguière, 1792	x	x	x	x	19	B
17914	Panglao 2005	CP2380	8°41.3′N, 123°17.8′E, 150–163 m	*Conus sulcatus* Hwass in Bruguière, 1792	x	x	x	x	21	B
17915	Panglao 2005	CP2381	8°43.3′N, 123°19.0′E, 259–280 m	*Toxicochlespira* Sysoev and Kantor, 1990	x	x	x	x	17	B
17916[a]	Panglao 2005	CP2385	8°51.0′N, 123°10.0′E, 982–989 m	*Comitas* Finlay, 1926	x	x	x	x	4, C	A
17917	Panglao 2005	CP2393	9°30.1′N, 123°41.6′E, 356–396 m	*Terebra polygyrata* Deshayes, 1859	x	x	x	x	6	A
17918	Panglao 2005	CP2388	9°26.9′N, 123°34.5′E, 762–786 m	*Comitas* Finlay, 1926	x	x	x	x	4, C	A
17919	Panglao 2005	CP2340	9°29.4′N, 123°44.4′E, 271–318 m	*Cochlespira* Conrad, 1865	x	x	x		8	A
17920	Panglao 2005	CP2340	9°29.4′N, 123°44.4′E, 271–318 m	*Cochlespira* Conrad, 1865	x	x	x		8	A
17921[a]	Panglao 2005	CP2340	9°29.4′N, 123°44.4′E, 271–318 m	*Conus orbignyi* Kilburn, 1975	x	x	x	x	21	B
17922	Panglao 2005	DW2400	9°32.5′N, 123°41.8′E, 111–115 m	*Conus wakayamaensis* Kuroda, 1956	x	x	x	x	21	B
17923	Panglao 2005	CP2395	9°36.2′N, 123°43.8′E, 382–434 m	*Cinguloterebra cf. fenestrata* Hinds, 1844	x	x	x	x	6	A
17924	Salomon 2	CP2184	8°16.9′S, 159°59.7′E, 464–523 m	*Thatcheria* Angas, 1877	x	x	x	x	10	B
17925	Salomon 2	CP2227	6°37.2′S, 156°12.7′E, 508–522 m	*Toxicochlespira* Sysoev and Kantor, 1990	x	x	x	x	17	B
17926[a]	Salomon 2	CP2269	7°45.1′S, 156°56.3′E, 768–890 m	*Borsonia* Bellardi, 1839	x	x	x	x	15	B
17927	Salomon 2	CP2260	8°03.5′S, 156°54.5′E, 399–427 m	*Daphnella* Hinds, 1844	x	x	x	x	10	B
17928	Salomon 2	CP2216	7°45.3′S, 157°39.4′E, 930–977 m	*Comitas* Finlay, 1926	x	x	x	x	3, C	A
17929	Salomon 2	CP2186	8°17.0′S, 160°00.0′E, 487–541 m	*Bathytoma* Harris and Burrows, 1891	x	x	x	x	20	B
17930	Salomon 2	CP2269	7°45.1′S, 156°56.3′E, 768–890 m	*Benthomangelia* Thiele, 1925	x	x	x	x	17	B
17931	Salomon 2	CP2269	7°45.1′S, 156°56.3′E, 768–890 m	*cf. Typhlomangelia* Sars, 1878	x	x	x	x	18	B
17932	Salomon 2	CP2197	8°24.4′S, 159°22.5′E, 897–1057 m	*Borsonia* Bellardi, 1839	x	x	x	x	15	B
17933	Salomon 2	CP2228	6°34.7′S, 156°10.5′E, 609–625 m	*Comitas* Finlay, 1926	x	x	x	x	3, C	A
17934	Salomon 2	CP2176	9°09.4′S, 158°59.2′E, 600–875 m	*Borsonia* Bellardi, 1839	x	x	x	x	16	B
17935	Salomon 2	CP2187	8°17.5′S, 159°59.8′E, 482–604 m	*Inquisitor* Hedley, 1918	x	x	x	x	2, C	A
17936	Santo 2006	LD28	15°35.4′S, 166°58.7′E, 3–8 m	*Conus generalis* Linne, 1758	x	x	x	x	19	B
17937	Santo 2006	NR52	15°35.6′S, 167°01.9′E, 15 m	*Conus gauguini* Richard and Salvat, 1973	x	x	x	x	19	B
17938[a]	Santo 2006	LD28	15°35.4′S, 166°58.7′E, 3–8 m	*Terebra textilis* Hinds, 1844	x	x	x	x	6	A
17939	Santo 2006	AT87	15°32.1′S, 167°16.1′E, 235–271 m	*Conus consors* Sowerby, 1833	x	x	x	x	19	B
17854	Norfolk 2	DW2034	23°41′S, 167°41′E, 485–505 m	*Nassaria*, Buccinidae	x	x	x	x		
17856	Norfolk 2	DW2081	25°54′S, 168°22′E, 500–505 m	*Cancellopollia*, Buccinidae	x	x	x	x		
GenBank				*Littorina*, Littorinidae	x	x	x	x		

Identification number (ID) corresponding to MNHN catalogue number, cruise and station of collection, with the coordinates and the depth, are given for each specimen. Specimens are identified at genus level, except *Conus* and Terebridae which are identified at species level. A cross indicates that the specimen was successfully sequenced for the gene. Allocation to clades A, B, C and 1–22, as defined by the molecular analysis, is given for each taxon.

[a] This specimen has been chosen to illustrate the clade to which it belongs in Fig. 1.

2.3. Sequencing

DNA was extracted from a piece of foot, using 6100 Nucleic Acid Prepstation system (Applied Biosystem) or DNeasy® 96 Tissue kit (Qiagen) for smaller specimens. A fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene was amplified using the universal primers LCO1490 and HCO2198 developed by Folmer et al. (1994). Three nuclear gene fragments were also analyzed: (i) 900 bp of the rDNA 28S gene, involving D1, D2 and D3 domains (Hassouna et al., 1984), using the primers C1 and D3 (Jovelin and Justine, 2001); (ii) 328 bp of the H3 gene using the primers H3aF and H3aR (Okusu et al., 2003); (iii) 1770 bp of the 18S gene using three pairs of primers: 1F and 5R, 3F and Bi, A2 and 9R (Giribet et al., 1996; Okusu et al., 2003). All PCR reactions were performed in 25 μl, containing 3 ng of DNA, 1× reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 μM of each primer, 5% DMSO and 1.5 U of Q-Bio Taq (Qbiogene) for all genes. Amplifications consisted of an initial denaturation step at 94 °C for 4′, followed by 30 cycles of denaturation at 94 °C for 30″, annealing at 52 °C for 28S gene and first and third fragment of 18S gene, and 53 °C for H3 gene and second fragment of 18S gene for 40″ and extension at 72 °C for 1′. The final extension was at 72 °C for 10′. COI gene amplifications followed description of Hebert et al. (2003). PCR products were purified using ExonucleaseI and Phosphatase and sequenced using BigDyeTerminator V3.1 kit (Applied biosystem) and the ABI3730XL sequencer. Because of the length of the 28S PCR product, two internal primers (D2 and C2′, Dayrat et al., 2001) were used for sequencing, in addition of primers used for PCR. All genes were sequenced for both directions to confirm accuracy of each sequence. The overlap of the three fragments of 18S gene made it possible to check for consistency. Sequences were deposited in GenBank (GenBank Accession Nos: EU015417–EU015858).

2.4. Phylogenetic analyses

COI and H3 genes were manually aligned whereas 28S and 18S genes were automatically aligned using ClustalW multiple alignment implemented in BioEdit version 7.0.5.3 (Hall, 1999). The accuracy of automatic alignments was confirmed by eye. Hyper-variable regions of 28S gene and 3′ extremity of 18S gene were excluded from further analyses due to ambiguities in the alignments. For protein coding genes (COI and H3), saturation according to codon position was tested by plotting genetic distances against patristic distances calculated from a Maximum Parsimony (MP) tree with a heuristic search option, 10 random taxon-addition (RA) and tree-bisection and reconnection (TBR) branch-swapping using PAUP 4.0b10 (Swofford, 2002).

Nucleotide substitution models were selected for each gene separately and for each combined dataset using the program Modeltest (Posada and Crandall, 2001), in conjunction with PAUP 4.0b10 (Swofford, 2002). Best models and parameters as estimated by the AIC criterion were used for Maximum Likelihood (ML) analyses; only the model was fixed for Bayesian Analyses (BA). Analyses were conducted using three different approaches. A heuristic MP search was executed with 100 RA, TBR branch-swapping, all sites equally weighted and indels treated as fifth states, using PAUP 4.0b10 (Swofford, 2002). ML heuristic search was conducted with 100 replicates with TBR branch-swapping using PhyML 2.4.4 (Guindon and Gascuel, 2003). Robustness of the nodes was assessed using non-parametric bootstrapping (Felsenstein, 1985) with 100 bootstraps replicates for MP analysis and 1000 for ML analysis, TBR branch-swapping and 100 RA replicates. BA consisted of six Markov chains (8,000,000 generations each with a sampling frequency of one tree each hundred generations) run in two parallel analyses using Mr. Bayes (Huelsenbeck et al., 2001). When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in. For the treatment of combined data using BA, the data were separated into four unlinked partitions corresponding to the four genes analyzed, each following the best fitting model of substitution estimated for each gene.

2.5. Turning the phylogeny into a classification

There are currently 41 available family-group names in the Conoidea, of which 19 are considered valid at family or subfamily ranks (Bouchet and Rocroi, 2005). In a nomenclatural perspective, only the occurrence of the type genus of a family-group name in a clade allows an unequivocal application of this name to that clade. For example, the clade containing the genus *Raphitoma* can unambiguously carry the name Raphitominae. However, many type genera are not represented in our taxon sampling and many of our molecular clades do not include a type genus. In such cases, we have relied on the traditional allocation of non-type genera to a subfamily to link clade and name. For example, a clade containing three genera classically classified in the family Drilliidae (Taylor et al., 1993; Tippet and Tucker, 1995) can carry the name Drilliidae, even though *Drillia* itself is not part of our taxon sampling. However, this approach does not lead to an unequivocal application of names when genera (or subfamilies) as traditionally construed prove to be non-monophyletic; in that case, only the type species (or the type genus) is the legitimate bearer of the name.

3. Results

For COI and H3 genes, 658 and 328 bp were sequenced, respectively, and no indels were found. After the alignment, we obtained a fragment of 933 and 1729 bp in length for the 28S and 18S genes, respectively. Sequencing of specimens belonging to genera *Clavatula*, *Pusionella* and *Perrona* was successful only for the COI gene: the prolonged conservation in the museum collections (more than 20 years) may have altered the quality of the DNA. Only

one specimen (17855) failed to sequence for COI gene, and three others (17842, 17919 and 17920, genus *Cochlespira*) for H3 gene. No bias was detected in base composition. The saturation analyses for the two protein coding genes revealed that the COI gene was highly saturated on the third position of codon, thus we used only the first and second positions in the phylogenetic analyses. Best model and parameters estimated for each gene and genes combinations are shown in Table 3. Independent analyses of each of the four genes provided very poorly resolved trees, with few well-supported clades (Table 4).

The only incongruencies found between the independent gene analyses corresponded to poorly supported nodes. The most supported incongruency concerned relationships between three specimens attributed to the genus *Bathytoma* (17700, 17865 and 17857). In the ML analysis of H3 gene 17700 was the sister-group of 17865 and 17857 whereas in the ML analysis of the 18S gene 17865 was the sister-group of 17700 and 17857. These two nodes were supported by bootstrap value of, respectively, 61 and 67, values weaker than the bootstrap value allowing the recognition of a supported clade (e.g., Hillis and Bull, 1993; Soltis and Soltis, 2003).

Table 3
Models of evolution and parameters estimated using AIC implemented in Modeltest for each gene separately and each combined dataset

Dataset	Model	Base frequencies	Substitution rates	I	G
COI	GTR+I+G	$\pi_A = 0.1922$ $\pi_C = 0.245$ $\pi_G = 0.2215$ $\pi_T = 0.3413$	r (A-C) = 0.8578 r (A-G) = 5.3343 r (A-T) = 0.3918 r (C-G) = 0.9449 r (C-T) = 35.0926	0.6915	0.6794
28S	GTR+I+G	$\pi_A = 0.1563$ $\pi_C = 0.3383$ $\pi_G = 0.3502$ $\pi_T = 0.1551$	r (A-C) = 0.7256 r (A-G) = 1.8046 r (A-T) = 1.5931 r (C-G) = 0.4122 r (C-T) = 7.8933	0.5957	0.6338
18S	TrNef+I+G		r (A-C) = 1 r (A-G) = 3.0918 r (A-T) = 1 r (C-G) = 1 r (C-T) = 9.2099	0.8620	0.5928
H3	GTR+I+G	$\pi_A = 0.2063$ $\pi_C = 0.3261$ $\pi_G = 0.3113$ $\pi_T = 0.1563$	r (A-C) = 1.4455 r (A-G) = 3.2261 r (A-T) = 2.663 r (C-G) = 0.9033 r (C-T) = 8.6701	0.6233	0.9671
CD1	GTR+I+G	$\pi_A = 0.2154$ $\pi_C = 0.2761$ $\pi_G = 0.2803$ π_T =0.2282	r (A-C) = 1 r (A-G) = 2.8258 r (A-T) = 1 r (C-G) = 1 r (C-T) = 10.8424	0.7230	0.4565
CD2	GTR+I+G	$\pi_A = 0.2062$ $\pi_C = 0.2772$ $\pi_G = 0.2874$ $\pi_T = 0.2292$	r (A-C) = 1.3887 r (A-G) = 3.1175 r (A-T) = 1.1091 r (C-G) = 0.986 r (C-T) = 11.5743	0.7192	0.4490

I, Proportion of invariable sites; G, gamma rate distribution; CD, combined dataset.

Since no incongruency was revealed among the single gene analyses, we constructed two combined datasets comprising the data of the 4 gene fragments resulting in a sequence of 3428 bp length. For both combined datasets we excluded the taxa attributed to *Clavatula*, *Pusionella* and *Perrona* for which only the COI gene was successfully obtained. For the first combined dataset (CD1) we also excluded the specimens 17855, 17842, 17919 and 17920, not sequenced for all genes, to avoid potential perturbation of phylogenetic reconstruction by missing data (Wiens, 1998). Thus, the CD1 included 104 ingroups and the second combined dataset (CD2) included 108 ingroups. In CD2, missing sequences were treated as missing characters in all analyses. For CD1 and CD2, respectively, 662 and 671 sites were variable among which 454 and 460 were parsimony informative.

The Conoidea were found monophyletic, at least with the two combined analyses, although not always strongly supported (for CD2, MP and ML bootstraps, respectively: 65 and 79, Posterior Probabilities PP: 1). Within the Conoidea, two clades could be distinguished: clade A (MP bootstraps: 58, ML bootstraps: 68, PP: 0.73) and clade B (MP bootstraps: 28, ML bootstraps: 52, PP: 1). Within the clade A, the clade C is found strongly supported with ML bootstraps (91) and PP (1). Each analysis of the two combined datasets allowed the definition of the same 21 higher level clades, each of them strongly supported: MP and ML bootstraps >80 and PP >0.99 (Mason-Gamer and Kellogg, 1996; Zander, 2004). They included from 1 to 12 genera each (Tables 4 and 5, Fig. 1). Clades were numbered according to their position in the tree (Fig. 1). Clades 1–9 are included in clade A, and among them clades 1–4 are included in clade C. Clades 10–21 are included in clade B.

As long branches, for example that displayed by clade 9, could potentially disturb phylogenetic reconstructions (Felsenstein, 2004), the three analyses (MP, ML and BA) were conducted for the whole CD2, excluding specimens 17701 and 17702 (clade 9). The Conoidea were again separated in two clades: A′ (including clades 1–8) and B. The boostraps and PP were increased for both clades A′ (MP Bootstraps: 60, ML bootstraps: 77, PP: 1) and clade B (MP Bootstraps: 37, ML bootstraps: 60, PP: 1).

The position of the representatives of *Clavatula*, *Pusionella* and *Perrona*, for which we obtained only the COI sequence, could be analyzed only in the single gene analysis. The taxa clustered in the weakly supported clade 22 in all the performed COI gene analyses (Table 4, tree not shown). The weak resolution of the trees obtained with the COI gene did not permit the placement of clade 22 in either clade A or B.

All representatives of a genus clustered together in 1 of the 22 clades, except representatives of *Borsonia*, *Comitas*, *Conus* and *Leucosyrinx*. The representatives of *Borsonia* and *Conus* splitted, respectively, in clades 15, 16 and 19–21, each including only specimens from a single genus. The relationships between the two clades were not resolved and thus the monophyly of each of these genera cannot be

Table 4
Node supports of ML, MP and BA analyses for the four genes separately and for the two combined datasets

	COI			28S			18S			H3			CD1			CD2		
	ML	MP	BA	ML	MP	BA	ML	MP	BA	ML	MP	BA	ML	MP	BA	ML	MP	BA
Conoidea							35		0.94				85	71	0.98	79	65	1
Clade A	1						13			6		0.6	85	70	0.88	68	58	0.73
Clade B						1							53	33	0.52	52	28	1
Clade C				34	30	0.98				41	40		89	81	1	91	76	1
Clade 1				99	92	1	99		1	66	43	0.54	100	100	1	100	100	1
Clade 2	11	9					7	28	0.96		32		84	84	1	84	85	1
Clade 3	57	46		95	95	0.83		63	0.96	100	98	0.95	100	100	1	100	100	1
Clade 4	93	92	0.96	100	100	1		1		100	99	0.98	100	100	1	100	100	1
Clade 5	27	17	1	54	34	0.98				32	72	0.99	100	93	1	100	97	1
Clade 6	53	26	1	86	93	1	34	44	0.96	58	70	0.63	100	98	1	100	100	1
Clade 7		36	0.57	90	87	0.94				41	21	0.53	100	97	1	100	99	1
Clade 8	98	100	1	99	99	1	62	55	1							100	100	1
Clade 9	100	100	1	100	100	1		0		100	100	1	100	100	1	100	100	1
Clade 10				90	80	1					16		100	100	1	98	95	1
Clade 11				99	99	1				42	23	0.77	100	99	1	100	100	1
Clade 12				97	87	1	59	52	1	45	26	0.95	100	95	1	100	92	1
Clade 13	23	27	0.55	97	98	1	92	88	1	46	46		100	100	1	100	100	1
Clade 14	75	62	1	72	66	0.92				70	81	0.65	99	100	1	100	100	1
Clade 15	100	98	1	100	98	1	88	62	0.98	100	100	1	100	100	1	100	100	1
Clade 16	38	56		31						56	48	0.99	97	93	1	98	95	1
Clade 17		15		24	20		56	57	0.98	96	94	1	99	100	1	99	100	1
Clade 18	91	90	1		16					64	51	0.87	98	93	1	96	86	1
Clade 19				100	100	1	100	98	1	99	100	1	100	100	1	100	100	1
Clade 20				100	100	1				87	74	0.94	100	100	1	100	100	1
Clade 21	99	97	1	89	84	1				95	97	1	100	100	1	100	100	1
Clade 22	56	52	0.88															

Bootstraps values and Posterior Probabilities are given for 26 nodes (all Conoidea, clades A, B, C and clades 1–22). CD, Combined dataset. Gray cells correspond to unavailable data (sequences for specimens attributed to clade 8 were not obtained for H3 gene, and sequences for those attributed to clade 22 were successfully sequenced only for COI gene).

rejected. Conversely, the monophyly of genera *Leucosyrinx* and *Comitas* (clades 3, 4 and 9) can be rejected, since representatives of the two genera clustered in the clade 4.

4. Discussion

4.1. Classification of the Conoidea

Although not strongly supported, our analysis suggests that the superfamily Conoidea is monophyletic. However, the Conoidea and two outgroups used here (*Cancellopollia* and *Nassaria*) both belong in the Neogastropoda, the phylogeny of which is not well resolved (Harasewych et al., 1997; Colgan et al., 2007), and the monophyly observed here could thus be an artifact due to under-sampling within Neogastropoda. Within Conoidea, the large amount of diversity included in our dataset allows us to discuss the current classification at genus, subfamily, and family levels.

4.2. Accuracy of taxonomic delimitations at genus level

The genus is the lowest level for which we can discuss taxonomic delimitations since most of our specimens are not identified at species level. Among the 57 genera identified in our dataset, monophyly can be rejected for only two of them (*Leucosyrinx* and *Comitas*), which indicates that in most cases shell morphology is an appropriate predictor of generic allocations. Two further genera (*Borsonia* and *Conus*) are found to be diphyletic, but the position of the two defined clades is unresolved and thus monophyly cannot be excluded. Similarly, the polyphyly of some genera within the clades 1–22 can not be confirmed because of the lack of support for intra-clade nodes (results not shown).

4.3. Position of the genera within the subfamilies

Our analysis confirms many previous assignments of genera to subfamilies as in Taylor et al. (1993) (Table 1) and subsequent refinements of their classification. We thus confirm a position of *Conopleura* in the Drilliidae (Tippet and Tucker, 1995), of *Anacithara* in the Crassispirinae (Kilburn, 1994), of *Turridrupa* in the Turrinae (Kantor et al., 1997), of *Toxicochlespira* in the Mangeliinae (Sysoev and Kantor, 1990), and of *Glyphostomoides* in the Raphitominae (Shuto, 1983). However, several results do not confirm established classifications (Tables 1 and 5). The genus *Otitoma*, tentatively retained by Kilburn (2004) in the Mangeliinae based on shell characters, is here found to be in the Crassispirinae. The genus *Lienardia*, earlier classified in the Mangeliinae, is here placed in clade 12, identified as a Clathurellinae. (Furthermore, specimens attributed to *Lienardia* display several types of protoconchs and *Lienardia* as currently understood is probably a highly polyphyletic assemblage of species, some belonging to Raphitominae—

Table 5
Genera included in the clades A, B, C and 1–22, and association to a taxonomic name proposed in previous classifications (see Table 1)

Clade 1 *Clavus* *Conopleura* *Splendrillia*		**DRILLIIDAE**	**C**	**A** **TURRIDAE** **+ DRILLIIDAE** **+ TEREBRIDAE ?**
Clade 2 *Crassispira* *Funa* *Inquisitor* *Otitoma* *Ptychobela*		**Crassispirinae**		
Clade 3 & 4 *Comitas* **Clade 3** *Leucosyrinx*		**"Cochlespirinae"**		
Clade 5 *Gemmula* *Lophiotoma* *Turris* *Turridrupa*		**Turrinae**		
Gemmuloborsonia		**"Turrinae"**		
Clade 6 *Cinguloterebra* *Terebra*		**TEREBRIDAE**		
Clade 7 *Anacithara* *Ceritoturris* *Horaiclavus* *Iwaoa*		**"Crassispirinae"**		
Clade 8 *Cochlespira*		**Cochlespirinae**		
Clade 9 *Leucosyrinx*		**"Cochlespirinae"**		
Clade 10 *Daphnella* *Eucyclotoma* *Glyphostomoides* *Gymnobela* *Kermia* *Pleurotomella*	*Raphitoma* *Rimosodaphnella* *Teretiopsis* *Thatcheria* *Tritonoturris* *Veprecula*	**Raphitominae**		**B** **CONIDAE**
Clade 11 *Eucithara* *Guraleus* *Leiocithara* *Macteola* *Mangelia*		**Mangeliinae**		
Clade 12 *Lienardia* *Etrema* *Nannodiella* **Clade 13** *Mitromorpha* *Anarithma* **Clade 14** *Tomopleura* **Clade 15 & 16** *Borsonia*		**Clathurellinae**		
Clade 17 *Benthomangelia* *Toxicochlespira*		**"Mangeliinae"**		
Clade 18 *Heteroturris* *Microdrillia* *Typhlomangelia*		**Clathurellinae**		
Benthofascis		**Conorbinae**		
Clade 19 *Conus*		**Coninae**		
Clade 20 *Bathytoma*		**Clathurellinae**		
Clade 21 *Conus*		**Coninae**		
Clade 22 *Clavatula* *Perrona* *Pusionella*		**CLAVATULIDAE**		**In TURRIDAE or CLAVATULIDAE**

Subfamilies are in bold, families in bold and capital. Type genera present in our dataset are underlined.

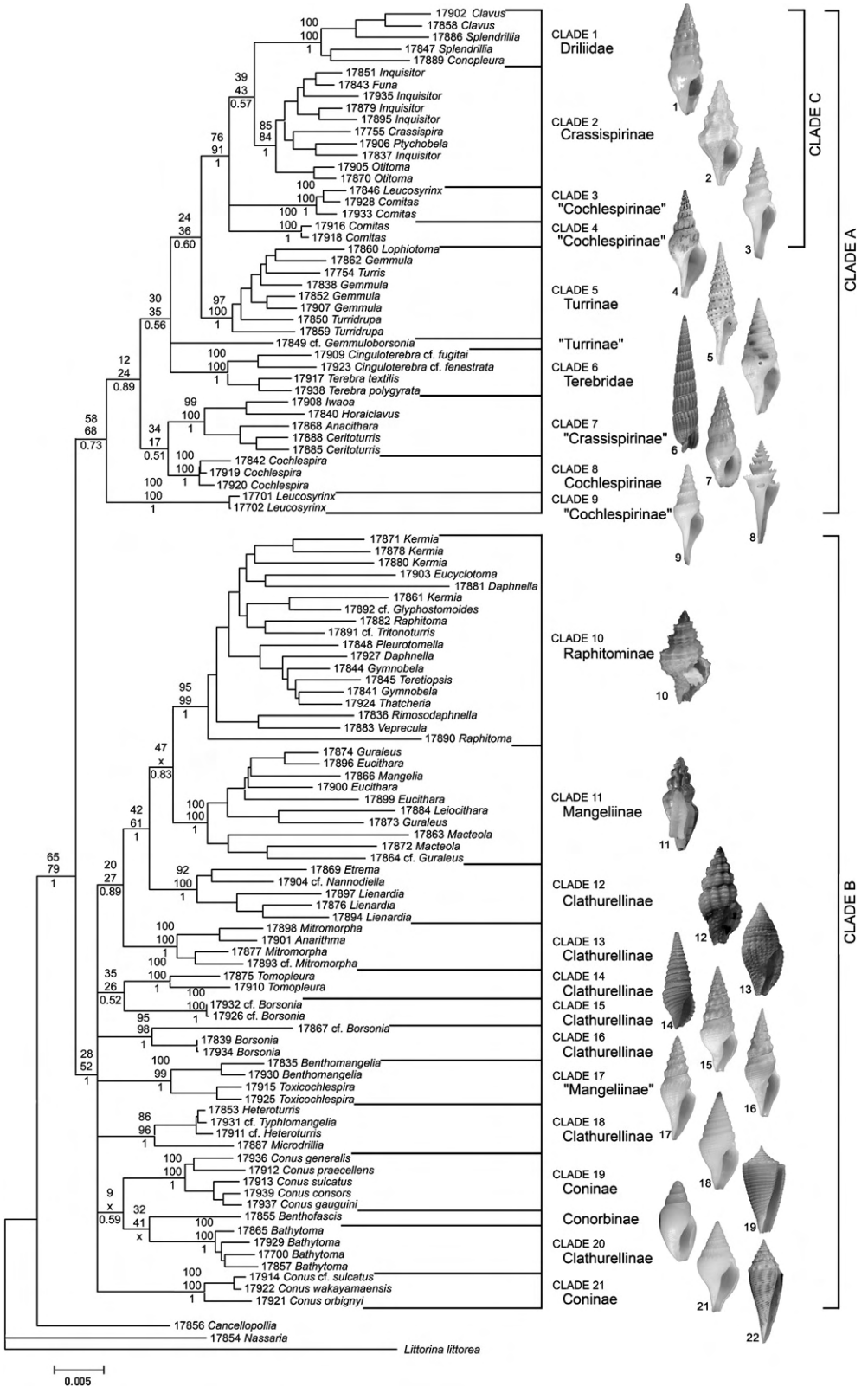


Fig. 1. Consensus tree of MP, ML and BA results obtained with CD2. Nodes presented here were found with at least two of the three methods used. Top downwards, MP bootrstraps, ML boostraps and Posterior Probabilities are specified for each node. Supports for intranodes of clades 1–21 are not presented. Taxonomic names are attributed for each of the clades 1–21, as explained in the text. One example of shell, corresponding to the type-genus when possible, is given for each clade. Illustrated specimens are quoted in the Table 2.

not represented in our molecular sampling—and others to Clathurellinae—as the specimens studied here). The position of *Gemmuloborsonia*, assigned to the Turrinae (Sysoev and Bouchet, 1996; Medinskaya, 2002), is unresolved.

4.4. Robustness of subfamilies delimitations

We found discrepancies between our phylogeny and previous classifications at the subfamily level. Thus, crassispirine genera are present in two clades (2 and 7), one of them (clade 2) containing the type genus. The polyphyly of this subfamily is supported by the existence of clade C, which includes clade 2, but excludes clade 7. Since the relationships between clade 7 and others clades within clade A are not resolved, it is unconclusive whether clade 7 must be ranked as its own subfamily or whether it must be grouped together with another existing subfamily. The subfamily Cochlespirinae as currently construed appears polyphyletic too, with four distinct clades (3, 4, 8 and 9), one of them (clade 8) containing the type genus. As for the Crassispirinae, the polyphyly of the Cochlespirinae is supported by the existence of clade C, which includes clades 3 and 4, but excludes clades 8 and 9. However, because of the limits of the resolution of the deeper nodes, it is inconclusive whether clades 3 and 4 should be allocated to the Crassispirinae or should constitute a new subfamily; the subfamily Cochlespirinae could be limited to clade 8, or could also include clade 9.

In the next three cases, polyphyly is possible but not demonstrated because of a general lack of support for deeper nodes in clade B. (a) Relationships between the two highly divergent clades (clades 11 and 17) of the Mangeliinae are not resolved and our results are inconclusive on the non-monophyly of the subfamily. (b) Coninae also ends up as two distinct clades (clades 19 and 21), a result already obtained by Duda and Kohn (2005). (c) The subfamily Clathurellinae is split into seven clades (clades 12, 13, 14, 15, 16, 18 and 20), but the non-monophyly of these clades is not demonstrated. With one exception, our molecular clathurelline clades correspond to intra-clathurelline "groups" defined by Taylor et al. (1993), suggesting that these may warrant formal naming as tribes. The exception is clade 18 which includes on one hand the genus *Typhlomangelia* (placed in the "borsoniid group" by Taylor et al., 1993) and on the other hand the genera *Heteroturris* and *Microdrillia* (placed in the "tomopleurid group" by Taylor et al., 1993).

4.5. Robustness of families delimitations

Finally, our results also permit a discussion of family classification within Conoidea. Taylor et al.'s (1993) anatomical study suggested a closer relationship of Clathurellinae, Conorbinae, Mangeliinae, Oenopotinae and Raphitominae to *Conus* than to other members of the family Turridae *s.l.* and their extension of Conidae included these turrid subfamilies. In our study, clade B, although weakly supported, corresponds to Taylor et al.'s (1993) family Conidae, thus supporting its monophyly.

Our study also revealed another weakly supported deep clade (clade A) that includes genera classified by Taylor et al. (1993) in three different families: Drilliidae, Terebridae and Turridae *s.s.* (consisting of Clavatulinae, Cochlespirinae, Crassispirinae, Turrinae and Zonulispirinae). Genera of the Drilliidae (clade 1) are included in clade C. This well-supported clade also contains taxa of the Turridae *s.s.* (Crassispirinae and *Comitas*), and excludes the other taxa of the Turridae *s.s.* Consequently, Turridae *s.s.* are not monophyletic. Furthermore, according to Kantor (2006), the radula of Drilliidae is not fundamentally different from that of Turridae *s.s.*

Within clade A, the monophyly of the Terebridae is supported but its relationships with other clades of Turridae *s.s.* is not resolved. The strong support obtained for clade A′ (clade A without clade 9) indicates that Terebridae are closely related to Turridae *s.s.* Moreover, the increase of clade support from A to A′ suggests an artifact effect of clade 9 on the phylogenetic reconstruction, e.g., a long branch attraction effect with the outgroups. This phenomenon could be avoided by increasing the amount of diversity included in the analysis (Bergsten, 2005). A close relationship between Terebridae and Turridae *s.s.* had already been suggested by Cossmann (1896), and Powell (1942, 1966), based on the resemblance of the shells of Terebridae and of the clavatuline genus *Pusionella*. Based on this observation and the fossil record, Powell (1966) speculated that Terebridae were derived from the Clavatulinae. Our results suggest that Turridae *s.s.* could be closer to Terebridae than to Conidae, but the question of whether Terebridae is included in Turridae *s.s.* or is its sister group still remains unresolved.

4.6. Towards a stabilized system for Conoidea

The weak support of neogastropod molecular phylogenies available in literature is supposed to be the consequence of an early radiation of the group (Harasewych et al., 1997; Colgan et al., 2003, 2007). Genes used in those studies were not adequate to resolve the relationships between clades that emerged during this radiation. In our study, we used the same genes, albeit at a lower taxonomic level, but deeper nodes are not resolved either. In view of the fact that most subfamilies of Turridae *s.l.* were already present in the Eocene, Powell (1966) dated their divergence before the Upper Cretaceous (before 65MY). As for other animal groups (e.g., Strugnell et al., 2005; Fry et al., 2006), resolving phylogenetic relationships between those early divergences seems to require slow-evolving genes. In this perspective, nuclear coding genes, rarely used in mollusk phylogenies, could be useful to resolve early relationships within Conoidea as well as deeper relationships within gastropods.

The taxonomic sampling used here allows an estimation of molecular variability within clades at each level: several genera are included in each subfamily, several subfamilies are included in each family, and most of the families defined by Taylor et al. (1993) are present. This strategy, where taxonomic sampling is hierarchically organized, is clearly required to discuss monophyly of each of those groups, and some problems are thus highlighted at each taxonomic level.

However, even with a dataset of 57 genera, covering most of the previously recognized families and subfamilies of Conoidea, the present study only brings preliminary results. At genus level, these 57 genera represent only 17% of the 340 already described recent genera and it is further clear that the shell-based current taxonomic extension of many genera will not stand after molecular testing. At subfamily and family levels, although a large part of the conoidean diversity is represented in this study, the families Strictispiridae and Pseudomelatomidae, the subfamilies Zonulispirinae and Zemaciinae in Turridae *s.s.*, the Pervicaciinae in Terebridae and the Oenopotinae in Conidae, are not part of our taxon sampling. The highly divergent clades found here in several subfamilies as previously defined demonstrate the need for further research in order to better restrict the taxonomic extensions of the already known subfamilies and probably formally name new subfamilies and/or tribes. Finally, at family level, new relationships are suggested. As a remake of the *Conus* story, it now appears that the long recognized family Terebridae does not stand alone apart from the rest of the Conoidea, but could be the sister-group or even part of the Turridae *s.s.*

Acknowledgments

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the *Tropical Deep Sea Benthos* programme on board R/V *Alis*, that generated the deep-sea samples used in this study. The Total Foundation supported the Panglao 2004 and Santo 2006 biodiversity surveys and Ellen Strong is thanked for her role in molecular sampling during these expeditions. We are also pleased to thank the staff of MNHN's "Service de Systématique Moléculaire" for technical facilities; the Consortium National de Recherche en Génomique, Génoscope for the sequencing facilities; S. Gofas for collecting African taxa; Barbara Buge and Pierre Lozouet for the pictures; and P. Lopez, N. Vidal, P. Gaubert, A. Warén and E. Strong for constructive comments on, and improvements of, the manuscript. Y. Terryn and A. Kohn identified specimens of Terebridae and *Conus*, respectively. The authors thank two anonymous referees for very constructive comments on the manuscript.

References

Bergsten, J., 2005. A review of long-branch attraction. Cladistics 21, 163–193.

Bouchet, P., 1990. Turrid genera and mode of development: the use and abuse of protoconch morphology. Malacologia 32, 69–77.

Bouchet, P., Rocroi, J.-P., 2005. Classification and nomenclator of Gastropods families. Malacologia 47.

Bouchet, P., Warén, A., 1980. Revision of the north-east Atlantic bathyal and abyssal Turridae (Mollusca, Gastropoda). J. Molluscan Stud. Suppl. 8, 1–119.

Chang, C.-K., 1995. Reevaluation of the classification of Turridae. Bull. Malacol. Taiwan 19, 49.

Chang, C.-K., 2001. Small Turrids of Taiwan. Thorsson, W.M., Taiwan.

Colgan, D.J., Ponder, W.F., Beacham, E., Macaranas, J., 2003. Gastropod phylogeny based on six fragments from four genes representing coding or non-coding and mitochondrial or nuclear DNA. Molluscan Res. 23, 123–148.

Colgan, D.J., Ponder, W.F., Beacham, E., Macaranas, J., 2007. Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). Mol. Phylogenet. Evol. 42, 717–737.

Cossmann, M., 1896. Essais de Paléoconchologie comparée, 2ème livraison. Paris.

Darlu, P., Tassy, P., 1993. La reconstruction phylogénétique. Concepts et méthodes. Masson, Paris.

Dayrat, B., Tillier, A., Lecointre, G., Tillier, S., 2001. New clades of Euthyneuran Gastropods (Mollusca) from 28S rRNA sequences. Mol. Phylogenet. Evol. 19, 225–235.

Duda, T.F., Kohn, A.J., 2005. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Mol. Phylogenet. Evol. 34, 257–272.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Sunderland.

Fischer, P., 1887. Manuel de conchyliologie et de paléontologie conchyliologique. F. Savy, Paris.

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. Mol. Mar. Biol. Biotechnol. 3, 294–299.

Fry, B.G., Vidal, N., Norman, J.A., Vonk, F.J., Scheib, H., Ramjan, S.F.R., Kuruppu, S., Fung, K., Hedges, S.B., Richardson, M.K., Hodgson, W.C., Ignjatovic, V., Summerhayes, R., Kochva, E., 2006. Early evolution of the venom system in lizards and snakes. Nature 439, 584–588.

Giribet, G., Carranza, S., Bagui, J., Riutort, M., Ribera, C., 1996. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. Mol. Biol. Evol. 13, 76–84.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Harasewych, M.G., Adamkewicz, S.L., Blake, J.A., Saudek, D., Spriggs, T., Bult, C.J., 1997. Neogastropod phylogeny: a molecular perspective. J. Molluscan Stud. 63, 327–351.

Hassouna, M., Michot, B., Bachellerie, J.-P., 1984. The complete nucleotide sequence of mouse 28SrRNA gene. Implications for the process of size increase of the large subunit rRNA of higher eukaryotes. Nucleic Acids Res. 12, 3563–3583.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA Barcodes. Proc. R. Soc. Lond. B 270, 313–321.

Hedley, C., 1922. A revision of the Australian Turridae. Rec. Aust. Mus. 13, 213–359.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Huelsenbeck, J.P., Ronquist, F., Hall, B., 2001. MrBayes: bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jovelin, R., Justine, J.-L., 2001. Phylogenetic relationships within the Polyopisthocotylean monogeneans (Plathyhelminthes) inferred

from partial 28S rDNA sequences. Int. J. Parasitol. 31, 393–401.

Kantor, Y.I., 2006. On the morphology and homology of the "central tooth" in the radula of Turrinae (Conoidea: Turridae). Ruthenica 16, 47–52.

Kantor, Y.I., Medinskaya, A.I., Taylor, J.D., 1997. Foregut anatomy and relationships of the Crassispirinae (Gastropoda, Conoidea). Bull. Nat. Hist. Mus. Lond. (Zool.) 63, 55–92.

Kantor, Y.I., Taylor, J.D., 1994. The foregut anatomy of *Strictispira paxillus* (Reeve, 1845) (Conoidea: Strictispiridae). J. Molluscan Stud. 60, 343–346.

Kilburn, R.N., 1983. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 1. Subfamily Turrinae. Natal Mus. Ann. 25, 549–585.

Kilburn, R.N., 1985. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 2. Subfamily Clavatulinae. Natal Mus. Ann. 26, 417–470.

Kilburn, R.N., 1986. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 3. Subfamily Borsoniinae. Natal Mus. Ann. 27, 633–720.

Kilburn, R.N., 1988. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 4. Subfamily Drilliinae, Crassispirinae and Strictispirinae. Natal Mus. Ann. 29, 167–320.

Kilburn, R.N., 1989. Notes on Ptychobela and Brachytoma, with the description of a new species from Mozambique (Mollusca: Gastropoda: Turridae). Natal Mus. Ann. 30, 185–196.

Kilburn, R.N., 1991. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 5. Subfamily Taraninae. Natal Mus. Ann. 32, 325–339.

Kilburn, R.N., 1992. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 6. Subfamily Mangeliinae, section 1. Natal Mus. Ann. 33, 461–575.

Kilburn, R.N., 1993. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 6. Subfamily Mangeliinae, section 2. Natal Mus. Ann. 34, 317–367.

Kilburn, R.N., 1994. Turridae [*s.l.*] (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 7. Subfamily Crassispirinae, section 2. Natal Mus. Ann. 35, 177–228.

Kilburn, R.N., 1995. Turridae (*s.l.*) of southern Africa and Mozambique (Mollusca: Gastropoda: Conoidea). Part 8. Conidae: subfamily Mangeliinae, section 3. Natal Mus. Ann. 36, 261–269.

Kilburn, R.N., 2004. The identities of *Otitoma* and *Antimitra* (Mollusca: Gastropoda: Conidae and Buccinidae). Afr. Invertebr. 45, 263–270.

Kohn, A.J., 1990. Tempo and mode of evolution in Conidae. Malacologia 32, 55–67.

Mason-Gamer, R.J., Kellogg, E.A., 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Graminae). Syst. Biol. 45, 524–545.

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., 2002. Foregut anatomy of the Turrinae (Gastropoda, Conoidea, Turridae). Ruthenica 12, 135–159.

Medinskaya, A.I., Sysoev, A., 2003. The anatomy of *Zemacies excelsa*, with a description of a new subfamily of Turridae (Gastropoda, Conoidea). Ruthenica 13, 81–87.

Morrison, J.P.E., 1965. On the families of Turridae. In: The 31st Annual Meeting of the American Malacologial Union, pp. 1–2.

Okusu, A., Schwabe, E., Eernisse, D.J., Giribet, G., 2003. Towards a phylogeny of chitons (Mollusca, Polyplacophora) based on combined analysis of five molecular loci. Org. Divers. Evol. 3, 281–302.

Posada, D., Crandall, K.A., 2001. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). Mol. Biol. Evol. 18, 897–906.

Powell, A.W.B., 1942. The New-Zealand recent and fossil Mollusca of the family Turridae. With general notes on turrid nomenclature and systematics. Bull. Auckland Inst. Mus. 2, 1–192.

Powell, A.W.B., 1966. The molluscan families Speightiidae ans Turridae. An evaluation of the valid taxa, both recent and fossil, with list of characteristic species. Bull. Auckland Inst. Mus. 5, 1–184.

Rosenberg, G., 1998. Reproducibility of results in phylogenetic analysis of mollusks: a reanalysis of the Taylor, Kantor and Sysoev (1193) data set for conoidean gastropods. Am. Malac. Bull. 14, 219–228.

Shimek, R.L., Kohn, A.J., 1981. Functional morphology and evolution of the toxoglossan radula. Malacologia 20, 423–438.

Shuto, T., 1983. New turrid taxa from the Australian waters. Mem. Fac. Sci. Kyushu Univ. Serie D Geol. 25, 16.

Soltis, P.S., Soltis, D.E., 2003. Applying the bootstrap in phylogeny reconstruction. Stat. Sci. 18, 256–267.

Strugnell, J., Norman, M., Jackson, J., Drummond, A.J., Cooper, A., 2005. Molecular phylogeny of coleoid cephalopods (Mollusca: Cephalopoda) using a multigene approach; the effect of data partitioning on resolving phylogenies in a bayesian framework. Mol. Phylogenet. Evol. 37, 426–441.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony* and Other Methods. Sinauer Associates, Sunderland.

Sysoev, A., Bouchet, P., 1996. Taxonomic reevaluation of *Gemmuloborsonia* Shuto, 1989 (Gastropoda: Conoidea), with a description of recent deep-water species. J. Molluscan Stud. 62, 75–87.

Sysoev, A., Kantor, Y.I., 1990. A new genus and species of "*Cochlespira*-like" turrids. Apex 5, 1–6.

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.

Taylor, J.D., Kantor, Y.I., Sysoev, A.V., 1993. Foregut anatomy, feedings mechanisms and classification of the Conoidea (= Toxoglossa)(Gastropoda). Bull. Nat. Hist. Mus. Lond. 59, 125–170.

Terryn, Y., 2007. A Collectors Guide to Recent Terebridae (Mollusca: Neogastropoda). ConchBooks, Hackenheim.

Thiele, J., 1929. Handbuch der systematischen weichtierkunde. G. Fischer, Jena.

Tippet, D.L., Tucker, J.K., 1995. Taxonomic notes on *Kenyonia* Brazier and *Conopleura* Hinds (Gastropoda: Conoidea). Nautilus 108, 37–38.

Tucker, J.K., 2004. Catalogue of recent and fossil turrids (Mollusca: Gastropoda). Zootaxa 682, 1–1295.

Wenz, W., 1938–1944. Teil 1: Allgemeiner Teil und Prosobranchia. In: Schindewolf, O.H. (Ed.), Handbuch de paläozoologie, Band 6, Gastropoda. Borntraeger, Berlin.

Wiens, J.J., 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? Syst. Biol. 47, 625–640.

Zander, R.H., 2004. Minimal values for reliability of bootstrap and Jackknife proportions, Decay index, and bayesian posterior probability. PhyloInformatics 2, 1–13.

3. Ajout de nouveaux caractères et de nouveaux spécimens

D'autres analyses ont depuis ce premier article été réalisées pour répondre aux questions qui n'y étaient pas abordées, ou étaient restées irrésolues. Ainsi, plusieurs familles de Néogastropodes ont été incluses dans l'analyse pour tester la monophylie des Conoidea et identifier son groupe-frère. De plus, plusieurs taxa au sein des Conoidea manquant dans la première étude, tels que la famille de Clavatulidae, et les sous-familles des Zemaciinae et des Oenopotinae, ont été ajoutés. Enfin, d'autres gènes ont également été séquencés pour la plupart des taxa (12S et 16S), permettant la résolution de plusieurs relations entre clades définis dans la première analyse. Ce jeu de données est également complété par des séquences fournies par Chris Meyer (Smithsonian Institution, Washington) et plusieurs de ses collaborateurs, ce qui permet d'inclure quelques taxa qui étaient manquants (comme les genres-types de certaines sous-familles, mais aussi des représentants de la famille de Pseudomelatomidae et de la sous-famille des Zonulispirinae).

Ce nouveau jeu de données inclus 120 spécimens, représentant 101 genres, 13 sous-familles et 6 familles. Les analyses sont actuellement encore en cours : seule une analyse bayésienne a été pour l'instant réalisée. Les résultats préliminaires issus de cette analyse sont présentés sur la figure 5.

L'analyse de ce nouveau jeu de données confirme les résultats obtenus lors de la première étude mais permet surtout de résoudre plusieurs relations phylogénétiques entre les différents clades. La famille des Terebridae est incluse dans la famille des Turridae *s.s.* La sous-famille des Zemaciinae, qui avait été placée dans la famille des Turridae *s.s.* sur la base de la forme de la coquille (les caractères anatomiques typiques des Conoidea étant perdus chez ce groupe – Medinskaya & Sysoev 2003), appartient à la famille des Conidae. D'autres résultats sont également à souligner, comme la polyphylie confirmée de plusieurs sous-familles, et l'inclusion des Oenopotinae dans les Mangeliinae et des Taraniinae dans les Raphitominae.

La phylogénie moléculaire obtenue permet également de discuter de la validité des caractères utilisés dans les précédentes classifications. Par exemple, l'utilisation du caractère « présence d'un opercule » conduit à la formation d'un groupe polyphylétique (Turridae *s.s.* + Oenopotinae). L'association des Mitromorphinae et du genre *Borsonia* (au sein des Borsoniinae, Powell 1942), sur la base de la présence d'un pli columellaire, correspond également à un groupe polyphylétique.

4. Vers une nouvelle classification

Même si toutes les relations phylogénétiques ne sont pas encore éclaircies au sein des Conoidea (entre certaines sous-familles au sein des Conidae et des Turridae *s.s.* essentiellement), les nouvelles hypothèses apportées par cette analyse permettent de proposer une nouvelle classification pour ce groupe. Dans le premier article (page 19), l'accent a été mis sur cette transition entre phylogénie et classification. L'association entre un clade et un nom de taxon doit en effet suivre des règles de nomenclature (Encadré 3) qui ne sont pas systématiquement suivies dans la litérature. En suivant ces règles, les clades définis par l'analyse phylogénétique peuvent être associés à des noms de sous-familles et de familles disponibles dans la littérature afin de proposer une nouvelle classification des Conoidea (Figure 5). Ce travail est actuellement en cours (tous les clades ne sont donc pas encore liés à un nom) et réalisé avec la collaboration de spécialistes du groupe (Philippe Bouchet, Alexander Sysoev, Zoological Museum of Moscow et Yuri Kantor, Russian Academy of Sciences).

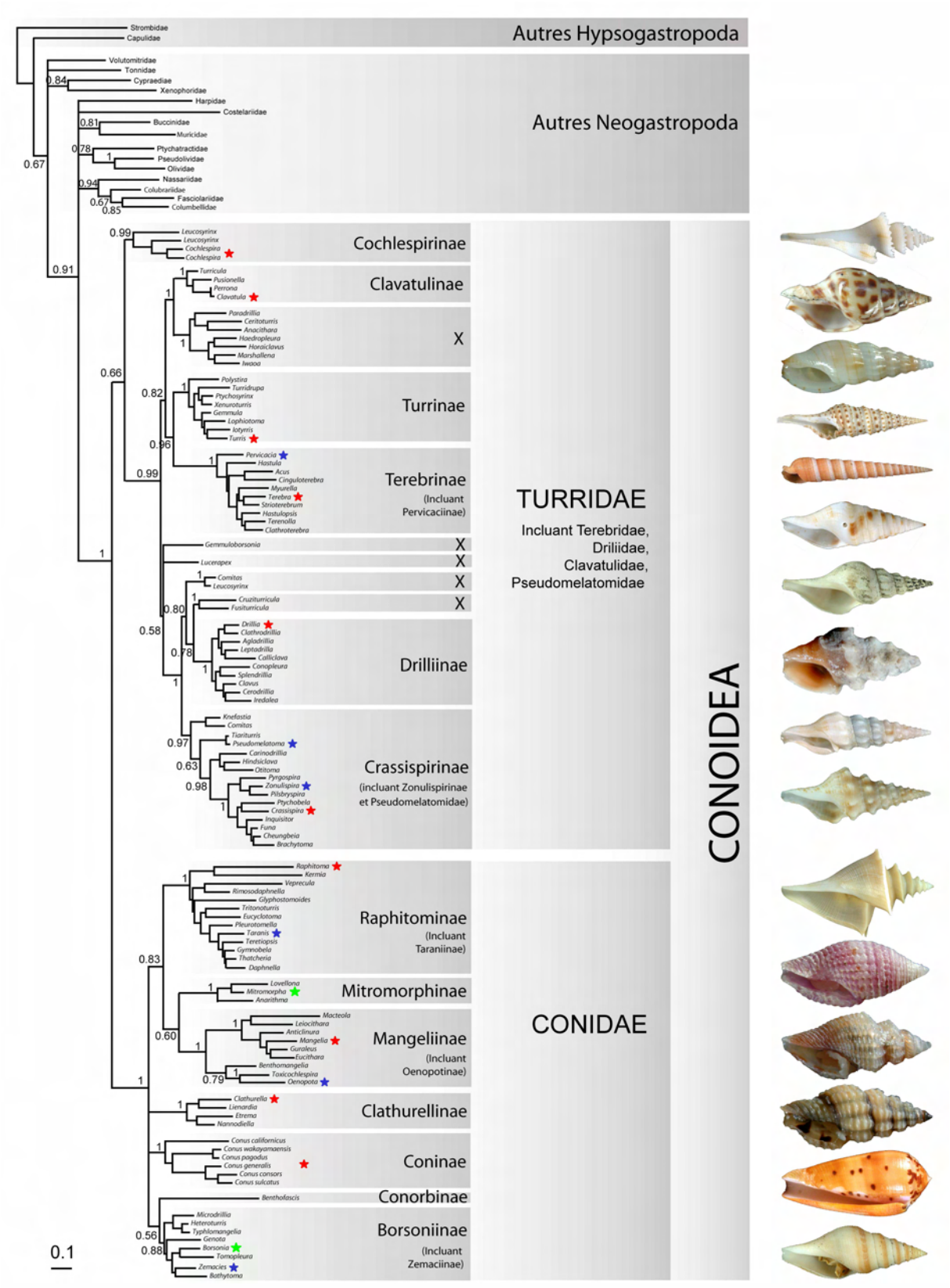


Figure 5 : Phylogénie moléculaire des Conoidea. Analyse bayésienne réalisée avec les gènes COI, 12S et 16S. Une nouvelle classification est proposée. Des noms de sous-familles et de familles sont attribués à chaque clade (Encadré 3). Etoile rouge : genre-type d'un groupe valide dans la dernière classification des Conoidea disponible (Bouchet & Rocroi 2005), et toujours valide dans cette nouvelle classification ; étoile verte : genre-type d'un groupe non valide dans la dernière classification et valide dans cette nouvelle classification ; étoile bleue : genre-type d'un groupe valide dans la dernière classification et non valide dans cette nouvelle classification. Certains clades, notés X, n'incluent pas de genre-type. A noter que l'inclusion de la sous-famille des Cochlespirinae dans les Turridae *s.s.* n'est pas soutenue dans cette analyse. De haut en bas : *Cochlespira*, *Pusionella*, *Horaiclavus*, *Turris*, *Terebra*, *Gemmuloborsonia*, *Comitas*, *Clavus*, *Inquisitor*, *Funa*, *Thatcheria*, *Mitromorpha*, *Leiocithara*, *Lienardia*, *Conus*, *Bathytoma*.

ENCADRE 3

Etablir le lien entre phylogénie et classification

D'après la dernière révision (Bouchet & Rocroi 2005), les Conoidea incluent sept familles et 13 sous-familles. Les noms de ces groupes (familles et sous-familles) sont considérés comme valides, c'est-à-dire qu'ils correspondent au nom le plus ancien disponible dans la littérature pour désigner chacun des groupes auquel ils correspondent. Par exemple, il existe plusieurs noms pour désigner la sous-famille des Raphitominae, chacun d'entre eux ayant été publiés dans des articles à des dates différentes. Le nom Raphitominae a par exemple été proposé en 1875 par Bellardi, et le nom Daphnellinae a été proposé en 1904 par Casey. Le nom Raphitominae étant plus ancien (principe d'antériorité), il sera considéré comme valide pour désigner ce groupe.

Pour attribuer chacun de ces noms valides à un des clades définis par l'analyse phylogénétique et ainsi établir une classification des Conoidea, la règle suivante a été appliquée. La présence d'un genre-type dans un clade peut permettre d'attribuer un nom à ce clade. Un genre-type est « le porte nom d'un taxon nominal » (ICZN, http://www.iczn.org/iczn/index.jsp), c'est-à-dire un genre dont le nom a servi de base à l'établissement d'un groupe taxonomique de rang supérieur (sous-famille, famille,…). Par exemple, le genre *Raphitoma* est le genre-type de la sous-famille des Raphitominae. Le clade qui contient un genre-type prendra donc le nom du groupe dont le genre en question est le genre-type. Par exemple, le clade qui contient le genre *Raphitoma* prendra le nom Raphitominae. De plus, dans le cas des Raphitominae, le clade correspondant ne contient que des genres placés dans les Raphitominae (Taylor *et al.* 1993), sauf le genre *Taranis* (Figure 5).

Toutes les sous-familles et familles considérées comme valides au sein des Conoidea (sauf la famille des Strictispiridae) sont représentées dans l'analyse phylogénétique. De plus, la plupart des genres-types de ces groupes valides sont inclus dans l'analyse (sauf pour la sous-famille des Conorbinae, dont le genre-type est fossile). Suivant la règle décrite, ces noms ont donc pu être attribués chacun à un clade.

Cependant, plusieurs clades n'incluent aucun genre-type de sous-familles ou de familles valides. Certains de ces clades incluent cependant des genres-types de groupes considérés comme invalides actuellement. C'est par exemple le cas du clade incluant les genres *Mitromorpha*, *Anarithma* et *Lovellona*. Le genre *Mitromorpha* est le genre-type de la sous-famille des Mitromorphinae, considérée actuellement comme synonyme de la sous-famille des Clathurellinae, et donc invalide (le nom Mitromorphinae est plus récent que le nom Clathurellinae). Le nom Mitromorphinae peut par conséquent être associé à ce clade, et il sera par la suite considéré comme valide, la synonymie avec le nom Clathurellinae étant annulée.

Enfin, d'autres clades (notés X sur la figure 5) n'incluent aucun genre-type, qu'ils correspondent à un groupe valide ou non. Dans ces cas, un nouveau nom devra être créé.

B. Phylogénie et évolution des caractères

La phylogénie moléculaire permet de proposer une nouvelle classification pour les Conoidea mais également d'analyser l'évolution de caractères particuliers. Ainsi, au sein des Terebridae, l'évolution de l'appareil venimeux a été analysée. Le séquençage de 3 gènes mitochondriaux (COI, 12S et 16S) de plus de 150 spécimens, représentant douze des quinze genres actuellement reconnus chez les Terebridae, par Mandë Holford (étudiante américaine en post-doc, sous la direction de Baldomero Olivera, Université de l'Utah), a permis de reconstruire une phylogénie de ce groupe. Cette phylogénie a ensuite été utilisée pour retracer l'évolution de ce caractère « présence/absence de l'appareil venimeux » dans la famille des Terebridae (Article 2).

Les résultats apportés par cette analyse permettent dans un premier temps de discuter la classification de ce groupe. La plupart des genres actuellement considérés comme valides (Terryn 2007) apparaissent polyphylétiques. En fait, seuls les deux genres monotypiques (*Terenolla* et *Impages*) sont monophylétiques. La phylogénie s'organise en cinq clades principaux (notés A à E dans l'article 2), avec notamment le clade A (qui n'inclut que l'espèce "*Terebra jungi*") en groupe-frère de tous les autres Terebridae. L'espèce *T. jungi* a été récemment décrite (Lai 2001), et les résultats de l'analyse moléculaire révèlent donc que cette espèce n'appartient pas au genre *Terebra* (clade C). Au niveau spécifique, les résultats semblent indiquer que plusieurs individus qui sont considérés comme conspécifiques possèdent des séquences très différentes. C'est par exemple le cas au sein de l'espèce *Terebra textilis*, représentée par quatre lignées très différentes, et qui pourrait donc inclure quatre espèces distinctes.

Pour plusieurs taxa identifiés dans notre jeu de données, et au moins pour quelques représentants de chacun des cinq clades principaux, l'information relative à la présence ou à l'absence d'un appareil venimeux était disponible dans la littérature, ou a pu être obtenue grâce à John Taylor (Natural History Museum of London) et Alexander Sysoev (Zoological Museum of Moscow). L'analyse de ce caractère a permis de retracer son évolution au sein des Terebridae. L'appareil venimeux serait donc présent chez l'ancêtre commun à tous les Terebridae, et aurait été perdu deux fois indépendamment au cours de l'évolution. Les deux clades dont les espèces ne possèdent pas d'appareil venimeux avaient déjà été identifiés par Miller (1970), mais la phylogénie moléculaire apporte ici un élément de plus en précisant que ces deux groupes ne sont pas groupes-frères.

Ce résultat a également des intérêts dans un autre domaine de recherche. Les Conoidea sont en effet caractérisés (pour la plupart) par la possession d'un appareil venimeux, et donc par leur capacité à produire des toxines. Certaines d'entre elles ont d'ores et déjà été synthétisées et une est actuellement commercialisée et utilisée comme analgésique (Letourneux 2004). Etant donnée la diversité de toxines trouvées chez une espèce et surtout le nombre d'espèces, connues ou non, au sein des Conoidea, le groupe est devenu une cible pour la découverte de nouveaux composés pharmaceutiques. Dans le genre *Conus*, chaque espèce possède entre 100 et 200 toxines qui lui sont propres, ce qui permet d'estimer que potentiellement, le genre *Conus* (600 espèces environ) constitue une réserve de plusieurs dizaines de milliers de toxines (Olivera 1997). Des estimations similaires peuvent être faites avec l'ensemble des Conoidea, qui inclut certainement plus de 10000 espèces (Bouchet *et al*. 2002).

La recherche biomédicale encourage la découverte de nouvelles toxines animales qui pourraient conduire à la synthèse de nouveaux médicaments. Plusieurs études se basent sur la reconnaissance de groupes phylogénétiquement éloignés pour maximiser les chances de découvrir ces nouvelles toxines (Olivera 2002 ; Olivera 2006 ; Olivera & Teichert 2007).

L'analyse phylogénétique des Terebridae indique que la recherche de nouvelles toxines sera plus efficace en se concentrant sur les trois clades distincts possédant un appareil venimeux, et non sur un seul d'entre eux. Par exemple, l'espèce *Terebra jungi*, moins commune que d'autres espèces du même genre, ne constituait pas une priorité pour la recherche de toxines. Les résultats de l'analyse phylogénétique montrent que cette espèce n'appartient pas au genre *Terebra*, et possède potentiellement des toxines différentes.

Cette approche est prometteuse pour tous les Conoidea. Plusieurs études se concentrent actuellement sur le genre *Conus*, et l'approche phylogénétique semble adaptée pour identifier les groupes susceptibles d'être les plus intéressants pour la recherche de toxines originales (Olivera & Teichert 2007). L'analyse de l'évolution du caractère présence/absence de l'appareil venimeux peut donc également être effectuée pour d'autres taxa au sein des Conoidea. Je collabore actuellement avec un étudiant russe, Alexander Fedosov (Russian Academy of Sciences), pour établir une phylogénie moléculaire des Raphitominae. Les premiers résultats indiquent que plusieurs évènements de perte de l'appareil venimeux ont également eu lieu dans cette sous-famille. Les Raphitominae ne font pas actuellement l'objet de recherche de toxines, mais cette analyse phylogénétique pourra servir de guide, comme pour les Terebridae, pour leur identification et leur synthèse.

Evolution of the Toxoglossa Venom Apparatus as Inferred by Molecular Phylogeny of the Terebridae

Mandë Holford[1*], Nicolas Puillandre[2], Yves Terryn[3], Baldomero Olivera[1], Corinne Cruaud[4] and Philippe Bouchet[2]

[1] University of Utah, Biology Department, Salt Lake City, Utah 84112, USA

[2] Muséum National d'Histoire Naturelle (MNHN), Departement Systematique et Evolution, UMR 7138 Systématique, Adaptation, Evolution, 55 rue Buffon, 75231 Paris Cedex 05, France

[3] NaturalArt, Kapiteinstraat 27, B-9000 Gent, Belgium

[4] GENOSCOPE, Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry Cedex, France

[*] Corresponding author: email: holford@biology.utah.edu

Molecular Biology and Evolution, in press

Abstract

Toxoglossate gastropods (conids, terebrids, and turrids) are one of the most successful animal groups that use venom to capture their prey. These marine animals are generally characterized by a venom apparatus made up of a muscular venom bulb and venom gland. The toxoglossan radula, often compared to a hypodermeric needle for its use as a conduit to inject toxins into prey, is considered a major anatomical breakthrough that assisted in the successful initial radiation of these animals in the Cretaceous and early Tertiary. The pharmacological success of toxins from cone snails has made this group a star among biochemists and neuroscientists, but very little is known about toxins from terebrids and other Toxoglossa, and the phylogeny of these families is largely in doubt. Here we report the very first molecular phylogeny for the Terebridae and use the results to infer the evolution of the venom apparatus for this group. Our findings indicate that most of the genera of terebrids are polyphyletic, and one species, ("*Terebra*" *jungi*) is the sister group to all other terebrids. Molecular analyses combined with mapping of the venom apparatus indicates the Terebridae have lost the venom apparatus at least twice during their evolution. Species in the genera *Terebra* and *Hastula* have the typical venom apparatus found in most toxoglossate gastropods, but all other terebrid species do not. For venomous organisms, the dual analysis of molecular phylogeny and toxin function is an instructive combination for unraveling the larger questions of phylogeny and speciation. The results presented here suggest a paradigm shift in the current understanding of terebrid evolution, while presenting a roadmap for discovering novel terebrid toxins, a largely unexplored resource for biomedical research and potential therapeutic drug development.

Introduction

Marine snails are not what initially come to mind when discussing venomous animals, but the toxoglossan gastropods that use venom to capture their prey are among the most highly successful groups of marine invertebrates. Toxoglossate (“poison tongue”) gastropods include cone snails (*Conus*, within the family Conidae), auger snails or Terebridae, and "Turridae" (a complex of families) [1] [2]. A venom apparatus made up of a muscular venom bulb and venom gland generally characterizes these marine mollusks. Toxins are injected into target animals via hollow disposable radular teeth, and act to immobilize prey or defend against predators [3]. The toxoglossan radula, often compared to a hypodermeric needle, is considered a major anatomical breakthrough that assisted in the successful initial radiation of these animals in the Cretaceous and lower Tertiary [4]. There are believed to be >10,000 species of toxoglossate mollusks [5]; Toxoglossa are the most abundant group of predatory snails in terms of species numbers. The pharmacological success of toxins from cone snails has led to their extensive use in characterizing cellular communication in the nervous system, but very little is known about toxins from terebrids and other Toxoglossa, and the phylogeny of these families is largely in doubt.
The Terebridae are a distinctive example of modular anatomical development within marine gastropods. All Terebridae are carnivorous and hunt prey, however, a significant number of terebrid species have lost the venom apparatus. Based on key structural differences in their foregut anatomy, Miller identified three distinct terebrid feeding types [6, 7]: Type I terebrids, of which there are two groups, IA and IB, have salivary glands, a long eversible labial tube, and a short and slightly retractable buccal tube (Figure 1A). Type I terebrids lack a radular and venom apparatus and therefore do not use toxins to subdue and immobilize their prey. These terebrids feed by grasping the prey with the anterior part of the labial tube and engulfing the prey whole. On the other hand, Type II terebrids exhibit characteristic toxoglossan feeding in that they impale the prey with a hypodermic radular tooth through which they deliver immobilizing toxins expressed in the venom gland, then grasp and engulf their catch with the sphincter of the labial tube. In addition to the hypodermic radular teeth and venom apparatus, these terebrids have salivary glands, an eversible labial tube of moderate length, and a long retractile buccal tube (Figure 1B). The final terebrid type, Type III, is similar to Type I, but possess an unusual accessory feeding organ, accessory proboscis structure, that is mostly uncharacteristic in other terebrids (Figure 1C). Type III terebrids use the accessory feeding organ to grasp and ingest tentacles of cirratulid polychates. To understand how the three strategies evolved, it is imperative to establish a molecular phylogeny for the Terebridae.
Compared to cone snails, terebrids are not very well characterized; their taxonomy is primarily based on shell morphology [8] and relatively few studies describe their anatomy [1, 6, 7, 9-11]. This is surprising, given the relative abundance of terebrids and their global tropical distribution and shallow water occurrence. The Terebridae comprise >300 known species [10, 11], and there may be another 100 unnamed taxa, especially in deeper waters (100-500 meters).

We used a three gene data matrix consisting of mitochondrial 12S and 16S rRNA, and cytochrome oxidase subunit I (COI) sequences from 67 species to reconstruct the first molecular phylogeny of the Terebridae (Figure 2). We also assessed the evolution of terebrid feeding strategies by mapping the presence or absence of the venom apparatus on the phylogeny. The results presented here suggest a major shift in understanding terebrid evolution: our findings indicate that the Terebridae have independently lost the venom apparatus twice during their evolution.

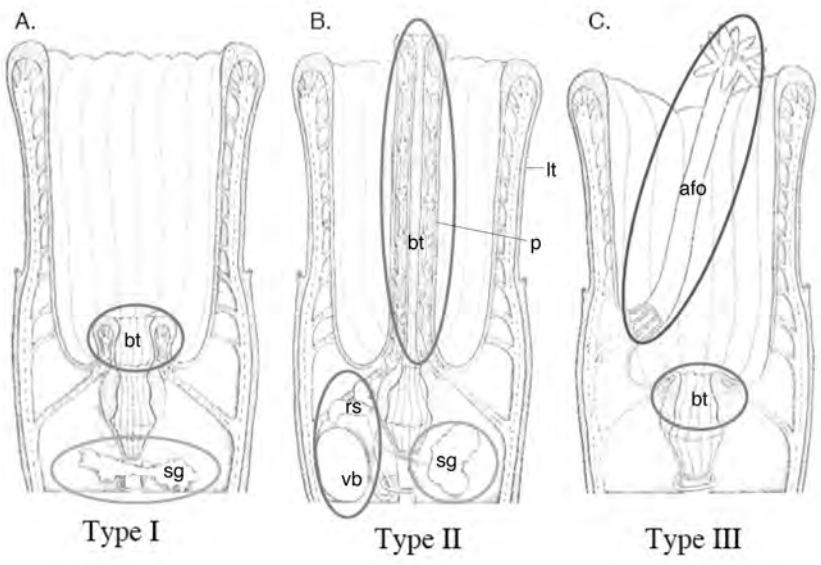


Figure 1. Schematic Representation of Miller's Foregut Anatomy of the Terebridae. The three types of foregut anatomy Miller used to define feeding strategies within the Terebridae are depicted (Types I, II, and III). Anatomical features are labeled: salivary glands (sg), buccal tube (bt), venom bulb (vb), radular sac (rs), accessory feeding organ (afo). Also labeled are the proboscis (p), found only in Type II terebrids, and the labial tube (lt). (After Miller, 1970).

The molecular phylogeny of the Terebridae additionally provides a roadmap for discovering novel terebrid toxins from those species with a venom apparatus. Given the similarities with cone snails, it is expected that the toxins from terebrids will be a rich, and at present, largely unexplored source of neuroactive peptides to investigate cellular communication in the nervous system. Reconstructing a reliable phylogeny for the Terebridae is the first step towards characterizing their toxins using the emerging biodiversity/exogenome combination strategy that has been developed for *Conus* venom peptides [12].

Results

Phylogenetic analyses

One hundred and fifty six samples of terebrids were used to reconstruct the molecular phylogeny of the Terebridae. For the COI gene, 658 bp were sequenced. After alignments, we obtained a fragment of 534 and 455 bp for 12S and 16S genes respectively. Trees obtained independently with COI, 12S, and 16S genes were only partly resolved, but no contradiction were found (results not showed). Consequently, 12S, 16S and COI mitochondrial genes, including 131 ingroups, were used to produce a combined data set for phylogenetic analyses (Figure 2).

Phylogenetic analyses strongly indicate the Terebridae is monophyletic (Posterior Probabilities (PP) = 1, Bootstraps (B) = 100), (Figure 2). "*Terebra*" *jungi* appears to be the sister group to all other terebrid (PP = 1, B = 92). Apart from "*T.*" *jungi* (Clade A), there are four major clades within the tree, designated in Figure 2 as Clades B-E. Clade B is comprised primarily of the genus *Acus* plus one species currently placed in *Terebra*, *Terebra aerolata,* which we define as *Acus aerolatus* (PP = 1, B = 100). Clade C includes *Cinguloterebra* and additional species currently placed in *Terebra,* including the type species of *Terebra*, *T. subulata* (PP = 1, B = 76). Clade D includes species currently placed in the genera *Hastula* and *Impages* (PP = 1, B =100*).*

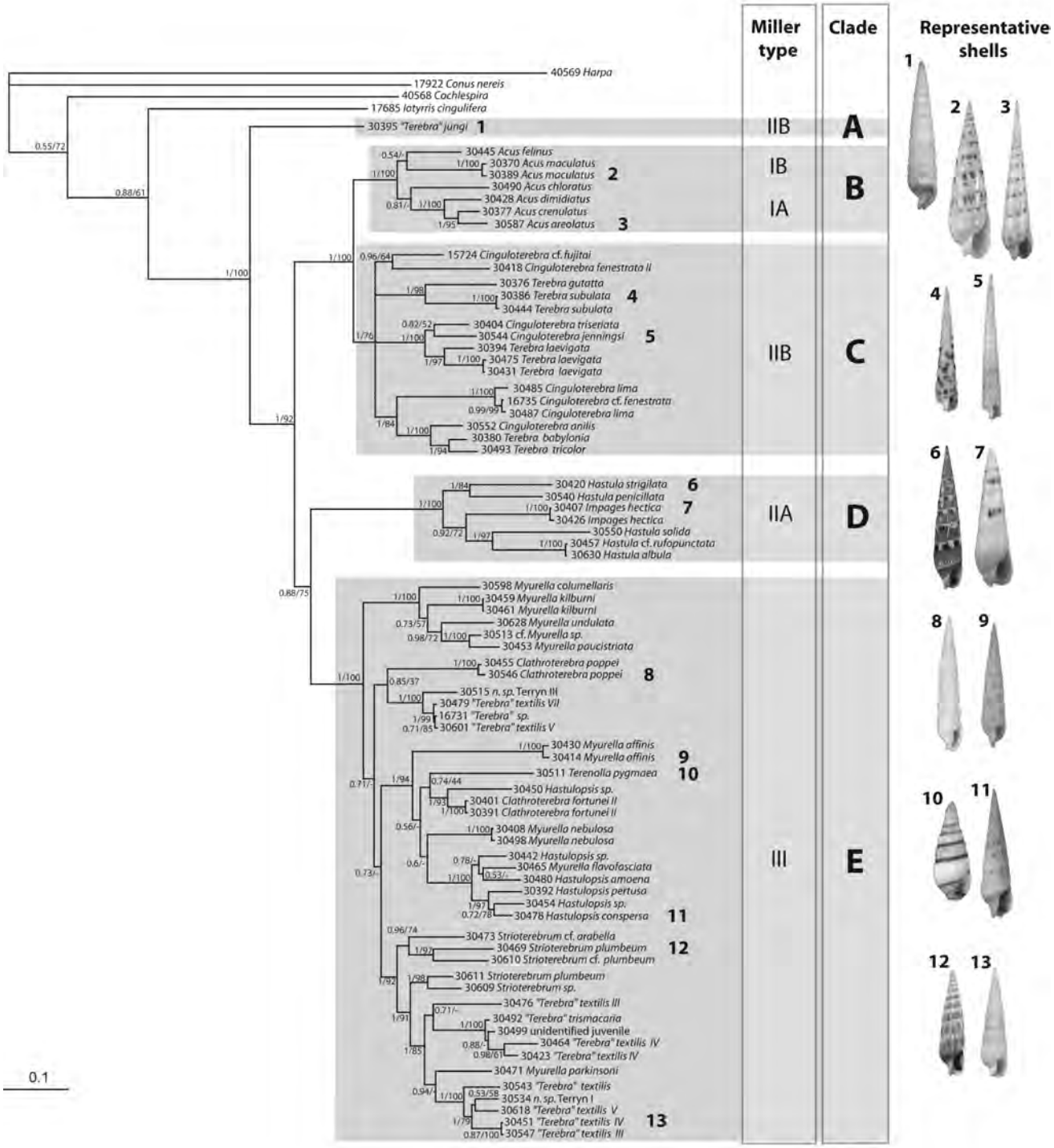


Figure 2. Combined Phylogenetic Tree. Consensus tree of ML and BA analyses using COI, 16S, and 12S dataset. Posterior Probabilities (PP) and Bootstraps (B) are specified for each node. Miller types, IA, IB, IIA, IIB, and III, as described in the text are highlighted in the tree. Molecular analyses divide the Terebridae into five distinct clades, Clades A-E, indicated by the shaded gray areas. Representative shells, numbered 1-13, are shown for each clade. For clarity, multiple samples of the same species are shown only when there is a geographic difference, e.g. 30370 and 30389-*Acus maculatus* from Panglao2004 and Santo2006 expeditions, respectively.

Clade E is the largest clade and includes the genera *Myurella, Clathroterebra, Terenolla, Hastulopsis, Strioterebrum,* and the *"Terebra" textilis-group* [11]. *Myurella* itself is polyphyletic, with several species placed as sister groups of other genera. For example, *Myurella affinis*, the type-species, is the sister group of *Terenolla*, *Hastulopsis*, some *Clathroterebra* species and two other *Myurella* species (PP = 1, B = 94). *Clathroterebra* is also polyphyletic, with the two representative species used in our analyses, *C. poppei* and *C. fortunei*, type species of *Clathroterebra*, appearing in separate well supported clades, (PP = 1, B = 100) for both. The distinctiveness of the monotypic genus *Terenolla* is confirmed in our analysis. *Hastulopsis* and *Strioterebrum* are paraphyletic. The *"Terebra" textilis-group* is dispersed within Clade E, and constitutes a group that includes a large amount of undescribed diversity, both at the genus and species levels, hence the various types (*textilis III, IV, V, VII*) listed in the tree.

Evolution of venom apparatus

Miller's type classifications are indicated in the tree in Figure 2 and Table 1. "*Terebra*" *jungi* has a venom apparatus and is identified as a Type IIB feeder (Sysoev, personal comm.). Clade B corresponds to Miller's Type I feeders, which were further subdivided between Type IA and Type IB feeders. Type IA feeders are *Acus crenulatus*, and *A. dimidiatus,* [6, 13]. *Acus areolatus*, not hitherto recognized as a species of *Acus*, is however similar to the other species in Clade B in not having a venom apparatus and was identified as a Type IA feeder by Miller [13]. The Type IB feeders in Clade B are: *A. felinus, A. maculatus, and A. chloratus*. The taxa that make up Clade C are all Type IIB feeders: *Terebra guttata, T. subulata, T. babylonia, T. tricolor, Cinguloterebra fenestrata, C. anilis*, and *C. jenningsi* (Taylor, personal comm.), [6, 9]. Clade D, made up exclusively of members of the genera *Hastula* (*H. strigilata, H. penicillata, H. solida, H. albula, H. cf. rufopunctata*) and *Impages* (*I. hectica*) are Type IIA feeders [9, 14]. Type III feeders are represented in the phylogenetic tree by Clade E, which includes species of *Myurella* (*M. columellaris, M. undulata, M. paucistriata, M. affinis, M. flavofasciata*, and *M. parkinsoni*), *Hastulopsis* (*H. conspersa* and *H. pertusa*), and *Terenolla* (*T. pygmaea*) (Taylor, personal comm.), [6, 9].

In order to efficiently characterize toxins from various terebrid species, it is essential to first identify those species that have a venom apparatus. Shown in Figure 3 is the mapping of the presence or absence of a venom apparatus in the terebrid species used to construct the molecular phylogeny in Figure 2. The map clearly indicates that terebrids have lost the venom apparatus twice during their evolution. It is inferred that the ancestor of all terebrids had a venom duct, and the groups corresponding to Clade B and Clade E have lost it. In addition to "*Terebra*" *jungi*, there are two distinct groups of terebrids, those in Clades C and D, which use the typical toxoglossate venom apparatus to hunt prey.

Discussion

The first molecular phylogeny of the Terebridae, presented in this report, significantly updates current hypotheses about the evolution of this group, and simultaneously facilitates efforts to characterize terebrid toxins. The results highlight several problematic propositions about the Terebridae, while confirming a few established hypotheses. One such hypothesis is that the Terebridae family is monophyletic. Our findings confirm that shell shape was intuitively a good character to group all terebrids in one taxon. Another substantiated hypothesis is that the monotypic genus *Terenolla* belongs in the Terebridae and is not a columbellid as suggested earlier [10]. A distinctive discovery revealed in Figure 2, is that "*Terebra*" *jungi* is a sister

group to all other terebrid species. This is an original finding unappreciated until the present molecular analysis.

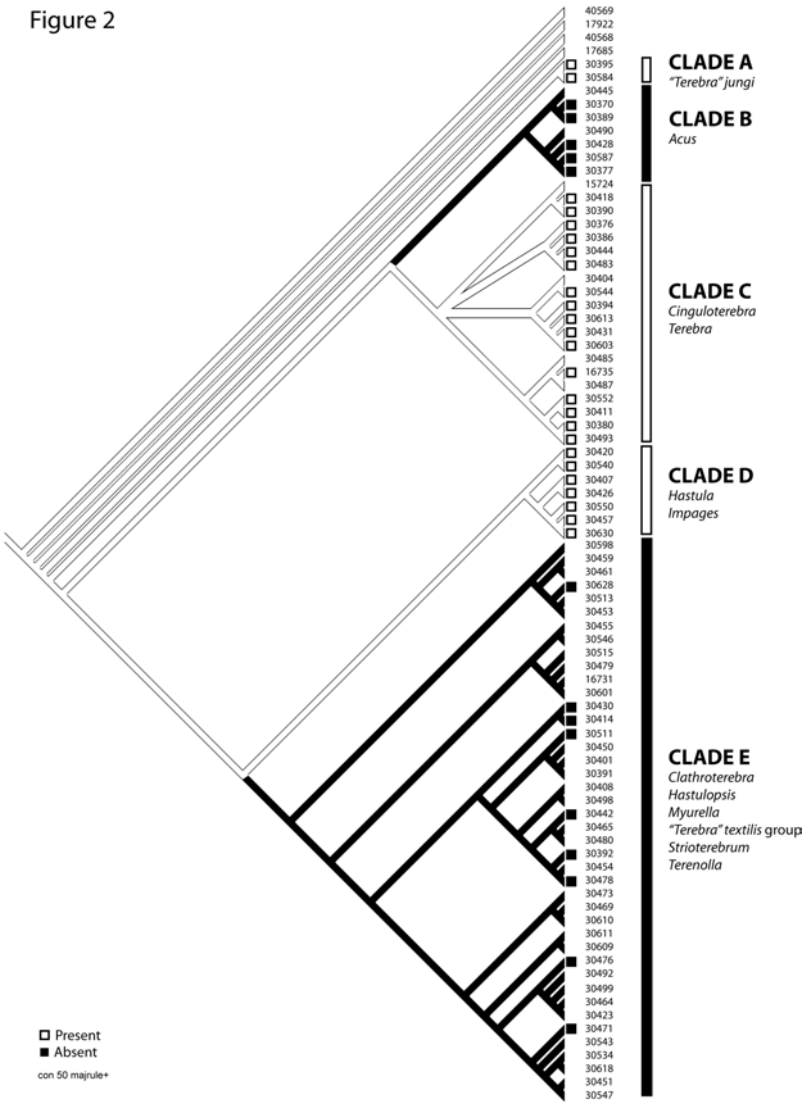


Figure 3. Possession of Venom Apparatus Mapped onto Terebrid Phylogeny. The presence or absence or a venom apparatus was mapped onto the molecular phylogeny of the Terebridae shown in Figure 2. Terebrid species with a venom apparatus are indicated by a white box, while terebrid species without a venom apparatus are indicated by a black box. The map indicates that terebrids have independently lost the venom apparatus twice during their evolution.

The prevailing implications of this phylogeny are that there is little congruence between the former genus-level classification based on shell morphology and clades recognized by molecular characters. The molecular phylogeny indicates that most of the genera proposed in the family Terebridae will have to be revised. This is evidenced by, e.g., species earlier assigned to *Terebra* now appearing throughout the tree in Clades B ("*Terebra*" *areolata*, here classified in *Acus*), C (*T. subulata*, type species of *Terebra*, and several others), E ("*Terebra*" *textilis* group), in addition to "*Terebra*" *jungi*, which will require the establishment of a new genus. Despite this non-monophyly of shell-defined genera, but based on the inclusion of their

type species (with the exception of the type species of *Acus*) in the analysis, we tentatively ascribe existing generic names to the four major clades: Clade B is thus *Acus* (H. & A. Adams, 1853), Clade C is *Terebra* (Bruguière, 1789), Clade D is *Hastula* (H. & A. Adams, 1853), and Clade E is *Myurella* (Hinds, 1844).

Evolutionary trends in the Toxoglossa have been reconstructed primarily through studies involving radular formation and anatomy of the digestive system [4, 9, 15-18]. The molecular phylogeny presented here paints a plausible picture of how terebrids evolved such a diversity of feeding strategies. Our findings suggest that all terebrids appear to be derived from a common ancestor with a venom apparatus of the Miller Type IIB (Figure 3). Furthermore, mapping of the venom apparatus indicates that two lineages of terebrids independently lost their ability to hunt prey using toxins, Clades B and E.

There is considerable correlation between our molecular phylogeny, Miller's anatomical groupings, and the ecological distribution of the terebrid species used in this study. Miller separated terebrids with a venom apparatus, Type II feeders, into two distinct groups, IIA and IIB. Type IIA and IIB terebrids differ in the shape of the buccal tube and shell morphology [6]. Type IIA terebrids have a long and slender buccal tube, small shining shells, with 7-10 whorls and a flared aperture. Terebrids of Type IIB have a short, thick buccal tube, the shells are large, long and slender, with 15 or more whorls and a constricted aperture. Our phylogenetic analyses supports this separation, as the shell shape of the species in Clade C, *Terebra*, are slender and multiwhorled, whereas those of Clade D, *Hastula*, are shiny with fewer whorls. The separation of the two clades is further supported by the ecological differences in their habitats. *Terebra* species of Clade C live buried in sandy or muddy subtidal flats, whereas *Hastula* species of Clade D live predominantly on surf beaches or in sand in reef pockets [6, 14]. Similarly, terebrids that feed without the use of a venom apparatus, Types I and III feeders, are represented by two different clades in the molecular phylogeny, Clades B and E, respectively. Our analyses goes further to identify that the non-venom apparatus feeders, Clades B and E, are not sister groups.

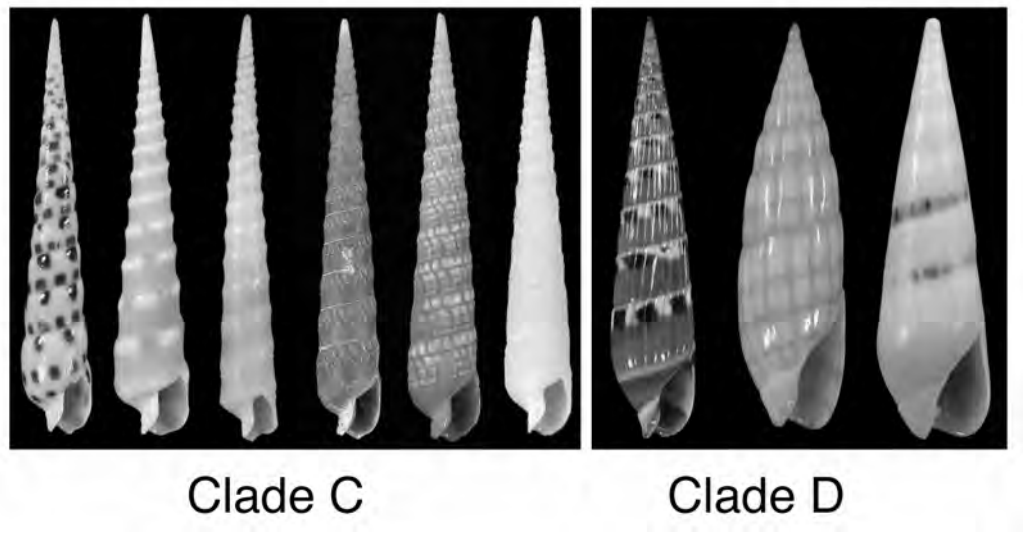


Figure 4. Terebrids with Venom Apparatus: Representative shell images of the terebrid species in Clades C and D that have a venom apparatus. The species from left to right are, Clade C: *T. subulata, T. guttata, C. jenningsi, C. anilis, T. babylonia, T. laevigata.* Clade D: *H. strigilata, H. solida,* and *I. hectica.*

The molecular phylogeny and venom apparatus map constructed for the Terebridae indirectly provides an instructive field guide for characterizing novel toxins from this group using the biodiversity-first, exogenomic strategy recently applied to cone snails [12, 19]. The exogenomic strategy was used to characterize cone snail toxins that target nicotinic receptors. In this strategy, phylogeny and molecular biology techniques are used to identify "exogenes," which are genes of the toxins expressed in the venom duct. Exogenes are expressed to be used exogenously, outside of the organism; hence, they rapidly evolve to respond to cues in their biotic environment, and are thus a powerful marker for differentiating ecological or evolutionarily distinct organisms. Clades C and D are two major terebrid groups from which it would be productive to investigate toxins for biochemical characterization (Figure 4). This detailed distinction of terebrids with and without toxins greatly facilitates efforts to characterize terebrid toxins. Furthermore, as Clades C and D are not sister clades, they may produce highly divergent toxins that could result in varied functional properties upon further characterization.

Initial reports on the toxins from terebrids [20, 21] indicate they have a comparable organization of precursor structures characteristic of conotoxin gene superfamilies, namely: a highly conserved signal sequence, followed by a propeptide region, then the mature toxin region. As has been demonstrated with cone snails, the conserved signal sequence and propeptide region in terebrids is an exploitable feature that can be used to facilitate their characterization using molecular biology techniques, e.g. as the basis for designing PCR primers. The Terebridae phylogeny in Figure 2 sets the stage for efficient characterization of terebrid toxins and identification of the gene superfamilies that encode the toxins. Preliminary results from the toxins characterized from *Terebra subulata* [20] and *Hastula hectica* [21] indicate terebrid toxins are not posttranslationally modified, facilitating the ability to chemically synthesize these peptides for testing of their function. The biochemical and genetic characterization of terebrid toxins, while identifying novel compounds useful for investigating cell communication in the nervous system, will also provide additional characters to further clarify the phylogeny and evolutionary biology of these organisms. For toxoglossa gastropods, the dual analysis of molecular phylogeny and venom function is an instructive combination for unraveling the bigger question of speciation. This work is a first attempt to address these issues for the Terebridae.

Methods

Material

All specimens used were collected during several expeditions to the West Pacific since 2004, and specifically fixed for molecular analysis. Living specimens were anesthetized using MgCl2; a piece of tissue (usually foot) was cut and fixed in 95% ethanol. Table 1 lists all specimens used in this study and the expedition sources. Vouchers are kept in the collection of the Muséum National d'Histoire Naturelle (MNHN). The species included in taxon sampling represent 12 of the 15 identified genera as defined in a recent shell morphology based revision of the Terebridae [11]. Outgroups were chosen to form a non-monophlyetic group as recommended by [22]. Representative species of Conidae and "Turridae" (s.l.) were chosen as closely related outgroups, and one species from the non-neogastropod family Harpidae was chosen as distant outgroup.

Sequencing
DNA was extracted from foot or other tissue using Qiagen QIAamp Dneasy Tissue kit. Fragments of mitochondrial genes 12S, 16S and COI were amplified using universal primers 12S1/12S3 [23], 16Sar/16Sbr [24] and LCO1490/HCO2198 [25] respectively. All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM MgCl2, 0.26 mM dNTP, 0.3 mM each primer, 5% DMSO, and 1.5 units of Qbiogene Q-Bio Taq. Amplification consisted of an initial denaturation step at 94o C for 4 min, followed by 37 cycles of denaturation at 94°C for 30 sec, annealing at 54o C for 12S gene, and 52o C for 16S, followed by extension at 72o C for 1 min. The final extension was at 72o C for 5 min. The amplification of COI genes were performed similarly, except there were two annealing cycles: the first repeated 5 times at an annealing temperature of 45o C, and the second repeated 30 times at 50o C as described by [26]. PCR products were purified using exonucleaseI and phosphatase and sequenced using the Applied Biosystem BigDyeTerminator V3.1 kit and the ABI3730XL sequencer. All genes were sequenced in both directions for increased accuracy of each sequence. Sequences were deposited in GenBank (Genbank accession numbers submitted but not yet assigned). COI sequences were also deposited in BOLD (Conoidea Barcode and Taxonomy project).

Phylogenetic analyses:
COI sequences were manually aligned and 12S and 16S were automatically aligned using ClustalW multiple alignment implemented in BioEdit version 7.0.5.3 [27]. The accuracy of automatic alignments was confirmed by visual inspection. Hyper-variable regions of 12S and 16S genes were excluded from further analyses due to ambiguities in the alignments. Nucleotide substitution models were selected for each gene separately using the program Modeltest [28], in conjunction with PAUP 4.0b10 [29]. Best model of evolution and parameters estimated for each gene are TrN + I (0.537) + G (0.885) for COI gene, TVM + I (0.310) + G (0.581) for 12S gene and GTR + I (0.402) + G (0.483) for 16S gene. Phylogenetic analyses were based on reconstructions using Maximum Likelihood (ML) and Bayesian Analysis (BA). ML heuristic search was conducted with 100 replicates with TBR branch-swapping using PhyML 2.4.4 [30], and robustness of the nodes was assessed using nonparametric bootstrapping [31] with 1000 bootstraps replicates for ML analysis.

BA consisted of six Markov chains (10,000,000 generations each with a sampling frequency of one tree each hundred generations) run in two parallel analyses using Mr. Bayes [32]. The number of swaps that are tried each time the chain stops for swapping was 4, and the chain temperature was set at 0.08. When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in.
For the combined analyses of the three genes, the same parameters were used for the ML analysis. For the BA, 1 different model was applied for each gene, each with 6 substitution categories. For the COI gene, as saturation was found on the third base of the codon, different models were applied for the 2 partitions (bases 1 and 2 vs base 3). Finally, we have four unlinked partitions (COI bases 1 and 2, COI base 3, 12S, and 16S).

Venom apparatus mapping
Data from cited literature and personal communications from A. Sysoev, and J. Taylor helped to determine the presence or absence of a venom apparatus in the species used for the phylogenetic analysis. The absence or presence of the venom apparatus was then mapped on the tree using Mesquite V. 2.01 [33], using the option “tracing character history.” The parsimony ancestral reconstruction method was used.

Table 1. List of Terebrid Samples Used in Molecular Phylogeny.

MNHN N°	Cruise	Coordinates, Depth	Genus	Species	COI	16S	12S	Miller Type	Clade
30482	BOA1	15°41.6'S, 167°02.1'E, 268-445 m	*Cinguloterebra*	*fujitai (Kuroda & Habe, 1952)*	x	x			C
30530	EBISCO	20°29'S, 158°42'E, 197-230 m		*n. sp. Terryn III*	x	x	x		E
30515	EBISCO	20°06'S, 160°23'E, 280-304 m		*n. sp. Terryn III*	x	x	x		E
30535	Panglao 04	9°33.0'N, 123°46.5'E, 8-14 m	*Hastula*	*lanceata (Linnaeus, 1767)*	x	x			D
30370	Panglao 04	9°37.4'N, 123°46.9'E, 3-20 m	*Acus*	*maculatus (Linnaeus, 1758)*	x	x	x	IB	B
30404	Panglao 04	9°35.3'N, 123°52.2'E, 84-87 m	*Cinguloterebra*	*triseriata (JE Gray, 1824)*	x	x	x		C
30464	Panglao 04	9°35.3'N, 123°52.2'E, 84-87 m	*"Terebra" textilis-group*	*textilis type IV*	x	x	x		E
30442	Panglao 04	9°33.5'N, 123°48.6'E, 80-120 m	*Hastulopsis*	*sp.*	x	x	x		E
30443	Panglao 04	9°37.4'N, 123°54.5E, 6-8 m	*Acus*	*felinus (Dillwyn, 1817)*	x	x	x	IB	B
30444	Panglao 04	9°37.4'N, 123°54.5E, 6-8 m	*Terebra*	*subulata (Linnaeus, 1767)*	x	x	x	IIB	C
30445	Panglao 04	9°37.4'N, 123°54.5E, 6-8 m	*Acus*	*felinus (Dillwyn, 1817)*	x	x	x	IB	B
30446	Panglao 04	9°37.4'N, 123°54.5'E, 4-5 m	*Terenolla*	*pygmaea (Hinds, 1844)*	x	x	x	III	E
30448	Panglao 04	9°37.4'N, 123°54.5'E, 4-5 m	*Terenolla*	*pygmaea (Hinds, 1844)*	x	x	x	III	E
30449	Panglao 04	9°37.4'N, 123°54.5'E, 4-5 m	*Terenolla*	*pygmaea (Hinds, 1844)*	x	x	x	III	E
30430	Panglao 04	9°37.4'N, 123°54.5E, 6-8 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30452	Panglao 04	9°37.4'N, 123°54.5E, 6-8 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30439	Panglao 04	9°33.4'N, 123°48.4'E, 3 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30511	Panglao 04	9°35.7'N, 123°44.4'E, 0-2 m	*Terenolla*	*pygmaea (Hinds, 1844)*	x	x	x	III	E
30510	Panglao 04	9°35.7'N, 123°44.4'E, 0-2 m	*Myurella*	*columellaris (Hinds, 1844)*	x	x		III	E
30459	Panglao 04	9°35.7'N, 123°44.4'E, 0-2 m	*Myurella*	*kilburni (RD Burch, 1965)*	x	x	x		E
30460	Panglao 04	9°35.7'N, 123°44.4'E, 0-2 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30587	Panglao 04	9°37.4'N, 123°46.9'E, 3-20 m	*Acus*	*areolatus (Link, 1807)*	x	x	x	IA	B
30472	Panglao 04	9°41.8'N, 123°51.1'E, 10 m	*Cinguloterebra*	*fenestrata type I*		x	x	IIB	C
30534	Panglao 04	9°42.1'N, 123°51.4'E, 3-4 m		*n. sp. Terryn I*	x	x	x		E
30465	Panglao 04	9°29.4'N, 123°56.0'E, 15-20 m	*Myurella*	*flavofasciata (Pilsbry, 1921)*	x	x	x	III	E
30408	Panglao 04	9°29.4'N, 123°56.0'E, 15-20 m	*Myurella*	*nebulosa (GB Sowerby I, 1825)*	x	x	x		E
30450	Panglao 04	9°43.3'N, 123°48.8'E, 123-135 m	*Hastulopsis*	*sp.*	x	x	x		E
30407	Panglao 04	9°32.8'N, 123°45.9'E, 0-2 m	*Impages*	*hectica (Linnaeus, 1758)*	x	x	x	IIA	D
30451	Panglao 04	9°36.8'N, 123°52.2'E, intertidal	*"Terebra" textilis-group*	*textilis type IV*	x	x	x		E
30431	Panglao 04	9°36.8'N, 123°52.2'E, intertidal	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30454	Panglao 04	9°36.4'N, 123°53.8'E, 60-62 m	*Hastulopsis*	*sp.*	x	x	x		E
30513	Panglao 04	9°36.4'N, 123°53.8'E, 60-62 m	*cf. Myurella*	*sp.*	x	x	x		E
30455	Panglao 04	9°36.4'N, 123°53.8'E, 60-62 m	*Clathroterebra*	*poppei (Terryn, 2003)*	x	x	x		E
30483	Panglao 04	9°32.8'N, 123°42.1'E, 3-35 m	*Terebra*	*subulata (Linnaeus, 1767)*	x	x	x	IIB	E
30481	Panglao 04	9°35.7'N, 123°44.4'E, 0-2 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30401	Panglao 05	9°39.2'N, 123°47.5'E, 255-268 m	*Clathroterebra*	*fortunei type II*	x	x	x		E
30582	Panglao 05	9°39.2'N, 123°47.5'E, 255-268 m	*"Terebra" textilis-group*	*textilis (Hinds, 1844)*	x		x		E
30418	Panglao 05	9°39.2'N, 123°47.5'E, 255-268 m	*Cinguloterebra*	*fenestrata type II*	x	x	x	IIB	C
30410	Panglao 05	9°39.2'N, 123°47.5'E, 255-268 m	*Cinguloterebra*	*fenestrata type I*	x	x	x	IIB	C
30395	Panglao 05	9°37.5'N, 123°40.2'E, 606-631 m	*"Terebra" bathyrhaphe-group*	*jungi (Lai, 2001)*	x	x	x	IIB	A
30390	Panglao 05	9°29.4'N, 123°44.4'E, 271-318 m	*Cinguloterebra*	cf. *fenestrata (Hinds, 1844)*	x	x	x	IIB	C
15724	Panglao 05	9°27.4'N, 123°49.4'E, 273-356 m	*Cinguloterebra*	cf. *fujitai (Kuroda & Habe, 1952)*	x	x	x		C
16731	Panglao 05	9°30.1'N, 123°41.6'E, 356-396 m	*"Terebra"*	*sp.*	x	x	x		E
16735	Panglao 05	9°36.2'N, 123°43.8'E, 382-434 m	*Cinguloterebra*	cf. *fenestrata (Hinds, 1844)*	x	x	x	IIB	C
30584	Panglao 05	9°34.3'N, 123°37.8'E, 729-733 m	*"Terebra" bathyrhaphe-group*	*jungi (Lai, 2001)*	x	x	x	IIB	A
30406	Panglao 05	6°24.1' S, 156°20.4' E, 1045-1207 m	*Cinguloterebra*	*sp.*	x				C
30479	Salomon 2	7°13.8' S, 158°29.4' E, 286-423 m	*"Terebra" textilis-group*	*textilis type VII*	x	x	x		E
30528	Salomon 2	7°43.5' S, 158°29.7' E, 336-341 m	*"Terebra" textilis-group*	*textilis type VII*		x	x		E
30424	Salomon 2	8°36.7' S, 157°21.0' E, 150-160 m	*"Terebra" textilis-group*	*textilis type VI*		x	x		E
30391	Salomon 2	7°59.3' S, 157°33.3' E, 260 m	*Clathroterebra*	*fortunei type II*	x	x	x		E
30501	Salomon 2	8°25.5' S, 159°26.4' E, 543-593 m	*"Terebra" bathyrhaphe-group*	*jungi (Lai, 2001)*		x	x	IIB	A
30487	Salomon 2	8°39.5' S, 157°23.0' E, 214-243 m	*Cinguloterebra*	*lima (Deshayes, 1857)*	x	x	x		C
30579	Salomon 2	8°39.5' S, 157°23.0' E, 214-243 m	*"Terebra" textilis-group*	*trismacaria (Melvill, 1917)*	x	x	x		E
30492	Salomon 2	8°39.5' S, 157°23.0' E, 214-243 m	*"Terebra" textilis-group*	*trismacaria (Melvill, 1917)*	x	x	x		E
30499	Salomon 2	8°39.5' S, 157°23.0' E, 214-243 m		Unidentified juvenile	x	x	x		E
30409	Santo 06	15°33.1'S, 167°17.8'E, 15-25 m	*Terebra*	*tricolor(GB Sowerby I, 1825)*	x	x	x	IIB	C
30389	Santo 06	15°28.7'S, 167°15.2'E, 19 m	*Acus*	*maculatus (Linnaeus, 1758)*	x	x	x	IB	B
30378	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Myurella*	*nebulosa (GB Sowerby I, 1825)*		x	x		E
30629	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Myurella*	*undulata (JE Gray, 1834)*	x	x	x	III	E
30628	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Myurella*	*undulata (JE Gray, 1834)*	x	x	x	III	E
30626	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30440	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30456	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30458	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30634	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30617	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30618	Santo 06	15°31.7'S, 167°09.7'E, 18-21 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30386	Santo 06	15°36.6'S, 167°10.1'E, 8-20 m	*Terebra*	*subulata (Linnaeus, 1767)*	x	x	x	IIB	C
30619	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Hastulopsis*	*conspersa (Hinds, 1844)*	x	x	x	III	E
30387	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Terebra*	*guttata (Röding, 1798)*	x	x	x	IIB	C
30376	Santo 06	15°33.1'S, 167°12.2'E, 3-40 m	*Terebra*	*guttata (Röding, 1798)*	x	x	x	IIB	C
30620	Santo 06	15°31.3'S, 167°10.4'E, 3-18 m	*Myurella*	*undulata (JE Gray, 1834)*	x	x	x	III	E
30494	Santo 06	15°34.4'S, 167°13.1'E, 9 m	*Acus*	*crenulatus (Linnaeus, 1758)*	x	x	x	IA	B
30377	Santo 06	15°34.4'S, 167°13.1'E, 9 m	*Acus*	*crenulatus (Linnaeus, 1758)*	x	x	x	IA	B
30623	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*conspersa (Hinds, 1844)*	x	x	x	III	E
30480	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*amoena (Deshayes, 1859)*	x	x	x		E
30463	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*conspersa (Hinds, 1844)*	x	x	x	III	E
30478	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*conspersa (Hinds, 1844)*	x	x	x	III	E
30388	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*pertusa (Born, 1778)*	x	x	x	III	E
30392	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Hastulopsis*	*pertusa (Born, 1778)*	x	x	x	III	E
30373	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Acus*	*dimidiatus (Linnaeus, 1758)*	x	x	x	IA	B
30394	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30475	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30476	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*"Terebra" textilis-group*	*textilis type III*	x	x	x		E
30621	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*"Terebra" textilis-group*	*textilis type IV*	x	x	x		E
30616	Santo 06	15°36.9'S, 167°10.5'E, 6-33 m	*"Terebra" textilis-group*	*textilis type IV*	x	x	x		E
30558	Santo 06	15°31'S, 167°09'E, intertidal	*Strioterebrum*	*plumbeum (Quoy & Gaimard, 1833)*	x	x	x		E
30469	Santo 06	15°31'S, 167°09'E, intertidal	*Strioterebrum*	*plumbeum (Quoy & Gaimard, 1833)*	x	x	x		E
30632	Santo 06	15°31'S, 167°09'E, intertidal	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30614	Santo 06	15°31'S, 167°09'E, intertidal	*Strioterebrum*	cf. *arabella (Thiele, 1925)*		x	x		E
30613	Santo 06	15°31'S, 167°09'E, intertidal	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30567	Santo 06	15°34.7'S, 167°13.8'E, 14-25 m	*Myurella*	*nebulosa (GB Sowerby I, 1825)*	x	x	x		E
30374	Santo 06	15°31.4'S, 167°09.7'E, intertidal	*Terebra*	*punctatostriata (JE Gray, 1834)*		x	x		C

30612	Santo 06	15°29'S, 167°14.94'E, 2-4 m	*Myurella*	*parkinsoni (Bratcher & Cernohorsky, 1976)*	x	x	x	III	E
30471	Santo 06	15°29'S, 167°14.94'E, 2-4 m	*Myurella*	*parkinsoni (Bratcher & Cernohorsky, 1976)*	x	x	x	III	E
30611	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Strioterebrum*	*plumbeum (Quoy & Gaimard, 1833)*	x	x	x		E
30610	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Strioterebrum*	cf. *Plumbeum (Quoy & Gaimard, 1833)*	x	x	x		E
30609	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Strioterebrum*	*sp.*	x	x	x		E
30608	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Hastula*	*strigilata (Linnaeus, 1758)*	x	x		IIA	D
30607	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Hastula*	*strigilata (Linnaeus, 1758)*		x	x	IIA	D
30420	Santo 06	15°35.2'S, 167°59.4'E,intertidal	*Hastula*	*strigilata (Linnaeus, 1758)*	x	x	x	IIA	D
30416	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Hastula*	*strigilata (Linnaeus, 1758)*		x	x	IIA	D
30414	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30412	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30598	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Myurella*	*columellaris (Hinds, 1844)*	x	x	x	III	E
30385	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*"Terebra" textilis-group*	*textilis (Hinds, 1844)*	x	x	x		E
30597	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30384	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Myurella*	*undulata (JE Gray, 1834)*	x	x	x	III	E
30383	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Terebra*	*argus (Hinds, 1844)*		x	x	IIB	C
30382	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Terebra*	*cingulifera (Lamarck. 1822)*		x	x	IIB	C
30552	Santo 06	15°35.2'S, 167°59.4'E, intertidal	*Cinguloterebra*	*anilis (Röding, 1798)*	x	x	x	IIB	C
30375	Santo 06	15°31.1'S, 167°10.5'E, 7 m	*Terebra*	*babylonia (Lamarck. 1822)*		x	x	IIB	C
30380	Santo 06	15°31.1'S, 167°10.5'E, 7 m	*Terebra*	*babylonia (Lamarck. 1822)*	x	x	x	IIB	C
30594	Santo 06	15°31.1'S, 167°10.5'E, 7 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30550	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Hastula*	*solida (Deshayes, 1857)*	x	x	x	IIA	D
30549	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Hastula*	*solida (Deshayes, 1857)*		x	x	IIA	D
30417	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Hastula*	*solida (Deshayes, 1857)*		x	x	IIA	D
30437	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Hastula*	*albula (MenkE, 1843)*	x	x		IIA	D
30438	Santo 06	15°26.6'S, 167°15.2'E, intertidal	*Hastula*	*albula (MenkE, 1843)*	x	x	x	IIA	D
30571	Santo 06	15°35.9'S, 167°01.3.1'E, 83-90 m	*Clathroterebra*	*fortunei type I?*		x	x		E
30498	Santo 06	15°27.6'S, 167°14.3'E, 6-35 m	*Myurella*	*nebulosa (GB Sowerby I, 1825)*	x	x	x		E
30453	Santo 06	15°29.6'S, 167°14..9'E, 2-5 m	*Myurella*	*paucistriata (EA Smith, 1873)*	x	x	x	III	E
30546	Santo 06	15°35.5'S, 167°02.7'E, 86-118 m	*Clathroterebra*	*poppei (Terryn, 2003)*	x	x	x		E
30544	Santo 06	15°28.6'S, 167°15.1'E, 3-31 m	*Cinguloterebra*	*jenningsi (RD Burch. 1965)*	x	x	x	IIB	C
30603	Santo 06	15°43.4'S, 167°15.0'E, 6 m	*Terebra*	*laevigata (JE Gray, 1834)*	x	x	x	IIB	C
30541	Santo 06	15°36.8S, 167°08.5'E, 1-42 m	*Myurella*	*affinis (JE Gray, 1834)*	x	x	x	III	E
30379	Santo 06	15°32.5'S, 167°10.5'E, 5-10 m	*Acus*	*dimidiatus (Linnaeus, 1758)*	x	x	x	IA	B
30372	Santo 06	15°32.5'S, 167°10.5'E, 5-10 m	*Acus*	*dimidiatus (Linnaeus, 1758)*	x	x	x	IA	B
30493	Santo 06	15°38.5'S, 167°15.1'E, 13 m	*Terebra*	*tricolor(GB Sowerby I, 1825)*	x	x	x	IIB	C
30485	Santo 06	15°32.5'S, 167°10.5'E, 5-10 m	*Cinguloterebra*	*lima (Deshayes, 1857)*	x	x	x		C
30461	Santo 06	15°42.7'S, 167°09.6'E, 2-3 m	*Myurella*	*kilburni (RD Burch. 1965)*	x	x	x		E
30601	Santo 06	15°40.7'S, 167°0.5'E, 517-614 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30570	Santo 06	15°38.1'S, 167°05.9'E, intertidal	*Myurella*	*undulata (JE Gray, 1834)*	x	x	x	III	E
30428	Santo 06	15°38.1'S, 167°05.9'E, intertidal	*Acus*	*dimidiatus (Linnaeus, 1758)*	x	x	x	IA	B
30547	Santo 06	15°31.3'S, 167°09.91'E, 1-6 m	*"Terebra" textilis-group*	*textilis type III*	x	x	x		E
30545	Santo 06	15°31.3'S, 167°09.91'E, 1-6 m	*"Terebra" textilis-group*	*textilis type III*	x	x	x		E
30426	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*Impages*	*hectica (Linnaeus, 1758)*	x	x	x	IIA	D
30543	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*"Terebra" textilis-group*	*textilis (Hinds, 1844)*	x	x	x		E
30457	Santo 06	15°22.6'S, 167°11.6'E, intertidal	*Hastula*	cf. *rufopunctata (EA Smith, 1877)*	x	x	x	IIA	D
30542	Santo 06	15°22.6'S, 167°11.6'E, intertidal	*Hastula*	*penicillata (Hinds, 1844)*	x	x	x	IIA	D
30540	Santo 06	15°22.6'S, 167°11.6'E, intertidal	*Hastula*	*penicillata (Hinds, 1844)*	x	x	x	IIA	D
30490	Santo 06	15°22.6'S, 167°11.6'E, intertidal	*Acus*	*chloratus (Lamarck. 1822)*	x	x	x	IB	B
30425	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*Strioterebrum*	cf. *arabella (Thiele, 1925)*	x	x	x		E
30473	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*Strioterebrum*	cf. *arabella (Thiele, 1925)*	x	x	x		E
30474	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*"Terebra" textilis-group*	*textilis type III*	x	x	x		E
17938	Santo 06	15°35.4'S, 166°58.7'E, 3-8 m	*"Terebra" textilis-group*	*textilis (Hinds, 1844)*	x	x	x		E
30381	Santo 06	15°35.4'S, 166°59.7'E, 3-37 m	*Acus*	*dimidiatus (Linnaeus, 1758)*	x	x	x	IA	B
30630	Santo 06	15°35.7'S, 166°59.3'E, 12 m	*Hastula*	*albula (Menke, 1843)*	x	x	x	IIA	D
30433	Santo 06	15°33.4'S, 167°12.4'E, 2-6 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30434	Santo 06	15°33.4'S, 167°12.4'E, 2-6 m	*"Terebra" textilis-group*	*textilis type V*	x	x	x		E
30624	Santo 06	15°33.4'S, 167°12.4'E, 2-6 m	*Hastulopsis*	*conspersa (Hinds, 1844)*	x	x	x	III	E
30635	Santo 06	15°33.4'S, 167°12.4'E, 2-6 m	*Myurella*	*parkinsoni (Bratcher &Cernohorsky, 1976)*	x	x	x	III	E
30423	Santo 06	15°33'S, 167°16.7'E, 92 m	*"Terebra" textilis-group*	*textilis type IV*	x	x	x		E
30403	Suva 4	18°26.4'S, 178°02.4'E, 50-51 m	*Cinguloterebra*	*cumingii (Deshayes, 1857)*			x		C
30409	Suva 4	18°26.4'S, 178°02.4'E, 50-51 m	*Cinguloterebra*	*cumingii (Deshayes, 1857)*		x	x		C

References:

1. Taylor, J.D., Kantor, Y.I. and Sysoev, A.V (1993) *Foregut anatomy, feeding mechanisms, relationships and classification of the Conoidea (=Toxoglossa) (Gastropoda).* Bull. Nat. Hist. Mus. Lond. (Zool.) **59**: p. 125-170.

2. Puillandre, N., Samadi, S., Boisselier, M.C., Sysoev, A.V., Kantor, Y. I., Cruaud, C., Couloux, A., Bouchet, P., *Starting to unravel the toxoglossan knot: molecular phylogeny of the "turrids" (Neogastropoda Conoidea).* Molecular Phylogenetics & Evolution, 2007. in press.

3. Olivera, B.M. (2002) *Conus venom peptides: reflections from the biology of clades and species.* Annu. Rev. Ecol. Syst. **33**: p. 25-42.

4. Shimek, R.L. and A.J. Kohn (1981) *Functional morphology and evolution of the toxoglossan radula.* Malacologia **20**: p. 423-438.

5. Bouchet, P., (1990) *Turrid genera and mode of development: the use and abuse of protoconch morphology.* Malacologia **32**: p. 69-77.

6. Miller, B.A. (1970) *Studies on the biology of Indo-Pacific Terebra (Ph.D. dissertation)*. University of New Hampshire: Durham, NH.

7. Miller, B.A. (1971) *Feeding mechanisms inf the family Terebridae.* Ann. Rep. Am. Mal. Union **1970**: p. 72-74.
8. Omaya, K. (1961) *On some new facts of the taxonomy of Terebridae.* Venus (Jap. Journ. Malac.) **21**(2): p. 176-189.
9. Taylor, J.D., (1990)*The anatomy of the foregut and relationships in the Terebridae.* Malacologia **32**: p. 19-34.
10. Bratcher, T. and W.O. Cernohorsky (1987) *Living Terebras of the world*. New York, NY: American Malacologists, Inc.
11. Terryn, Y. (2007) *A collectors guide to Recent Terebridae (Mollusca: Neogastropoda)*. Hackenheim, Germany: Conchbooks & NaturalArt.
12. Olivera, B.M., Teichert, RW (2007) *Diversity of the neurotoxic Conus peptides: A model for concerted pharmacological discovery.* Molecular Interventions **7**(5): p. 253-262.
13. Miller, B.A. (1975) *The biology of Terebra gouldi Deshayes, 1859, and a discussion of life history similarities among other terebrids of similar proboscis type.* Pacific Science **29** (3): p. 227-241.
14. Miller, B.A. (1979) *The Biology of Hastula inconstans (Hinds, 1844) and a discussion of life history similarities among other Hastulas of similar proboscis type.* Pacific Science **33** (3): p. 289-306.
15. Kantor, Y., Sysoev, A.V. (1989) *The morphology of toxoglossan gastropods lacking a radula, with a description of new species and genus of Turridae.* J. Moll. Stud. **55**: p. 537-549.
16. Kantor, Y., Taylor, J.D. (2000) *Formation of marginal radular teeth in Conoidea (Neogastropoda) and the evolution of the hypodermic envenomation mechanism.* J. Zool. Lond. **252**: p. 251-262.
17. Mills, P. (1979) *Radular tooth structure of three species of Terebridae (Mollusca: Toxoglossa).* Veliger **19**: p. 259-265.
18. Simone, L. (1999) *Comparative morphology and systematics of Brazilian Terebridae (Mollusca, Gastropoda, Conoidea), with descriptions of three new species.* Zoosystema **21** (2): p. 199-248.
19. Olivera, B.M. (2006) *Conus peptides: biodiversity-based discovery and exogenomics.* Journal of Biological Chemistry **281**: p. 31173-31177.
20. Imperial, J.S., Watkins, M., Chen, P., Hillyard, D.R., Cruz, L.J., Olivera, B.M. (2003) *The augertoxins: biochemical characterization of venom components from the toxoglossate gastropod Terebra subulata.* Toxicon **41**: p. 391-398.
21. Imperial, J., Kantor, Y., Watkins, M., Heralde III, F. M., Stevenson, B., Chen, P., Hansson, K., Stenfo, J., Ownby, J.P-.Bouchet, P., Olivera, B. M. (2007) *Venomous auger snail Hastula (Impages) hectica (Linnaeus, 1758): Molecular phylogeny, foregut anatomy, and comparative toxinology.* J. Experimental Zoology **308B**: p. 1-13.
22. Darlu, P., Tassy, P (1993) *La reconstruction phylogénétique. Concepts et méthodes*. Paris: Masson.
23. Simon, C., Franke, A., Martin, A. (1991) *The Polymerase Chain Reaction: DNA extraction and amplification*, in *Molecular Techniques in Taxonomy*, G. Hewitt, Johnson, AWB, Young, JPW, editors. Springer-Verlag: New York. p. 329-355.
24. Palumbi, S. (1996) *Nucleic Acids II: The Polymerase Chain Reaction*, in *Molecular Systematics*, D. Hillis, Moritz, C, Mable, BK, editors. Sinauer Associates: Sunderland, Massachusetts. p. 205-247.
25. 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.* Mol. Mar. Biol. Biotechnol. **3**: p. 294-299.
26. Herbert, P., Cywinska, A., Ball, S.L., deWaard, J.R. (2003) *Biological identifications through DNA barcodes.* Proc. R. Soc. Lond. **270**: p. 313-21.
27. Hall, T. (1999) *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* Nucl. Acids. Symp. Ser. **41**: p. 95-98.
28. Posada, D., Crandall, K.A. (2001) *Selecting Models of Nucleotide Substitution: An application to Human Immunodeficiency Virus (HIV-1).* Mol. Biol. Evol. **18**: p. 897-906.
29. Swofford, D. (2002) *PAUP,* Phylogenetic Analysis Using Parsimony* and other methods*. Sunderland, Massachusetts: Sinauer Associates.
30. Guindon, S., and Gascuel, O. (2003) *A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood.* Syst. Biol. **52**: p. 696-704.
31. Felsenstein, J., *Inferring Phylogenies*. 2004, Sunderland, Massachusetts: Sinauer Associates.

32. Huelsenbeck, J.P., R.onquist, F., Hall, B. (2001) *MrBayes: Bayesian inference of phylogeny*. Bioinformatics **17**: p. 753-755.
33. Maddison, W., Maddison, D.R. (2007) *Mesquite: A modular system for evolutionary analysis Version 2.01*. http://mesquiteproject.org

Acknowledgements:
The authors thank Yuri Kantor for processing many of the terebrid specimens in the field; Virginie Heros and Philippe Maestrati for assistance throughout, from the field to curating samples at MHNH; the staff of MNHN's "Service de Systématique Moléculaire" for technical facilities; and John Taylor and Alexander Sysoev for providing unpublished anatomical information on terebrid taxa used in this study.

Funding:
The Panglao 2004 and Santo 2006 expeditions, which were the source of many specimens, were supported among others by the Total Foundation. Joint funding from National Science Foundation (NSF) CHE and OISE postdoctoral fellowship (0610202) for MH also supported this work.

CHAPITRE II

TAXONOMIE INTEGRATIVE : CONCEPTS, CARACTERES ET METHODES

Speciation is a process of emerging genealogical distinctness, rather than a discontinuity affecting all genes simultaneously.

Mallet 2001

(Species are) more or less impossible for humans to perceive entirely simply by looking at them, as they do for cells and organisms, which is why biologists have symposia devoted to the topic of species delimitation.

De Queiroz 2007

Afin de proposer des hypothèses de délimitation d'espèces robustes au sein des Turridae, l'approche d'α-taxonomie qui sera suivie s'appuie sur un cadre conceptuel et méthodologique détaillé dans les deux premières parties de ce chapitre. Dans ce contexte, un des principaux caractères qui sera utilisé est le barcode ADN, présenté dans la troisième partie. Deux applications de barcoding seront détaillées : l'expertise taxonomique, illustrée par l'analyse du genre *Eumunida* et de pontes de gastéropodes, et la taxonomie intégrative, dont le principe est exposé à la fin de ce chapitre et illustré dans le chapitre suivant.

A. Un concept d'espèce unifié

L'espèce constitue l'un des exemples les plus frappants de concept universellement reconnu en biologie mais difficile, au moins en apparence, à définir et à utiliser. D'Aristote à Linné, et des tribus de l'Amazonie aux scientifiques de tous pays, l'espèce est considérée comme le seul niveau taxonomique comparable entre tous les groupes d'organismes, ou en tout cas le seul qui ait une véritable réalité biologique (Mayr 1982 ; De Haan *et al*. 2007 ; Begossi *et al*. 2008 ; Samadi 2008). Elle se différencie ainsi des autres rangs taxonomiques (genres, familles,…), plus ou moins arbitrairement définis (Minelli 2000 ; Bertrand & Arlin 2006 ; mais voir aussi Dubois 1988). Elle est la principale unité de décompte de la biodiversité (Gaston 2000 ; Agapow *et al*. 2004), et est utilisée dans tous les domaines de la biologie, et même, surtout dans le contexte actuel de crise écologique, dans d'autres domaines tels que l'éthique, l'économie ou la politique (avec par exemple les notions de propriétés biologiques).

Plusieurs dizaines de concepts d'espèce ont été proposées (Lherminier & Solignac, 2000), avec par exemple le concept biologique de l'espèce (BSC, Mayr 1942) ou le concept phylogénétique (PSC, Nixon & Weehler 1990), suivis de nombreux autres plus ou moins dérivés (pour une revue de ces concepts : Mayden 1997 ; Adams 1998 ; Harrison 1998 ; Agapow *et al*. 2004 ; Hey 2006 ; De Queiroz 2007). Cependant, ces différents concepts correspondent non pas à une définition, mais à des propriétés des espèces qui ne sont pas partagées par l'ensemble des organismes.

Comme proposé par plusieurs auteurs (Mayden 1997 ; Marshall 2006 ; De Queiroz 1998 ; Samadi & Barberousse 2006), il est primordial de faire la distinction entre le concept attaché au mot « espèce » et les différentes propriétés utilisées pour délimiter ces espèces dans la pratique. Il est indispensable que la définition de l'espèce s'applique à tous les organismes et qu'elle permette de les classer chacun dans une et une seule espèce. De Queiroz (1998) propose ainsi un concept unifié pour l'espèce, formalisé par la suite par Samadi & Barberousse (2006), le Lineage Species Concept (LSC) : une espèce est un réseau d'organismes, définitivement divergents des autres. En revanche, les différentes propriétés qui seront utilisées pour délimiter dans la pratique ces espèces peuvent ne pas être partagées par l'ensemble des organismes. C'est le cas des concepts proposés précédemment (BSC, PSC,…), et qui constitueront donc les propriétés utilisées pour la délimitation d'espèce dans la pratique.

La distinction entre concept et délimitation d'espèce permet d'appuyer une analyse taxonomique sur un cadre conceptuel rigoureux et d'utiliser différentes propriétés pour délimiter les espèces. Cette distinction est de plus en plus acceptée dans la communauté scientifique, et commence à être appliquée (Bichain *et al*. 2007 ; Shaffer & Thomson 2007 ; Fu & Zeng 2008 ; Giraud *et al*. 2008).

B. Délimitations des espèces

1. Critères de délimitation

Dans la pratique, la délimitation des espèces ne peut pas s'appuyer directement sur le LSC. En effet, il faudrait pour cela avoir accès à l'ensemble du réseau constitué par les organismes (cela reviendrait à connaître les parents de chaque organisme), de façon à identifier les fragments qui correspondent aux espèces. Délimiter des espèces consistera donc à inférer le réseau des organismes indirectement, et les espèces ainsi constituées seront par conséquent des hypothèses (Baum 1998 ; Godfray 2002). Pour proposer ces hypothèses, les taxonomistes se basent sur deux conséquences directes du LSC. En effet, selon le LSC, les espèces sont des communautés de reproduction, et ces communautés sont définitivement divergentes les unes des autres. De ces deux conséquences peuvent être dérivées trois catégories de critères de délimitation d'espèce, qui correspondent en fait aux méthodes classiquement utilisées en taxonomie pour délimiter des espèces :

(i) Critère de ressemblance : une communauté de reproduction partage un pool de gènes commun et les individus d'une même communauté se ressemblent donc entre eux. En analysant sur un ou plusieurs caractères l'ensemble de la variabilité d'un groupe, l'objectif sera donc de délimiter des espèces en minimisant la variabilité au sein de chacune d'entre elle. La connaissance que le taxonomiste aura acquise sur un taxon donné est indispensable pour discriminer les caractères qui permettent d'identifier les spécimens d'une même espèce de ceux qui ne le permettent pas. C'est par exemple le cas des caractères qui varient avec l'environnement, ou encore ceux qui sont impliqués dans le dimorphisme entre sexes ou entre les différents stades de développement.

(ii) Critère biologique : les individus d'une même espèce se reproduisent entre eux et échangent du matériel génétique. Ce critère peut-être testé par des expériences de croisement (Gomez *et al*. 2007), mais elles sont dans la plupart des cas difficilement réalisables. Un isolement reproducteur peut également être détecté de manière indirecte, en utilisant par exemple les outils de la génétique des populations pour estimer les flux de gènes. L'absence de recombinaison entre plusieurs loci indépendants est également une conséquence d'un isolement reproducteur (Knowlton 2000, Mallet 2001). Ce critère peut être analysé en comparant les généalogies de plusieurs gènes (« multiple genes genealogies » - Andjic *et al*. 2007 ; Arzanlou *et al*. 2008) : il y a recombinaison si un groupe d'individus est monophylétique à plusieurs loci, et que les relations entre individus au sein du groupe varient selon les gènes (« genealogical species concept » - Le Gac & Giraud, in press).

(iii) Critère phylogénétique : les individus d'une même espèce partagent une histoire commune, différente des autres espèces. Deux espèces divergentes seront donc réciproquement monophylétiques. La difficulté consiste cependant à identifier des caractères héritables qui permettront de retracer l'histoire phylogénétique des espèces. Ces caractères doivent avoir eu le temps de fixer des différences depuis l'évènement de spéciation, c'est-à-dire que le temps de coalescence du caractère doit être inférieur au temps de spéciation. Une phylogénie d'un caractère qui n'aura pas eu le temps de coalescer (parce que la spéciation est trop récente, ou parce que la taille de la population est trop importante, ou encore simplement par hasard) ne sera pas en accord avec la phylogénie des espèces (Nichols 2001).

2. Limites des critères et choix des caractères

Les informations apportées par les différents critères énoncés précédemment ne sont cependant pas toutes du même type : certains critères apportent des preuves (ce sont donc des critères suffisants), d'autres uniquement des arguments (ce ne sont donc pas des critères suffisants). Par exemple, si des échanges reproducteurs sont établis entre deux groupes d'individus, c'est la preuve qu'ils appartiennent à la même espèce. De la même façon, si deux individus sont interstériles, ils ne peuvent pas appartenir à la même espèce. En revanche, l'absence d'échange reproducteur entre deux groupes n'est pas une preuve mais seulement un argument en faveur de l'hypothèse : « ces deux groupes correspondent à deux espèces différentes ». En effet, sous certaines conditions, deux groupes qui n'échangent pas de gènes peuvent appartenir à la même espèce : par exemple, deux groupes en allopatrie, et donc qui n'échangent pas de gènes, peuvent se reproduire s'ils sont remis en contact (Beaumont *et al*. 2004 ; Bozikova *et al*. 2005 ; Gay *et al*. 2007).

De la même façon, les autres critères de délimitation d'espèces constituent des arguments et non des preuves (la ressemblance et la monophylie ne sont des conditions ni nécessaires ni suffisantes). En effet, ces deux critères peuvent être vérifiés entre deux groupes d'individus qui appartiennent à la même espèce, lorsque par exemple deux populations distantes géographiquement fixent des différences morphologiques ou moléculaires. Ainsi, Ferguson (2002) montre que généralement une distance génétique suffisamment élevée est corrélée à un isolement reproducteur, mais il cite également de nombreux contre-exemples. En allopatrie, des différences peuvent se fixer alors que la reproduction est encore possible (alors qu'en sympatrie, la fixation de différences génétiques implique un isolement reproducteur - Knowlton 2000 ; Barluenga *et al*. 2006). De plus, l'acquisition de ces différences n'est pas spontanée, et, selon le moment où l'observateur se place, les différentes propriétés seront ou ne seront pas vérifiées (Gomez *et al*. 2007). L'absence de monophylie réciproque entre deux groupes n'est pas forcément la preuve qu'ils correspondent à la même espèce, et réciproquement, la monophylie de deux groupes n'est pas la preuve que leur séparation soit définitive (Rosenberg & Tao 2008).

La mise en place de ces critères n'est donc pas systématique entre espèces différentes, et ils peuvent également se mettre en place entre deux groupes d'une même espèce. L'objectif consiste donc à identifier des caractères pour lesquels les critères sont vérifiés au niveau spécifique, et pas à un niveau inférieur (groupes au sein des espèces) ou supérieur (groupe de plusieurs espèces). Utiliser des caractères héritables permet par exemple d'éviter de regrouper des spécimens sur des ressemblances liées à l'environnement (plasticité phénotypique). L'ADN est à ce titre un caractère de plus en plus utilisé. Cependant, il faut également que le gène choisi présente plusieurs propriétés (détaillées ci-dessous) qui lui permettront de fixer des différences au niveau spécifique (et encore une fois ni au niveau populationnel, ni à un niveau supra-spécifique).

Le schéma (Figure 6) illustre le processus de coalescence de trois gènes après un évènement de spéciation. Ces gènes sont neutres, c'est-à-dire que leur évolution ne dépend que du taux de mutation et de la dérive (qui dépend elle-même de la taille efficace de la population de gènes). Le deux premiers gènes sont mitochondriaux, et diffèrent par leur taux de mutation : il est plus faible pour le second gène. A taille efficace égale, le second gène aura besoin de plus de temps pour que des différences apparaissent entre les deux espèces. Le troisième gène est nucléaire, et possède donc une taille efficace quatre fois plus élevée que les deux premiers (il est diploïde et transmis par les deux parents, contrairement aux gènes mitochondriaux qui sont haploïdes et transmis maternellement). Même si le taux de mutation est équivalent à celui du premier gène, un effet de dérive plus faible empêche les mutations de

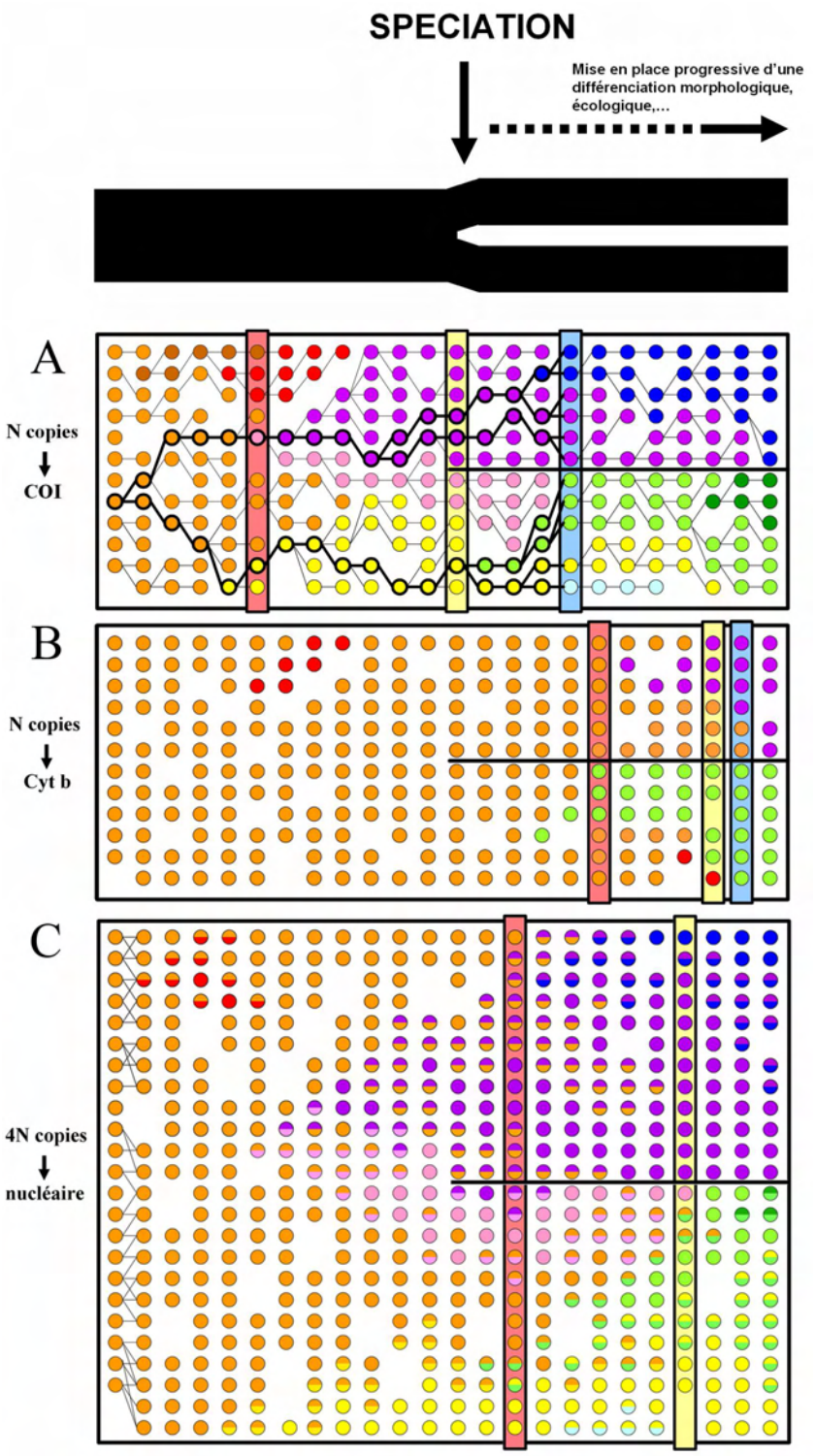


Figure 6 : Coalescence hypothétique de populations de gènes lors d'un évènement de spéciation, pour trois loci différents. A. Un gène avec N copies dans la population, et évoluant rapidement (gène COI par exemple). B. Un gène avec N copies dans la population, et évoluant plus lentement (gène cytochrome b chez les mollusques). C. Un gène avec 4N copies dans la population (gène nucléaire). Le gène 28S peut être considéré comme un intermédiaire entre les cas B et C : son taux de mutation est comparable au cas B, il est transmis par les mâles et les femelles, mais seule une copie est présente dans chaque individu, à cause des phénomènes d'évolution concertée. La barre rose marque la génération où une structuration génétique commence à être détectée ; la jaune la génération où aucun haplotype n'est partagé entre les deux espèces ; la bleue la génération où les deux espèces sont réciproquement monophylétiques pour le gène considéré. Ces critères de délimitation d'espèces ne sont pas vérifiés à la même génération selon le gène observé.

se fixer rapidement. Ces différentes caractéristiques des gènes utilisés (taux de mutation et taille efficace) influent donc sur la capacité des gènes à fixer des différences entre espèces. Les caractères choisis devront donc présenter des taux de mutation et des temps de coalescence permettant la fixation de différences entre les espèces pour permettre leur délimitation. Les caractères soumis à sélection, qu'ils soient morphologiques ou moléculaires, peuvent également fixer rapidement des différences entre les espèces.

3. Méthodologie et applications

Les différents critères de délimitation d'espèces présentent donc chacun plusieurs limites, et aucun n'est à la fois nécessaire et suffisant. Les délimitations d'espèces qu'ils permettent de proposer ne sont donc que des hypothèses. Le travail du taxonomiste consiste dans un premier temps à proposer des hypothèses primaires de délimitation d'espèces. Ces hypothèses primaires sont traditionnellement proposées après l'analyse de la variabilité morphologique. Tout autre caractère peut également être utilisé à cette étape, sachant que, comme on l'a vu, certains caractères (ceux qui sont héritables par exemple), permettront de proposer des hypothèses primaires plus robustes. Ces hypothèses peuvent ensuite être testées à l'aide d'autres caractères, comme l'ADN, ou d'autres critères, comme la monophylie, pour aboutir à des hypothèses secondaires. Ces hypothèses sont ainsi engagées dans un processus de modification/validation, au fur et à mesure que de nouveaux caractères sont analysés ou que de nouveau critères sont appliqués. L'hypothèse de délimitation d'espèce qui ne sera pas rejetée par les différents critères sera la plus robuste.

Dans le cas des Turridae, les caractères morphologiques ne permettent pas proposer des hypothèses primaires de délimitation d'espèces satisfaisantes (Introduction page 9), ou en tout cas moins satisfaisantes que celles qui pourront être proposées avec d'autres caractères. En effet, les connaissances que nous avons de l'ADN (par exemple, c'est un caractère héritable) permettent de penser que les différents caractères qui le composent, telles que les séquences nucléotidiques, constituent des caractères intéressants pour proposer des hypothèses primaires (Page *et al*. 2005). L'ADN est d'ailleurs maintenant couramment utilisé dans les analyses taxonomiques, notamment avec le développement récent des projets barcoding. Je présente dans la partie suivante de ce chapitre le principe de cette méthode et le caractère qui a été choisi comme barcode ADN universel (le gène COI). J'insiste notamment sur les intérêts que présente ce caractère en taxonomie en le replaçant dans le contexte conceptuel et méthodologique que je viens de décrire dans cette première partie du chapitre. Je développerai ensuite deux applications du barcoding : expertise taxonomique, avec les exemples du genre *Eumunida* et des pontes de gastéropodes collectées aux Philippines, et proposition de nouvelles hypothèses de délimitation d'espèces. Cette deuxième application sera illustrée dans le chapitre 3 par différentes études réalisées sur différents groupes au sein des Turridae.

C. Barcoding

1. Le principe du barcoding

Un barcode ADN est un fragment d'ADN utilisé pour identifier l'organisme porteur de ce fragment. Cette identification repose sur le fait que la séquence ADN d'un individu ressemble plus à celle des autres individus de son espèce qu'à celle des individus d'autres espèces. Si on dispose d'une base de données qui contient des séquences ADN de toutes les espèces vivantes sur Terre, il suffit donc de trouver quelle espèce possède la séquence la plus proche de l'individu inconnu pour pouvoir l'identifier (Figure 7). Un barcode ADN remplit donc le rôle d'un code-barre, comme celui d'un produit manufacturé. En effet, le système de code-barre mis en place sur ces produits (Universal Product Code) permet, par simple lecture du code, de connaître les informations liées au produit (son prix, sa date de fabrication,…). De la même façon, la lecture du barcode ADN d'un organisme permet son identification taxonomique.

Comme dans le cas d'un produit manufacturé, où les informations du produit sont regroupées à l'avance dans une base de données, l'identification taxonomique se fait en comparant la séquence inconnue avec une base de données. Cette base de données contient en effet des séquences d'ADN issues de spécimens qui ont été associés à un nom d'espèce par un taxonomiste. Par conséquent, dans cette base de données de référence, le lien entre le spécimen et le barcode ADN est univoque et indiscutable, mais le lien entre le nom et la séquence d'ADN est une identification fournit par un taxonomiste, identification qui doit pouvoir être vérifiée. Elle ne peut l'être que par l'observation du spécimen en question (Ebach & Holdredge 2005, Hajibabaei *et al*. 2005 ; Savolainen *et al*. 2005, Ward *et al*. 2005). Ce spécimen, ou voucher, constitue donc l'individu de référence auquel est attaché la séquence d'ADN et toutes les autres informations liées au spécimen et à la séquence. Son intérêt est primordial, notamment pour réexaminer les hypothèses d'identification proposées, mais aussi lorsque l'analyse de la séquence remet en question les limites de l'espèce qui y est associée (Hebert *et al*. 2004a / Brower 2006).

La notion de voucher peut-être de plus généralisée à toutes les informations liées au spécimen physique : photographies, ADN matriciel, informations liées à la collecte, prélèvements de tissus… Ces vouchers physiques et électroniques génèrent une quantité d'informations qu'il n'est possible de gérer que grâce à la mise en place de bases de données adaptées (Savolainen *et al*. 2005). BOLD (Barcode Of Life Database, Hajibabaei *et al*. 2006 ; Ratnasingham & Hebert 2007) a été développée dans cet objectif, et constitue la première base de données dédiée à l'identification de spécimen par la méthode barcoding.

L'utilisation d'un marqueur moléculaire en expertise taxonomique n'est pas une idée nouvelle (voir par exemple Bradley & Baker 2001 ; Roca *et al*. 2001 ; Floyd *et al*. 2002 ; Tautz *et al*. 2003 ; Moritz & Cicero 2004 ; Will & Rubinoff 2004 ; Will *et al*. 2005 ; Rubinoff 2006). Dès les années 70, les techniques d'analyse moléculaire sont utilisées en taxonomie (Ward *et al*. 2005). Le mot « barcode » a d'ailleurs déjà été utilisé en systématique (Arnot *et al*. 1993 ; Rash & Gusfield 2002). La nouveauté du principe proposé par Hebert repose principalement sur deux points :

- la standardisation de la méthode : l'analyse d'un seul et même gène pour tous les organismes (« Horizontal Genomics », Yassin 2007) permet l'acquisition d'un savoir-faire qui réduit la durée et le coût nécessaires pour l'obtention d'un barcode à partir d'un spécimen (Godfray 2007). Il existe d'ailleurs actuellement un projet de réalisation d'un barcodeur portable (Janzen 2004 ; Janzen *et al*. 2005), de la taille d'un téléphone cellulaire, et qui devrait inclure

à terme l'ensemble des étapes nécessaires au séquençage d'un spécimen (extraction d'ADN, amplification par PCR et séquençage en lui-même). Par communication avec une base de données de référence, comme BOLD par exemple, n'importe quel utilisateur pourrait identifier un organisme.

- le séquençage de l'ensemble des organismes vivants : la grande majorité des projets barcoding en cours ont en effet pour but de « barcoder » à terme l'ensemble des organismes vivants sur Terre. Notamment, le projet I-Bol (http://www.dnabarcoding.org) planifie de séquencer 5 millions de spécimens, correspondant à 500000 espèces, au cours des cinq prochaines années. L'objectif affiché de ces projets est de décrire l'ensemble de la diversité sur Terre. Cependant, à l'instar du séquençage du génome humain, ils fournissent également un ensemble de données dont l'analyse est le point de départ à de nombreuses études plus approfondies.

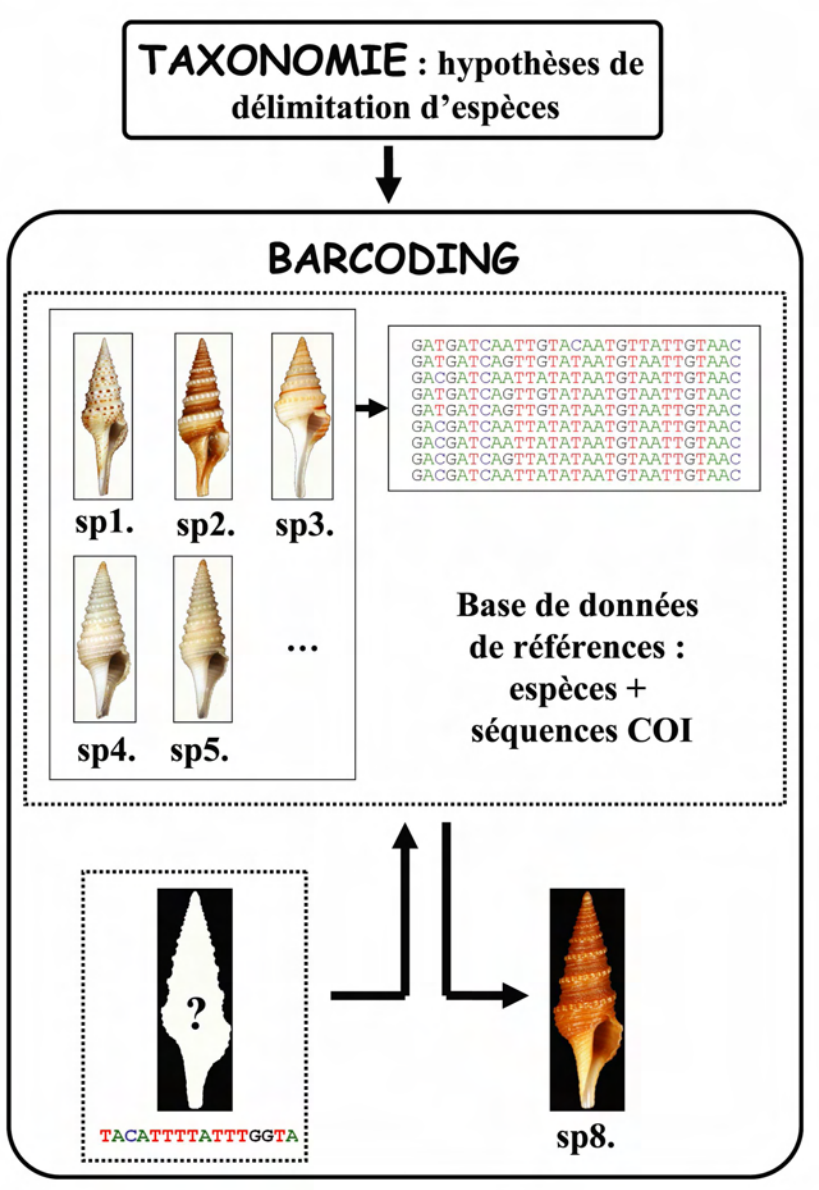


Figure 7 : Principe du barcoding comme outil d'identification. Pour chaque espèce délimitée au préalable par une analyse α-taxonomique, plusieurs spécimens sont séquencés pour le gène COI. L'ensemble des associations entre les séquences et les noms d'espèces constitue la base de données de référence. Un spécimen inconnu peut ensuite être identifié en comparant sa séquence d'ADN à la base de données de barcodes.

2. Le choix du gène COI

Hebert *et al*. (2003a, b) proposent d'utiliser un fragment de 650pb (5' terminal) du gène mitochondrial codant pour la première sous-unité du cytochrome oxydase, ou COI, comme barcode universel. Dans la plupart des groupes où il a été testé (Tableau 1), le gène COI possède les propriétés nécessaires à un barcode ADN, avec un niveau de variabilité faible au sein des espèces et élevé entre espèces différentes. Il offre de plus de nombreux avantages techniques : il est présent en plusieurs copies dans chaque cellule et il est facilement amplifiable, d'abord parce que la taille du fragment choisi est relativement courte, et ensuite parce qu'il existe des zones conservées en amont et en aval du fragment qui permettent de définir facilement des amorces permettant d'amplifier le gène par PCR chez une grande diversité d'organismes (amorces proposées par Folmer *et al*. 1994 et fréquemment utilisées – Tableau 1). De plus, ce gène était déjà utilisé couramment, notamment pour des analyses phylogénétiques ou phylogéographiques, et de nombreuses séquences sont déjà disponibles dans les bases de données classiques (comme Genbank). Enfin, contrairement aux caractères morphologiques traditionnellement utilisés pour identifier un spécimen, l'ADN possède la propriété d'être identique à tous les stades de vie, quel que soit le sexe et dans toutes les parties de l'organisme (sauf dans quelques cas, comme pour les mutations somatiques) (Stoeckle 2003 ; Blaxter 2004 ; Janzen *et al*. 2005). Il peut également être séquencé à partir de spécimens anciens, tels que ceux préservés depuis plusieurs décennies dans les muséums (Fang *et al*. 2002 ; Hajibabaei *et al*. 2005, 2006 ; Skage & Schander 2007).

Pour un caractère morphologique, l'expérience du taxonomiste lui apporte des connaissances sur son déterminisme et ses modalités d'évolution, et lui permet donc de savoir s'il peut être utilisé pour proposer des hypothèses de délimitation d'espèces robustes. A l'inverse, le barcode ADN, en particulier le gène COI, est un caractère héritable. Il permet donc de limiter l'expérience taxonomique nécessaire à l'identification d'un spécimen. De plus, le gène COI est un gène codant impliqué dans la chaîne respiratoire des mitochondries et présente à la fois : (i) une faible variabilité sur les deux premières bases du codon due à la sélection liée à sa fonction, et qui permet l'alignement des séquences indispensable à l'analyse phylogénétique ; (ii) un taux de mutation élevé sur la troisième base du codon (due à la dégénérescence du code génétique), qui, associé à un temps de coalescence rapide (Figure 6 page 55), permet l'accumulation rapide de différences entre espèces. Ces caractéristiques correspondent aux propriétés que doit posséder un caractère pour reconnaître des espèces selon les critères de ressemblance, biologique ou phylogénétique (héritabilité, fort taux de mutation, temps de coalescence court – B. page 54) : le gène COI consitue donc un caractère intéressant en taxonomie.

3. Limites du barcode COI

Cependant, principalement trois limites ont été avancées quant à l'utilisation du gène COI en taxonomie, que ce soit comme barcode universel pour l'identification de spécimens ou pour proposer des hypothèses de délimitation d'espèces (voir ci-dessous : *D. applications du barcoding*). Tout d'abord, plusieurs auteurs ont proposé que des copies nucléaires du gène COI (NUMTs : NUclear MiTochondrial sequences ; Bensasson *et al*. 2000) pourraient être séquencées par erreur à la place du COI mitochondrial (Thalmann *et al*. 2004 ; Moritz & Cicero 2004 ; Lorenz *et al*. 2005). Ces copies nucléaires perdent en général rapidement leur fonction (Ohno 1970), et sont donc rapidement modifiées, y compris sur les zones où les amorces utilisées pour amplifier le gène COI pourraient se fixer. Ces copies nucléaires, reconnaissables notamment par la présence de codons-stop dans la séquence, n'ont jamais été

détectées dans le cadre d'une étude barcoding, notamment chez les Gastéropodes (Remigio & Hebert 2003).

	Domaine/Taxon	Références
1	Agronomie	Armstrong & Ball 2005
	Biologie de la conservation	Blaxter 2003 ; De Ley *et al*. 2005 ; Rubinoff 2006
	Biomédecine	Besansky *et al*. 2003
	Contenus stomacaux	Bourlat *et al*. 2008
	Espèces invasives	Siddall & Budinoff 2005
	Sang pompé par les moustiques, oiseaux écrasés contre les avions	Hebert *et al*. 2004b
	Trafic de viande ou d'animaux	Lorenz *et al*. 2005 ; Smith *et al*. 2008b
2	Algues rouges	Saunders 2005
	Araignées	Barrett & Hebert 2005
	Autres insectes	Hogg & Hebert 2004 ; Whiteman *et al*. 2004 ; Ball *et al*. 2005 ; Monaghan *et al*. 2005 ; Smith *et al*. 2005 ; Scheffer *et al*. 2006 ; Smith *et al*. 2006 ; Rach *et al*. 2008 ; Stahls & Savolainen 2008
	Bryozaires	Gomez *et al*. 2007
	Chondrychtiens et Actinoptérygiens	Ward *et al*. 2005 ; Spies *et al*. 2006 ; Smith *et al*. 2008b ; Ward *et al*. 2008
	Crustacés	Costa *et al*. 2007
	Lézards	Marshall 2006
	Mammifères	Lorenz *et al*. 2005 ; Clare *et al*. 2006
	Méiofaune du sol	Powers 2004 ; De Ley *et al*. 2005
	Microorganismes	Markmann & Tautz 2005 ; Summerbell *et al*. 2005 ; Maslov *et al*. 2007 ; Scicluna *et al*. 2006 ; Evans *et al*. 2007 ; Seifert *et al*. 2007
	Mollusques	Kelly *et al*. 2007 ; Mikkelsen *et al*. 2007, Campbell *et al*. 2008
	Oiseaux	Lambert *et al*. 2005 ; Hebert *et al*. 2004b ; Kerr *et al*. 2007
	Papillons	Hebert *et al*. 2004a ; Janzen *et al*. 2005 ; Brower 2006 ; Hajibabaei *et al*. 2006 ; van Velzen *et al*. 2007
3	Amphibiens	Vences *et al*. 2005, Smith *et al*. 2008a
	Autres insectes	Trewick 2008
	Cnidaires	Hebert *et al*. 2003b ; Fukami 2004 ; Erpenbeck *et al*. 2005 ; Flot & Tillier 2007 ; Shearer & Coffroth 2008
	Diptères	Meier *et al*. 2006
	Gastéropodes	Meyer & Paulay 2005
	Lépidoptères	Elias *et al*. 2007 ; Wiemers & Fiedler 2007
	Plantes	Chase *et al*. 2005 ; Kress *et al*. 2005
	Rotifères	Blaxter *et al*. 2005
4	Champignons	Seifert *et al*. 2007
	Cnidaires	Moura *et al*. 2008
	Plantes	Newmaster *et al*. 2006 ; Rubinoff *et al*. 2006 ; Kress & Erickson 2007 ; Lahaye *et al*. 2008

Tableau 1 : liste non exhaustive des articles comparant la variabilité intra et interspécifique du barcode COI et contribuant ainsi à la création des bases de données de référence. Les articles sont classés en quatre catégories : 1. analyses limitées à un domaine de la biologie ou à une question appliquée ; 2. analyses limitées à un taxon, avec en général une variabilité intraspécifique inférieure à la variabilité interspécifique ; 3. analyses limitées à un taxon, avec des variabilités intraspécifiques et interspécifiques chevauchantes ; 4. recherche d'un barcode autre que le gène COI pour certains des taxa analysés en 3. Les problèmes rencontrés dans les articles de la catégorie 3 sont liés soit à un niveau de variabilité du gène COI non adapté (variabilité intraspécifique très importante - cas des amphibiens), soit à une remise en cause des hypothèses de délimitation d'espèces et qui ne permettent pas l'identification des variabilités intra et interspécifiques (Hogg & Hebert 2004 ; Meyer & Paulay 2005).

Ensuite, le gène COI ne respecte pas dans certains groupes les conditions requises pour être utilisé comme un barcode, c'est-à-dire que la variabilité intraspécifique n'est pas inférieure à la variabilité interspécifique dans des cas où les délimitations d'espèces sont considérées comme robustes. Chez les amphibiens, la variabilité intraspécifique est très importante et comparable à la diversité interspécifique. Inversement, chez les Cnidaires, la variabilité intraspécifique est très faible (Tableau 1).

Enfin, le mode de transmission, généralement maternel, des gènes mitochondriaux peut également entraîner des biais. Si les femelles ne dispersent pas entre deux populations

d'une même espèce, ces deux populations seront considérées comme appartenant à deux espèces différentes si seul le gène COI (ou un autre gène mitochondrial) est analysé. Cette différenciation génétique entre ces deux populations disparaîtra avec un gène nucléaire, les migrations des mâles permettant le mélange des copies nucléaires entre les deux populations (Moritz & Cicero 2004 ; Fu & Zeng 2008). Plus généralement, l'arbre phylogénétique du gène COI (ou de tout autre gène mitochondrial – Fu & Zeng 2008) peut être différent de l'arbre phylogénétique des espèces (Nichols 2001 ; Funk & Omland 2003 ; Degnan & Rosenberg 2006 ; Rosenberg & Tao 2008). Il peut notamment présenter du polymorphisme ancestral dans le cas des espèces jeunes, ou tout simplement par hasard (dû à la variance dans les temps de coalescence) (Mallet & Willmott 2003 ; Moritz & Cicero 2004 ; Fu & Zeng 2008) : le gène observé n'aura pas eu le temps de coalescer depuis l'évènement de spéciation (incomplete lineage sorting - Gompert *et al*. 2006 ; Salter Kubatko & Degnan 2007). Ce phénomène semble atténué pour le gène COI, un gène mitochondrial coalescant en moyenne quatre fois plus vite qu'un gène nucléaire (Zink & Barrowclough 2008 ; Figure 6 page 55). De plus, la sélection qui s'applique sur le génome mitochondrial peut accélérer la fixation des mutations accumulées sur le gène COI (Bazin *et al*. 2006).

D. Applications du barcoding

1. Expertise taxonomique

Le barcoding est avant tout un outil d'identification taxonomique (Lipscomb *et al*. 2003 ; Mallet & Willmott 2003 ; Seberg *et al*. 2003 ; Moritz & Cicero 2004 ; Ebach & Holdredge 2005 ; Wheeler 2005 ; Godfray 2006). Il vient à ce titre s'ajouter aux autres outils d'expertise taxonomique, que sont par exemple les clés d'identification basées sur des caractères morphologiques. Les caractères utilisés dans une clé de détermination n'ont pas forcément participé à l'élaboration d'hypothèses de délimitations. Ils sont choisis *a posteriori*, quand ces hypothèses sont établies, selon le critère suivant : ils doivent permettre de replacer un spécimen inconnu dans l'espèce à laquelle il appartient. De la même façon, à partir des hypothèses de délimitation d'espèces qui ont été proposées au préalable, la variabilité du barcode ADN au sein et entre les espèces peut être évaluée pour vérifier que chaque espèce est caractérisée par une ou plusieurs séquences d'ADN qui lui sont propres.

La première étape pour mettre en place un outil barcoding pour l'expertise taxonomique consiste donc à séquencer plusieurs spécimens, représentatifs de chacune des espèces délimitées, pour constituer la base de données de référence. Si le gène choisi est diagnostique et permet de discriminer chacune des espèces, il permettra d'assigner un individu à une espèce en utilisant sa séquence d'ADN. Les méthodes qui permettent l'assignation d'une séquence inconnue à une espèce peuvent se baser sur les différents critères, en se basant soit sur l'utilisation d'un seuil de distances génétiques (règle des 10X – Hebert *et al*. 2004b ; MOTU_Define – Blaxter *et al*. 2005 ; DNA-Surveillance – Baker *et al*. 2003 ; TaxI – Steinke *et al*. 2005 ; BOLD-IDS – Ratnasingham & Hebert 2007), soit sur l'utilisation de la théorie de la coalescence (Matz & Nielsen 2005 ; Abdo & Golding 2007), soit sur le partage d'apomorphies (De Salle *et al*. 2005 ; Rach *et al*. 2008) ou encore sur une méthodologie dérivée de la théorie des réseaux neuronaux (Zhang *et al*. 2008). Les méthodes basées sur l'utilisation de seuils de distances génétiques restent critiquées, notamment parce que ce seuil peut varier fortement d'un taxon à l'autre et donc difficilement généralisable (Ferguson 2002 ; Holland *et al*. 2004 ; Gomez *et al*. 2007 ; Bichain 2007). Cependant, plusieurs méthodes, notamment basées sur une approche phylogénétique, permettent d'obtenir de bons scores d'assignement (supérieurs à 95 %, voire 99 % pour certaines – Little & Stevenson 2007), et les problèmes subsistants sont plutôt liés à la qualité de l'échantillonnage qu'aux méthodes employées (Ross *et al*. 2008).

Les deux exemples qui sont présentés ci-dessous permettent d'illustrer deux aspects de l'application du barcoding comme outil d'expertise taxonomique : constitution de la base de données et identification de spécimens. Tout d'abord, une base de données de référence a été constituée pour le genre *Eumunida*. Elle permet de vérifier que dans la plupart des cas le barcode COI est un caractère diagnostic, mais aussi de modifier dans deux cas les hypothèses de délimitation d'espèces qui étaient disponibles pour ce genre. Ensuite, les bases de données déjà disponibles (BOLD et GenBank), ont été utilisées pour identifier des pontes de gastéropodes marins. Deux méthodes ont été appliquées, l'une basée sur la similarité entre séquences (mais qui ne permet que rarement d'identifier les spécimens), et l'autre sur une analyse phylogénétique, et qui permet de replacer le spécimen dans une famille, une sous-famille ou un genre quand il n'a pas pu être identifié au niveau spécifique.

2. Le cas du genre Eumunida

J'ai participé à l'analyse d'un groupe de crustacés décapodes marins, le genre *Eumunida*, en collaboration avec Sarah Samadi, Josie Lambourdière, Marie-Catherine Boisselier et Enrique McPherson (Article 3). Pour ce genre, des hypothèses de délimitation d'espèces basées sur des caractères morphologiques, auxquelles ont été associés des noms, étaient déjà disponibles (pour la plupart proposées par E. MacPherson, taxonomiste spécialiste du groupe). De plus, quand un taxonomiste propose une hypothèse d'espèce associée à un nom, il désigne un (l'holotype) ou plusieurs (les paratypes) spécimens comme porte-noms. Même si les hypothèses de délimitation d'espèces changent, les noms d'espèces resteront associés à ces spécimens. Nous avons séquencé les spécimens-types du genre *Eumunida* conservés dans les collections du museum. Ce séquençage a été possible parce que ces collections sont conservées en alcool depuis leur collecte. De cette façon, les séquences d'ADN de ces spécimens pourront être associées aux noms d'espèces disponibles. Plusieurs autres séquences de spécimens non types, récoltés lors des campagnes MUSORSTOM dans le Pacifique, ont également été inclues à l'analyse. Plus de 230 spécimens ont été séquencés pour le gène COI, incluant 9 holotypes et 24 paratypes représentant 13 espèces. En suivant les descriptions morphologiques disponibles dans la littérature, les spécimens non types ont été identifiés comme appartenant à 17 espèces différentes. Plusieurs spécimens correspondant à 9 espèces ont également été séquencés pour le gène 28S. L'ajout d'un outgroup à également permis d'analyser les relations phylogénétiques au sein du genre.

Dans la plupart des cas, les groupes définis génétiquement sont parfaitement congruents avec les hypothèses morphologiques : chaque groupe n'inclut que des spécimens qui ont reçu la même identification spécifique, ainsi que, pour la plupart, un ou plusieurs-spécimens-types qui permettent de lier un nom au groupe correspondant. Cependant, dans deux cas les résultats ne sont pas congruents :
- deux groupes génétiques ont été identifiés morphologiquement comme *E. annulosa*. De plus, ces deux groupes ne sont pas groupes-frères. Seul un caractère morphologique, très ténu, permet de différencier ces deux groupes, mais ils ne s'appliquent qu'aux spécimens adultes. Ces deux groupes correspondent donc très certainement à des espèces cryptiques (Knowlton 2000 ; Sáez & Lozano 2005). Les hypothèses taxonomiques correspondant au nom *E. annulosa* ont donc changé : ce qui était considéré comme une seule espèce correspond maintenant à deux espèces diférentes. Cependant, le lien entre le nom E. annulosa et l'holotype est conservé : son inclusion dans l'une de ces deux nouvelles espèces permet de la lier au nom ; la seconde devra être décrite sous un nouveau nom.
- trois noms d'espèces, représentées par leurs types (*E. karubar*, *smithii* et *parva*), sont inclus dans un seul et même groupe génétique, et devront donc vraisemblablement être mis en synonymie.

Cette analyse a permis de vérifier que le gène COI était un bon marqueur diagnostic au sein du genre *Eumunida*, et qu'il pouvait donc être utilisé pour l'identification de spécimen. Quelques hypothèses taxonomiques ont cependant dû être modifiées, permettant ainsi de constituer une base de données de référence incluant des hypothèses de délimitation d'espèces robustes.

DNA barcode, type specimens and species delimitation in the genus *Eumunida*

Authors: Sarah SAMADI[1], Nicolas PUILLANDRE[1], Enrique MACPHERSON[2], Josie LAMBOURDIERE[3], Corinne CRUAUD[4], Arnaud COULOUX[4,] Marie-Catherine BOISSELIER-DUBAYLE[1].

[1]"Systématique, Adaptation et Evolution", UMR 7138 UPMC-IRD-MNHN-CNRS (UR IRD 148), Muséum National d'Histoire Naturelle, Département Systématique et Evolution, CP 26, 57 Rue Cuvier, F-75231 Paris Cedex 05, France.
[2] Centro de Estudios Avanzados de Blanes (CSIC), C. acc. Cala Sant Francesc 14, 17300 Blanes, Girona, Spain
[3] Service de systématique moléculaire (CNRS, IFR101), Muséum National d'Histoire Naturelle, Département Systématique et Evolution, CP 26, 57 Rue Cuvier, F-75231 Paris Cedex 05, France.
[4] GENOSCOPE, Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry Cedex France.

Corresponding author: S. Samadi. UMR 7138 CNRS-IRD-MNHN-UMPC, Muséum National d'Histoire Naturelle (USM 603), Département Systématique et Evolution, Case postale 26, 57 Rue Cuvier, F-75231 PARIS CEDEX 05, *e-mail :* sarah@mnhn.fr ; *tel :* 33 1 40 79 37 59; *fax :* 33 140 79 38 44

Soumis à Biological Journal of the Linnean Society

Running title: DNA barcode, type specimens & species delimitation

Abstract

Achieving the primary purpose of DNA barcode projects – generating an efficient expertise tool – is an important challenge to the taxonomy of the 21st century. Since, the expertise lies on the hypotheses provided by researchers, new expertise tools must be directly linked to research activities. As a consequence the constitution of DNA barcode datasets must follow the requirements of specimens sampling of the taxonomic research, including naming practices. In this perspective, it is useful to clarify the conceptual and methodological framework of the taxonomic research. With a case study on the decapod genus *Eumunida* we illustrate that in taxonomy, research and expertise are the two faces of the same coin. Our sampling covers a large part of the described species of this genus with numerous specimens from several locations but also type specimens. Genetic clusters are identified using two unlinked genes (COI and 28S rDNA). The inclusion of type specimens in the dataset offers an unambiguous link between a genetic cluster and a taxonomic name. The distribution of the states of the morphological characters used in the determination key over the identified genetic clusters is examined. This analysis reveals a cryptic species, morphological characters that are polymorphic within genetic clusters, and other characters considered as diagnostic at the species level, that differ between adult and juvenile specimens within a genetic cluster. Thus our analyses enhance the alpha-taxonomy of the genus and provide a very efficient expertise tool. This case study highlights that when a DNA-barcode dataset is constituted following the procedures of an integrative taxonomic research, the resulting expertise tool is very effective: it do not only permit to propose species names to newly examined specimens but also to suggest, thanks to the data from these newly examined specimens, the need of taxonomic revisions.

Key-words: DNA-barcode, Decapoda, integrative taxonomy, species concept, taxonomic expertise.

Introduction

The primary purpose of DNA barcode projects is usually not to produce new descriptions but to facilitate taxonomic expertise by developing a global standard for the identification of biological species based on molecular data (Schindel & Miller 2005; Hebert & Gregory 2005). However, taxonomic research and taxonomic expertise are the two faces of the same coin: taxonomic research brings out hypotheses of species delimitations, and gives names to these hypotheses, whereas taxonomic experts try to attribute such names to a given specimen. In this context one should reminds the peculiar role of type specimens and more precisely how the species names are linked to holotypes. Indeed, first of all the holotype associated a name to a real organism. The name is chosen by a taxonomist when he proposes a new hypothesis of species delimitation. However when thereafter other taxonomists revise the group and propose new hypotheses of species delimitation the name remains linked to the holotype. Thus, when taxonomist reconsider species limits by using new data such as examination of new specimens and/or new characters, the assignment of name depends on which cluster (the new delimitations hypotheses) the type specimen falls in. In the case of molecular taxonomy (*e.g.* the use of molecular characters to delimit species) and the development of the Barcoding of life projects, we claim that the DNA barcode datasets must be set up following the requirements of specimens sampling of the taxonomic research, including naming practices. Following such requirements, the barcoding libraries would effectively avoid potential mistakes in species identifications and subsequent biases, for example, in molecular studies (i.e. phylogenetic reconstructions). After clarifying these sampling requirements, we present a case study on the genus *Eumunida* (Decapoda) for which COI and 28S rDNA data were gathered for 229 specimens from several locations including type specimens for a large part of the described species. We also compare the distribution of the states of the morphological characters used in the key for species identification over the identified genetic clusters. The inclusion of type specimens in the data set links unambiguously a genetic cluster to a taxa name. Finally, we use this dataset to examine the reliability of the taxonomy of the genus and test the performance of DNA-barcode as a tool for taxonomic expertise and its ability to pinpoint the necessity of taxonomic revision when additional specimens and characters are examined.

Species delimitations in the barcoding context

In order to clarify the requirements of specimens sampling of the taxonomic research, we should go back to the conceptual framework that justifies the methods used to delimit species in practice. In 1998, de Queiroz, contributed to clarify this framework by distinguishing the question “what is a species” from the question “how species are recognized?”. This distinction, clearly stated in de Queiroz (1998, 1999, 2005 a, 2005b, 2005c, 2007) as well as in Frost and Kluge (1995) and subsequently used by e.g. Knowles and Carstens (2007), Raxwourthy et al. (2003), Samadi & Barberouse (2006), Rissler and Apodaca (2007) or Giraud et al (2008) can now be considered classical. Considering this distinction, most delimitation practices can be unified under a definition of species close to ideas first proposed by Simpson (1951) or Wiley (1981) in which species are viewed as definitively diverging lineages. Following this way of thinking, we can derive three families of criteria that can be used to delineate species in practice and that actually correspond to the three major fields of biology involved in species delimitations (*i.e.* alpha-taxonomy, populations biology and phylogeny).

The classical way to propose primary species hypotheses is that of alpha-taxonomy., Classically the taxonomist, specialized for a given taxon, using his knowledge on the morphological characters (determinism, linkage to gender, etc…), examine as many

specimens as possible and, propose primary hypotheses about species limits that minimize the intra species variability. Mallet (1995) named this criterion “the genotypic cluster definition”. In the conceptual framework of the general lineage species concept this criterion is justified by the fact that since the individuals within a species are linked by genealogical relationships, they share a common pool of heritable characters. Thus, they display a certain degree of similarity that can be used to propose primary delimitation hypotheses. Note that in such approach, the background knowledge used by the taxonomists is essential to avoid to wrongly interpreting the measured similarity as an indication of belonging to the same reproductive community (for example in the case of characters that are subject to phenotypic plasticity or if they evolve under sexual selection or balancing selection). Concerning the naming practices, the taxonomist designated, from the set of specimens he used to propose his primary species hypothesis, one specimen as the holotype and, when the description is based on more material, several paratypes. These type specimens are thereafter definitively associated to the chosen name. However, for many characters and taxa, such knowledge is very scarce and thus such hypotheses based solely on clustering upon similarity may be not very robust and need to be reinforced by other evidences. In this context other data and criteria may be used to reconsider the primary hypotheses and propose new ones.

Among these criteria, one is that derived from the so-called Biological Species Concept. In the conceptual framework of the general lineage species concept, individuals within a species belong to a reproductive community, thus they must be inter-fecund and reproductive exchanges must be effective among them. Direct evidence of inter-fecundity is usually difficult to obtain. Thus, this criterion is generally used through indirect evidences (for example estimation of gene flows or evidence of recombination between unlinked characters). Population biologists, using the tools of populations genetic on heritable variable characters, but also experimental procedures or field and behavioural observations, produce indirect evidences of inter-fecundity within, or reproductive isolation between, sets of organisms. However, producing such evidences is not an easy task and can not be applied for many taxa.

The last family of criteria is the phylogenetic one. Indeed, as a species is a reproductive community that is definitively divergent from such other communities, individuals within such a community share a common history which can be inferred using heritable characters. In that purpose, the pattern of distribution of characters states is used to tell the phylogenetic history of a set of organisms. This criterion (reciprocal monophyly) applies very well when time since speciation is enough to allow fixation of different state of the characters in both sister species. Indeed, the coalescence time of the used characters must be shorter than the time since the speciation event. Thus, when speciation is recent, or when populations sizes are huge, shared ancestral polymorphism and the variance of coalescent time can scramble the phylogenetic signal and the phylogenetic criterion will not give adequate hypotheses of species delimitations. In order to detect recently diverging lineages new tools, based on the coalescent theory, have been recently proposed (e.g. Knowles, 2004; Knowles and Carstens, 2007) but are not yet commonly used.

These two last criteria are a way to test the primary hypotheses proposed following the classical way used in alpha-taxononomy. Using such criteria secondary species hypotheses may be proposed that may modify the attribution of species names. Once again, the attribution of names depends on where the type specimens fall following these additional criteria.

As we emphasized below each of these criteria may be used to delimit species but, whatever the criterion we use, we need characters that are not only variable, but above all, for which the variability (*i.e.* the different states of the character) is heritable. In this context, DNA offers many variable characters that can be used to delimit species using either of the derived criteria. When species are described upon other characters (usually morphological characters), DNA can be used to test the formulated hypotheses using the others criteria. To apply these

criteria some sampling requirements must be fulfilled. Indeed, to measure similarity or to estimate gene flows sampling must cover as well as possible characters variability and the geographic distribution within the species. The closest-related known species must also be included in order to determine if the species is an autonomous evolutionary lineage.
For barcoding projects an important additional question is how to link a given specimen to a hypothesis of species delimitation and then to a species name. This link between species hypotheses and specimens is materialized through the position within these hypotheses of the type specimens and more precisely through the holotype. Thus to rigorously link DNA-barcodes (*e.g.* molecular characters) to the available species names, we need to determine the DNA-barcode of the corresponding type-specimens. However, even if this point has been often emphasized (e.g. Hebert et al. 2004; Janzen et al. 2005, Smith et al., 2007), only very few studies included such material in their datasets. Indeed, obtaining sequences from old museum material is difficult because, depending on the methods used to preserve specimens and the age of the collection, DNA may undergo more or less important damages. Moreover DNA-barcode for holotype materials are only found in description of new species (*e.g.* Burns et al., 2007) although it is the only way to unambiguously link specimens to a species name. For example in the study of Smith et al. (2007) some of the scientific names are given in quotation marks because, without the DNA-barcode of the holotypes, it was not possible to definitively link the genetic clusters to the names.

Materials and methods

We used 229 *Eumunida* specimens from the South East Pacific and Indian Ocean belonging to the collections of the Museum National d'Histoire Naturelle (MNHN), Paris (Appendix). Among them, 9 are holotypes and 24 are paratypes of 13 different species. The 196 remaining specimens were morphologically identified to the species level by Enrique Macpherson and are attributed to 17 valid names of *Eumunida* species. These 17 names are represented by 1 to 95 specimens, with a mean of 12.05 specimens per name (Appendix). These morphological identifications were used as primary species hypotheses. The morphological characters used in species identification of all of the species of the genus were listed and used to build a morphological matrix (Table 1).
DNA was extracted from a piece of muscle tissue using the DNeasy® 96 Tissue kit (Qiagen), and specimens were kept as vouchers. A fragment of Cytochrome Oxidase I (COI) mitochondrial gene and a fragment of 28S rDNA nuclear gene were amplified using respectively universal primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer *et al*., 1994), and C1' (5'- ACCCGCTGAATTTAAGCAT-3', (Jovelin & Justine 2001) and D2 (5'-TCCGTGTTTCAAGACGG-3', (Dayrat et al. 2001). All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq, QBiogene for COI gene and Taq Core Kit 2, Qbiogene for 28S rDNA gene. Thermocycles consisted of an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 48°C for COI gene and 56°C for 28S rDNA gene for 40'' and extension at 72°C for 1'. The final extension was at 72°C for 10'.

Table 1: Description of morphological characters, and state for each species. The characters are numbered from 1 to 16. The different states of the characters are described (for example Yes/No for presence absence). Species for which molecular data were obtained are indicated in bold.

1 Thoracic spines, YES=1, NO=0
2 Posterior region of carapace with complete striae, YES=1, NO=0
3 Number of anterolateral spines on each side, one· spines=1, two spines=0
4 Pad on palm of cheliped, Yes=1, NO=0
5 Epigastric spines, YES=1, NO=0
6 Posterior part of abdominal tergites, after last stria, smooth, YES=1, NO=0
7 Depressed area on branchial region of carapace, YES=1, NO=0
8 Mesiodorsal row of spines on cheliped palm, YES=1, NO=0
9 First anterolateral spine less than half lateral supraorbital, YES=1(less), NO=0 (more)
10 Distal spines on carpus of chelipeds, 1=2 sp, 0=3 sp
11 Distal spine on merus of third maxilliped, YES=1, NO=0
12 Male Pleopods, YES=1, NO=0
13 Six to seven spines upper margin of propodus walking legs, YES=1, NO=0
14 Row of ventral spines on merus of chelipeds, 1=5-8sp, 0=1sp
15 Ocular peduncles short, not reaching end of lateral supraorbital spines, YES=1, NO=0
16 Lateral surface of 4th pereiopod with spines, YES=1, NO=0

	Characters															
SPECIES names	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
E. ampliata	0	1	1	1	0	1	0	0	1	0	0	0	0	1	0	1
E. annulosa	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	1
E. australis	1	1	0	1	0	1	0	0	0	0	0	0	0	1	1	1
E. balssi	0	1	0	0	0	1	0	1	1	0	1	0	0	0	0	1
E. bella	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	1
E. bispinata	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1
E. capillata	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0
E. debilistriata	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	1
E. depressa	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0
E. dofleini	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
E. funambulus	1	1	1	1	1	0	0	1	0	0	1	0	1	1	0	1
E. gordonae	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
E. karubar	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1
E. keijii	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1
E. laevimana	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0
E. macphersoni	1	1	1	0	0	0	0	1	1	0	0	0	0	0	1	0
E. marginata	0	1	0	1	1	0	0	1	1	0	1	0	0	0	0	0
E. minor	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1
E. multilineata	1	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0
E. pacifica	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1
E. parva	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1
E. picta	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1
E. similior	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0
E. smithii	0	1	0	0	0	0	0	1	1	0	0	1	0	1	1	1
E. spinosa	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E. squamifera	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
E. sternomaculata	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
E. treguieri	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1
Eumunida n. sp.	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1

Part of the PCR products were purified using Montage™ PCR Centrifugal Filter Devices (Millipore) and sequenced on a Ceq2000™ automated sequencer (Beckman) – corresponding to Genbank accession numbers AY800009-800046, AY800048, AY800050, AY800051, AY800055-800065 and DQ011181-011220. The other part of the PCR products were purified and sequenced by the Genoscope (genbank accession numbers EU243337 to EU243562 for COI gene and EU243574 to EU243663 for 28S rDNA gene). In all cases, both directions were sequenced to confirm accuracy of each haplotype sequences.
Sequences were manually aligned for COI gene, and automatically for 28S rDNA gene using Bioedit software (Hall 1999). A genetic distance matrix including all sequences was calculated for COI gene under the K2P model (Hebert et al 2004). Forr each gene, a

phylogenetic tree was built, using Bayesian Analysis (BA), consisting of two Markov chains (2000000 generations each with a sampling frequency of one tree each hundred generations) run in four parallel analyses using Mr. Bayes (Huelsenbeck, Ronquist & Hall, 2001). When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in. Only the number of nucleotide substitutions categories was fixed (n = 6). For both genes, we use *Munida acantha* as outgroup, a species of the same superfamily than *Eumunida* (genbank accession numbers: AY800033 for COI gene and EU249347 for 28S rDNA gene).

Results

Mitochondrial dataset

We obtained 226 COI sequences of 658 bp long with 219 polymorphic sites corresponding mostly to the third codon position (164) but also to the first codon position (47). This dataset is available in a general project in BOLD untitled " *Eumunida* barcodes and taxonomy" under the accessing numbers EUMU001-07 to EUMU226-07. The maximum K2P distance between pairs of COI sequences is 0.158, with a minimum of 0 and a mean of 0.079. The genetic distance between two type-specimens of the same species are always inferior to 0.033. This upper boundary of intra-specific distances between type-specimens corresponds to the distance between two paratypes of the species *E. bispinata*. The pairwise distance involving the holotype of *E. parva* and each of the five paratypes of *E. karubar*, ranges between 0.011 and 0.012. All other pairwise distances between two type-specimens of different species-names are superior to 0.059 (this lower boundary corresponding to the distance between holotypes of *E. annulosa* and *E. multilineata*). The histogram representing all the distances between types and non-types specimens displays two groups (Figure 1): the first, with an upper boundaries of 0.033, includes all the distances between two type specimens of the same species, but also distances between holotype of *E. parva* and each of the five paratypes of *E. karubar*; the second, characterized by a lower boundary of 0.043, includes only comparisons between types of different species. The NJ tree calculated on the COI dataset displays 16 terminal entities (data not shown). Each of these 16 terminal entities includes only sequences separated by a genetic distance inferior to 0.033. Among these 16 entities 10 contain 1 or several type specimens.

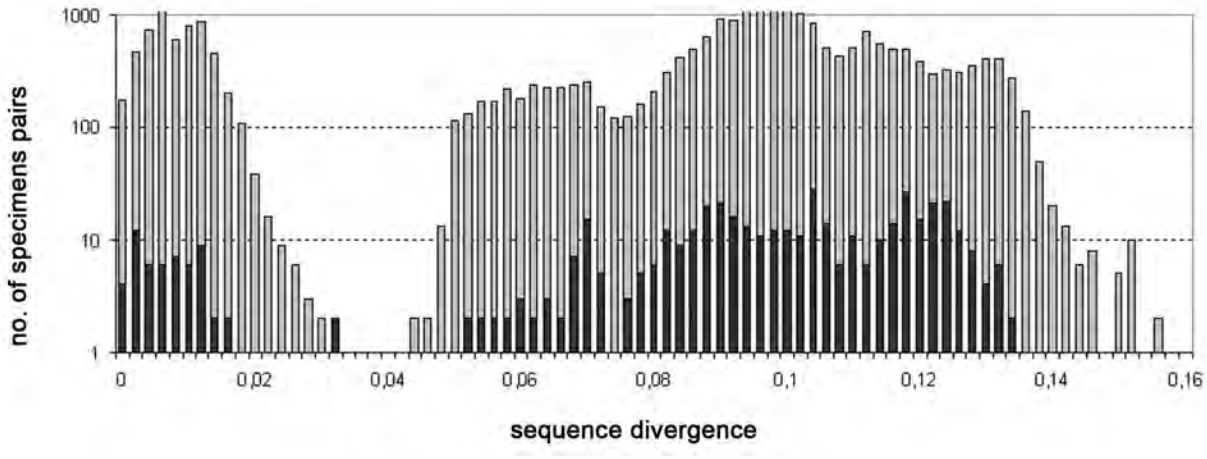


Figure 1: Histogram of genetic distances for the COI gene dataset. Black bars: pairs of type-specimens. Grey bars: pairs of non-type-specimens.

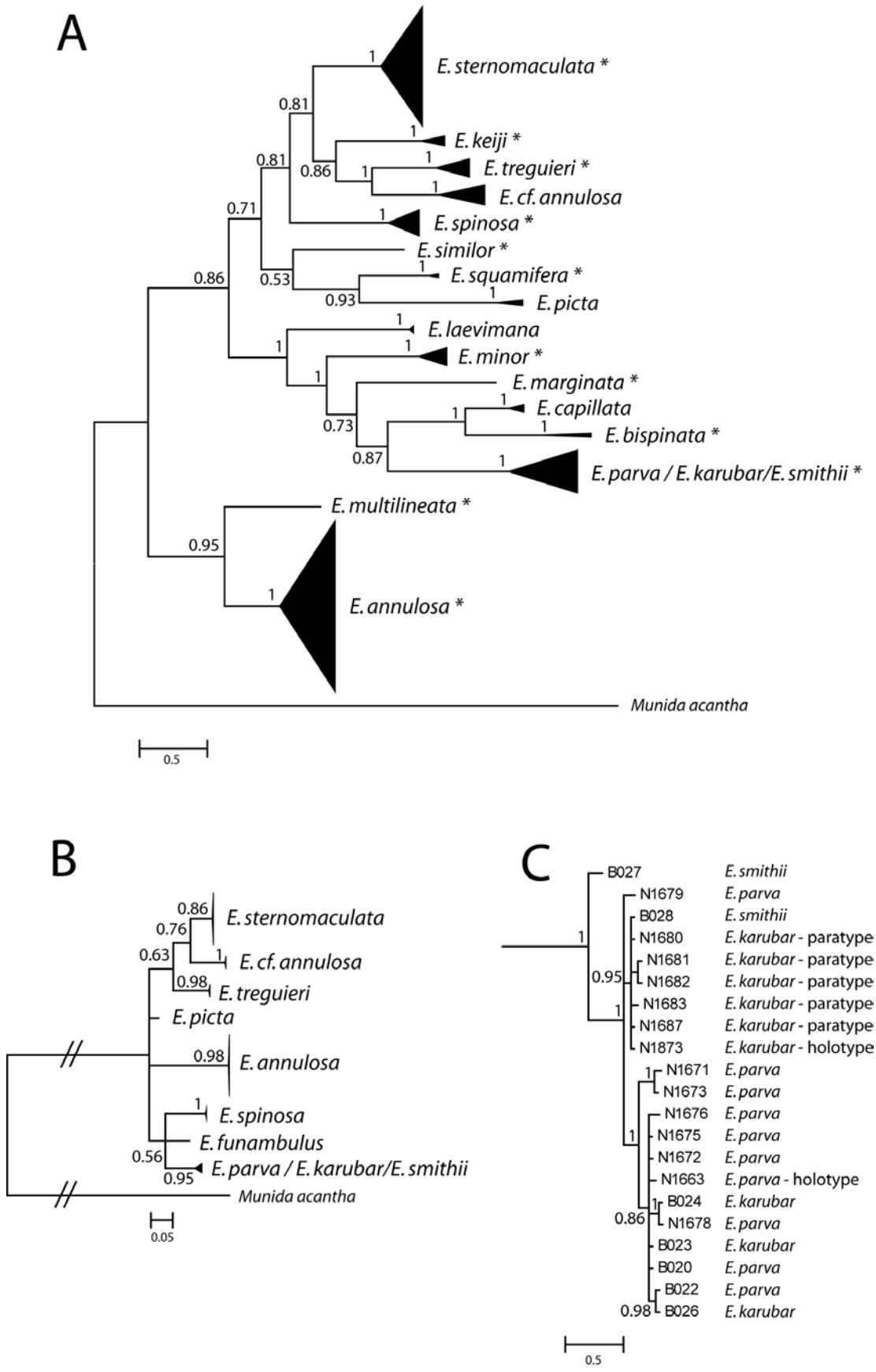


Figure 2: (A). Bayesian tree for COI gene dataset, with posterior probabilities indicated for each node. Clades are collapsed in triangles, with the height representing the number of specimens and the width the length of the branches. (B) Bayesian tree for the 28S gene dataset (C) Detail of the COI gene tree for the *E. parva/E. karubar/E. smithii* clade.

Nuclear dataset
For the 28S rDNA gene, we obtained 90 sequences of 867 bp long. The 28S dataset reveals the same monophyletic lineages than the COI dataset. Two groups of genetic distances are separated by a gap on the genetic distances histogram. The short-distance group has an upper bound of 0.002 and the long-distance group has an upper bound of 0.028. For each pair of specimens, a genetic distance inferior to 0.002 for this dataset corresponds to a genetic distance inferior to 0.033 when calculated with the COI dataset. Conversely, when the genetic distance between two 28S rDNA sequences is superior to 0.028, then the genetic distance between COI sequences corresponding to the same specimens is superior to 0.043. The distances between type specimens of the same species fall in the short-distance group whereas distances between type specimens of different species fall in the long-distance group.

Phylogenetic analyses
In the Bayesian tree (Figure 2A) obtained with the COI gene the 13 genetic entities that include several specimens, are monophyletic and are highly supported (PP=1). All the entities for which we obtained the data for several specimens for the 28S gene are also found monophyletic with this gene (Figure 2B). The phylogenetic relationships among these clusters are not fully resolved but some nodes are well supported with the COI gene.

Genetic entities and species names
Combining the two molecular data sets we are able to define 17 genetically distinct entities (Figure 2). Eleven of these entities can be directly linked to a species name, because they include at least the sequence of one type-specimen (holotype and/or paratypes) for at least one of the two genes (*E. annulosa*, *E. bispinata*, *E. keiji*, *E. marginata*, *E. minor*, *E. multilineata*, *E. similor*, *E. spinosa*, *E. squamifera*, *E. sternomaculata*, *E. treguieri*). For four of these entities specimens are morphologically attributed to a same species name (*E. capillata*, *E. funambulus*, *E. laevimana*, *E. picta*). One clade (*E. cf. annulosa*) includes specimens morphologically attributed to *E. annulosa* but no type specimen. This clade is not closely related to the *E. annulosa* clade, Indeed, the sister clade of *E. cf. annulosa* is *E. treguieri*. The remaining entity puts together specimens morphologically identified as belonging to three different species (*E. parva*, *E. karubar* and *E. smithii*). For the COI dataset, the holotype of *E. parva* and five paratypes and the holotype of *E. karubar* are included within this genetic cluster (Figure 2C).

Discussion

Identifying the gap between intra and inter-specific distances
This analysis, in which we included a large part of the known species of the genus *Eumunida*, reveals that within this genus, the distribution pattern of genetic distances among individuals has a bimodal pattern. We can indeed define clusters of genetically very close specimens. These clusters are separated from each other by larger distances and the intra-cluster distances do not overlap with inter-cluster distances. The phylogenetic analysis reveals that these genetic clusters are monophyletic. Thus these sets of organisms may be considered as species following two different criteria (similarity and reciprocal monophyly). The extensive sample size analyzed for some of the identified clusters suggests that the observed gap is not an artifact resulting from a sampling bias. We here point out, as others (Costa et al. 2007; Hajibabaei et al. 2007; Meyer & Paulay 2005; Wiemers & Fiedler 2007), the importance of the number of samples examined within species to identify a “species threshold”. However, the link between these genetic clusters and the morphological hypotheses needs a more

accurate analysis. The material links between classical morphological hypotheses and species names are the type specimens. Thus to link morphological and genetic hypotheses of species delimitation the states of the genetic characters should be also determined for type specimens. For this reason we included in our analysis as many holotype and paratype specimens as possible. The genetic distances between paratypes for a given name follows the distribution pattern of genetic distances obtained with a larger number of specimens but the gap between genetic distances among paratypes of a same species and type specimens of different species is greater. The inclusion of type specimens into the genetic clusters is the only way to unambiguously attribute species names to such genetic clusters, and thus to the included specimens, avoiding mistakes in future taxonomic or phylogenetic studies. As pointed out by Hajibabaei et al. (2006), obtaining the complete fragment for old museum specimens is not always easy, and in such cases the attribution of species names to genetic cluster can be based under some conditions on a smaller set of characters (shorter sequences).

Concordance of most of the genetic entities with primary species hypotheses

In this case study, morphological taxonomy provided a primary hypothesis of species belonging for each analyzed specimens and, on the other hand, the molecular analysis of type specimens permitted to link species names to DNA sequences. The analysis of molecular diversity among specimens morphologically attributed to a same species name – including type specimens – allows us to propose secondary hypotheses of species delimitations. In this analysis 15 of the 17 genetic entities identified with two genes fit with the primary hypotheses. These 15 entities include specimens strictly attributed to a unique species name, and 10 of them also include type specimens. For these 15 species names the morphological and the molecular identifications are in complete accordance and the data added in our study do not permit the rejection of primary delimitation hypotheses. Thus, these hypotheses are the best hypotheses given the data available. The sampling used in our analysis also permitted to enhance delimitations hypotheses in two opposite situations.

A cryptic species hidden under the name E. annulosa

Our analysis revealed two clades of specimens morphologically attributed to the same species name. Indeed, the two genes consistently divided the specimens attributed to *E. annulosa* into two distant genetic clades. The specimens falling in the clade that includes the holotype of *E. annulosa* should thus be considered as *E. annulosa*, whereas the other clade (*E. cf. annulosa*) corresponds to a species not yet detected by morphologists (Figure 2). Moreover in the phylogenetic analysis of the COI dataset, *E. cf. annulosa* is more related to the *E. treguieri* clade than to *E. annulosa* clade. The genetic divergence found between each of them largely exceeds the average intraspecific divergence found for the other species, as well as in other galatheoids (Machordom & Macpherson 2004). Therefore, the newly detected species has a clear molecular support. We examined the morphology of the specimens of this potentially new species and detected slight morphological differences with *E. annulosa* or *E. sternomaculata*. These two species are easily distinguished by the relative length of the first pair of anterolateral spines, longer in *E. sternomaculata* than in *E. annulosa*, and the presence of two (*E. annulosa*) or three (*E. sternomaculata*) distal spines on the carpus of the chelipeds (Table 1). The new species (*E. cf. annulosa*) displays intermediate characters. Indeed, the relative size of the first anterolateral spine is between *E. annulosa* or *E. sternomaculata* and a 3rd distal spine is present on the cheliped carpus but can be very small. This result enhances the importance of molecular analyses to detect cryptic species not only in this genus, but also in others crustacean decapods (see the review by Knowlton 2000). Moreover, these characters are difficult to observe on small specimens and the morphological characters distinguishing these species seems diagnostic only for adult specimens. In such a situation the DNA-barcode

expertise is more efficient than the traditional morphological approach since it is informative for all the stages of the life cycle. Such an expertise tool has indeed broad applications (*e.g.* (De Ley et al. 2005; Savolainen et al. 2005; Vences et al. 2005).

Synonymy of E. karubar, E. parva and E. smithi

Our analysis also suggests that among the 17 species names used for the primary hypotheses of species belonging, three correspond to a same genetic entity (*E. karubar, E. parva, E. smithi*) that includes five paratypes and the holotype of *E. karubar* and the holotype of *E. parva* (Figure 2C). Considering the COI dataset, the genetic distances between sequences falling into this well supported clade are lower than that obtained between others paratypes of a given species falling in a same clade (for example between the two paratypes of *E. bispinata*). However, the others specimens attributed to one of the three species names using morphological determination are scattered in the different sub-clades without any obvious significant pattern (Figure 2C). Indeed, type specimens and localities are mixed for the three species.

The morphological distinction among *E. parva*, *E. karubar* and *E. smithii* are based on the presence (*E. smithii* and *E. karubar*) or absence (*E. parva*) of ventral spines on the merus of the chelipeds (Table 1). Furthermore *E. smithii* and *E. karubar* have some ventromesial spines on the palm of the chelipeds, which are usually absent in *E. parva* (Table 1). The difference between *E. smithii* and *E. karubar* is based on the length of the ocular peduncles, shorter in *E. smithii* than in *E. karubar* (see De Saint Laurent & Poupin 1996). The molecular analyses undertaken show that the morphological differences found in these species may be considered as intraspecific variability, and that this variability does not give diagnostic features at the species level. Thus, from the combination of data (morphology, independent genetic characters and geography) we suggest that the three names are synonymous (*E. parva* and *E. karubar* are thus considered as junior synonyms of *E. smithii*). The alternative hypothesis is that both the speciation events are very recent (leading to weak genetic structure and small genetic distances) and that the morphological distinction between the three species is very weak (leading to misidentification of specimens others than type material).

Conclusions: from genetic entities to species delimitations

Our study illustrates the integrative taxonomy strategy advocated by several authors (*e.g.* Dayrat 2005; DeSalle 2006; Sites & Marshall 2003; or Vogler & Monaghan 2007). The first step on this strategy is to test the primary morphological hypotheses of species delimitation using independent characters (DNA polymorphism) and various species delimitation criteria. The inclusion of type specimens in the molecular study allows us to directly link a species name to a DNA-barcode. An appropriate sampling effort within species and the large taxonomic coverage within the genus permitted to support most of the primary species hypotheses based on morphology using the genotypic clustering criterion on molecular data but also to bring up new hypotheses. The use of a close out-group outside the genus allows to reveal that each of the proposed species has it own evolutionary history and thus to test the species hypotheses under the phylogenetic criterion. We here not test species delimitations with the biological criterion but, for some of the species a previous study (Samadi et al. 2006) permitted to show that gene flows occur among populations over the geographic range of the species and not among species suggesting that species hypotheses should stand up to the reproductive isolation criterion.

Our study highlights that such a sampling strategy allows DNA-barcoding to be used as an efficient determination key. In this purpose sampling must cover the intra and inter specific variability, the geographic distribution, and when possible, include type specimens in order to enhance taxonomy and bring out new hypotheses. However to enhance taxonomy, one should

go back to specimens and morphology. Thus, DNA-barcode is an efficient expertise tool only if sampling is directly and adequately linked to taxonomical research. The strategy followed in this case study when applied to major museum collections (when preservation of specimens allows molecular study) will greatly enhance the quality of the "DNA-barcoding" expertise tool.

Acknowledgments

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the *Tropical Deep Sea Benthos* programme on board R/V *Alis*, that generated the deep-sea samples used in this study. We are also pleased to thank the staff of MNHN's "Service de Systématique Moléculaire" and the Consortium National de Recherche en Génomique, Genoscope for technical facilities; Simon Tillier for constructive discussions and Mande Holford and Line Le Gall for constructive comments on, and English improvements of, the manuscript.

References

Burns JM, Hajibabaei M, Hallwachs W, and Herbert PDN. 2007. DNA barcodes of closely related (but morphologically and ecologically distinct) species of Skipper Butterflies (Hesperiidae) can differ by only one to three nucleotides. *Journal of the Lepidopterists' Society* **61:** 138-153.

Carstens BC, and Knowles LL. 2007. Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology* **16:** 619-627.

Costa FO, deWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M, and Hebert PD. 2007. Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* **64:** 272-295.

Dayrat B, Tillier A, Lecointre G, and Tillier S. 2001. New Clades of Euthyneuran Gastropods (Mollusca) from 28S rRNA Sequences. *Molecular Phylogenetics and Evolution* **19:** 225-235.

Dayrat B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society* **85:** 407-415.

De Ley P, De Ley IT, Morris K, Abebe E, Mundo-Ocampo M, Yoder M, Heras J, Waumann D, Rocha-Olivares A, Jay Burr AH, Baldwin JG, and Thomas WK. 2005. An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360:** 1945-58.

de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, eds. *Endless forms: species and speciation*. Oxford: Oxford University Press, 57-75.

de Queiroz K. 1999. The general lineage concept of species and the defininig properties of the species category. In: Wilson RA, ed. *Species: New interdisciplinary essays*. Cambridge, MA: MIT Press, 79-89.

de Queiroz K. 2005a. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America*. **102:** 6600-6607.

de Queiroz K. 2005b. A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences*. **56:** 196-215.

de Queiroz K. 2005c. Different species problems and their resolution. *BioEssays*. **27:** 1263-1269.

de Queiroz K. 2007. Species concept and species delimitation. *Systematic Biology*. **56**: 879-886.

DeSalle ROB. 2006. Species Discovery versus Species Identification in DNA Barcoding Efforts: Response to Rubinoff. *Conservation Biology* **20:** 1545-1547.

Frost J, Kluge AG. 1995. A consideration of epistemology in systematic biology, with special reference to species. *Cladistics* **10:** 259-294.

Giraud T, Refrégier G, Le Gac M, de Vienne DM, and Hood ME. 2008. Speciation in fungi. *Fungal Genetics and Biology*.

Hajibabaei M, Singer GAC, Hebert PDN, and Hickey DA. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics* **23:** 167-172.

Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, and Hebert PDN. 2006. A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes* **6:** 959-964.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95-98.

Hebert PD, and Gregory TR. 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology* **54:** 852-9.

Hebert PD, Penton EH, Burns JM, Janzen DH, and Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 14812-7.

Huelsenbeck JP, and Ronquist F. 2001. MrBayes: a program for the Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.

Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, and Hebert PD. 2005. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360:** 1835-45.

Jovelin R, and Justine JL. 2001. Phylogenetic relationships within the polyopisthocotylean monogeneans (Platyhelminthes) inferred from partial 28S rDNA sequences. *International Journal for Parasitology* **31:** 393-401.

Knowles LL, Carstens BC. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology*. **56**: 887-895.

Knowles LL. 2004. The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology* **17:** 1-10.

Knowlton N. 2000. Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* **420:** 73-90.

Machordom A, and Macpherson E. 2004. Rapid radiation and cryptic speciation in squat lobsters of the genus Munida (Crustacea, Decapoda) and related genera in the South West Pacific: molecular and morphological evidence. *Molecular Phylogenetics and Evolution* **33:** 259-79.

Mallet J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution* **10:** 294-299.

Meyer CP, and Paulay G. 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLoS Biology* **3**.

Saint Laurent M de, and Poupin J. 1996. Crustacea, Anomura: les espèces indo-ouest pacifiques du genre Eumunida Smith, 1880(Chirostylidae): description de six espèces nouvelles. *Mémoires du Muséum national d'histoire naturelle* **168**: 337-385.

Samadi S, and Barberousse A. 2006. The tree, the network, and the species. *Biological Journal of the Linnean Society* **89:** 509-521.

Samadi S, Bottan L, Macpherson E, De Forges B, and Boisselier M-C. 2006. Seamount endemism questioned by the geographic distribution and population genetic structure of marine invertebrates. *Marine Biology***:** 1-13.

Savolainen V, Cowan RS, Vogler AP, Roderick GK, and Lane R. 2005. Towards writing the encyclopedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360:** 1805-11.

Schindel DE, and Miller SE. 2005. DNA barcoding a useful tool for taxonomists. *Nature* **435:** 17-17.

Simpson GG. 1951. The Species Concept. *Evolution* **5:** 285-298.

Sites JJW, and Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution* **18:** 462-470.

Smith MA, Wood DM, Janzen DH, Hallwachs W, and Hebert PDN. 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proceedings of the national Academy of Science* **104:** 4967-4972.

Vences M, Thomas M, Bonett RM, and Vieites DR. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360:** 1859-1868.

Vogler AP, and Monaghan MT. 2007. Recent advances in DNA taxonomy. *Journal of Zoological Systematics and Evolutionary Research* **45:** 1-10.

Wiemers M, and Fiedler K. 2007. Does the DNA barcoding gap exist?-a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* **4:** 8.

Wiley EO. 1981. *Phylogenetics: the theory and practice of phylogenetic systematics*. John Wiley & Sons.

3. Identification d'œufs de gastéropodes

Dans le cas des Turridae, les hypothèses de délimitation d'espèces sont trop controversées pour pouvoir effectuer le même genre d'approche que pour le genre *Eumunida*. Avant de pouvoir constituer une base de données de référence, une analyse d'α-taxonomie intégrative (détaillée à la fin de ce chapitre et dans le suivant) devra en effet être réalisée. Sachant que les pontes de gastéropodes analysées ici ne pourront pas être identifiées précisément, les objectifs de cette étude étaient de :
- tester les outils d'identification (BOLD-IDS et BLASTn) et évaluer la représentativité des bases de données où des séquences d'ADN de gastéropodes sont disponibles (BOLD et GenBank) ;
- utiliser la reconstruction phylogénétique pour replacer les séquences des pontes dans la phylogénie des gastéropodes et ainsi identifier leurs positions systématiques.

Plusieurs pontes, dont la morphologie permet de penser que certaines appartiennent au groupe des Conoidea, ont été collectées lors de la mission Aurora 07 (Philippines). Le séquençage du gène COI a permis par comparaison avec les bases de données disponibles une identification de la ponte, même si le niveau spécifique n'a pas pu être atteint dans la plupart des cas (Article 4). Pour cela, plusieurs méthodes ont été comparées : une approche basée sur la ressemblance entre les séquences d'ADN, en comparant les séquences des pontes avec les séquences disponibles sous GenBank ou dans BOLD, mais également une approche phylogénétique (en utilisant les séquences disponibles dans GenBank) pour positionner les pontes dans la phylogénie des gastéropodes. Deux autres gènes mitochondriaux ont également été séquencés (16S et 12S) pour augmenter les chances de trouver une séquence correspondant à celle de la ponte dans les bases des données.

Les séquences d'ADN utilisées ont permis de positionner dans l'arbre phylogénétique les pontes inconnues, permettant de les identifier au niveau familial, sous-familial ou générique pour la plupart. Quelques pontes ont également pu être identifiées aux niveaux générique ou spécifique en les comparant avec les bases de données à l'aide des outils BOLD-IDS et BLASTn. Cependant, même si certaines familles de gastéropodes correspondent à des groupes dont la monophylie est supportée, ces résultats doivent être interprétés avec prudence au vu des changements qui ont pu être proposés aux niveaux familial et sous-familial au sein des Conoidea (Chapitre 1).

Dans la plupart des cas, seul un gène sur les trois analysés permet d'identifier la ponte. Cette analyse aura surtout permis de montrer que les bases de données sont actuellement trop incomplètes pour pouvoir être utilisées comme un outil d'identification. Seules quelques familles de gastéropodes sont représentées correctement dans GenBank ou dans BOLD et un effort d'échantillonnage et de séquençage important doit être mis en place afin de pouvoir utiliser ces bases de données pour de l'expertise taxonomique.

Identification of neogastropod eggs using DNA barcodes

N. Puillandre*[,1], A. Couloux[2], S. Samadi[1]

[1] Université Pierre et Marie Curie (UPMC), UMR 7138, Systématique, adaptation, évolution (UPMC/IRD/MNHN/CNRS), CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.
[2] GENOSCOPE, Centre National de Séquençage, Evry, France

Keywords: DNA Barcoding, egg capsules, neogastropod, BOLD, GenBank database

* Corresponding author: UMR 7138, Systématique, adaptation, évolution, UPMC; IRD; MNHN; CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.; puillandre@mnhn.fr; +33 1 40 79 37 43

En préparation ; pour soumission à Molecular Ecology Resources

Running title: Barcoding neogastropod eggs

Abstract

Identifying species based on early life stages can be problematic, especially if only adult morphology is described. DNA based identification, known as DNA barcoding, can be used to link several life-stages of the same species, and thus to propose a taxonomic identification of eggs, larvae or juveniles. We propose here to use such approach to identify several marine neogastropods eggs collected in France and Philippines, by comparing their DNA to available sequences in GenBank and BOLD databases. Identification are based on similarity methods (BLAST in GenBank, Identification Engine in BOLD) to propose a species name. Only a few egg capsules are identified at species level, but several others are putatively identified at genus level. When such identification is impossible, because no matching sequences are found in the databases, phylogenetic reconstruction using all available sequences in GenBank is used to position the unknown egg capsules in a higher-ranked taxa. Our results highlight the poor taxonomic coverage achieved at present in genetic databases, emphasizing the need to develop DNA barcoding projects on molluscs.

Introduction

Identification of early life stages is particularly difficult for marine life, where in situ studies or ex situ development from larvae to adults are difficult to perform. To answer this problem, DNA barcodes have now demonstrated in several animal groups their capacity to link contrasted life stages of a same species (Blaxter 2004, Steinke et al. 2005, Pegg et al. 2006, Thomas et al. 2005, Vences et al. 2005, Ahrens et al. 2007), and they have already been used for marine fauna to link larvae to adults (Victor 2007). Such approaches were also used to explore biodiversity in poorly-known groups, leading to the discovery of previously unrecognized species, when only adults were examined (Barber & Boyce 2006). Among mollusks, and especially bivalves, DNA has already been used to identify early stages, using rRNA probes for in situ hybridization (Le Goff-Vitry et al. 2007; Pradillon et al. 2007; Jones et al. 2008) or PCR-SSCP combined with sequencing (Livi et al. 2006). Recently, DNA barcodes were also used to link egg masses, larvae and adults, but the analysis were limited to one family of gastropods, the Naticidae (Huelsken et al. 2008).
Here, we test the capacity of the animal barcode, the COI gene, and of two other mitochondrial genes (16S and 12S) to identify molluscan eggs. Egg capsules were collected in France and Philippines, and most of them are difficult to identify solely using morphological characters. Eggs capsules morphologies are described only for some species in a few groups of gastropods (Thorson 1935, Thorson 1940, Bandel 1976). The lack of data concerning the variability of egg cases shapes for closely related species, or for different genera in one family, does not facilitate the morphological identification, even at the familial level. Thus, DNA sequences were employed to identify the egg capsules, using two methods based on similarity: the identification engine of BOLD (Barcode Of Life Database) and the BLAST algorithm on GenBank. However, taxonomic diversity in mollusc is certainly underepresented in gene databases, as less than 16000 molluscan COI sequences are currently published in GenBank, corresponding potentially to only 3688 species over the around 100000 valid species already described (Bouchet 2006). We thus relied on phylogenetic reconstruction (using all available sequences from GenBank), in order to tentatively determine the sister taxa of unidentified egg-capsules, when similarity-based methods were not effective. Using this two-steps approach (similarity-based methods, and then phylogenetic reconstruction when taxonomic coverage is too weak), most of egg capsules are identified, but only because several genes and several methods are used, emphasizing the need for a more complete assessment of marine biodiversity to complete sequences databases.

Material and methods

Sampling
Two egg capsules (EC1 and EC2 - Fig. 1A) were collected on French coasts in 2006. As French coastal fauna is well-known, these two capsules were identified at the species level based on the presence of adults near the egg capsule (Fig. 1A) or on the morphology of the egg capsule (*Erosaria spurca* and *Coralliophila meyendorffi*). They will allow testing the capacity of DNA barcodes to identify determined egg capsules. Twenty-four egg capsules were collected in Philippines, between 150 and 1450 m deep, during the Aurora 2007. Based on the morphology, they were identified as neogastropods capsules: they all share the cupola shape typical of neogastropods egg cases (Bandel 1976) (Fig. 1C-H). One egg capsule shows a different shape, a bilaterally flattened capsule, typical of the genus *Conus* (Conidae, Neogastropoda – Fig. 1B). The capsules were first photographed on the substrate, and then placed in 95 % ethanol.

Sequencing

DNA was extracted from the whole egg capsule, using 6100 Nucleic Acid Prepstation system (Applied Biosystem). Three genes fragments, corresponding to some of the most represented genes for mollusks in GenBank, but also to genes classically used at the species level (Hebert et al. 2003, Remigio & Hebert 2003) were amplified: (i) a fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene using universal primers LCO1490 and HCO2198 (Folmer et al. 1994), (ii) a fragment of 550 bp of 16S mitochondrial gene using primers 16SH (CGTGATCTGAGTTCAGACCGG) and 16SL (GTTTACCAAAAACATGGCTTC) and (iii) a fragment of 600 bp of 12S mitochondrial gene using primers 12SI (TGCCAGCAGYCGCGGTTA) and 12SIII (AGAGYGRCGGGCGATGTGT). All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq, QBiogene for all genes. Thermocycles used for COI gene are those described in Hebert et al. (2003); for 16S and 12S genes, they consisted in an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 52°C for 16S and 54°C for 12S. The final extension was at 72°C for 10'. PCR products were purified and sequenced by the Genoscope. In all cases, both directions were sequenced to confirm accuracy of each haplotype sequence (GenBank Accession Numbers: EU870520-EU870589).

Species identification

Two identification tools based on genetic similarity were used to identify egg capsules. First, each sequence was compared to sequences available in GenBank, using BLASTn algorithm as implemented in GenBank (defaults parameters). The best hit, associated with the percentage of maximum identity, was conserved. Second, each sequence was compared to all available barcode records in BOLD, using the identification engine BOLD-IDS, using the option "searching all barcode records in BOLD". It provides a list of similar sequences, with the associated taxon name and the percentage of similarity. Contrary to the BLAST algorithm, identification in BOLD is based on genetic distances, and is not influenced by sequence length (Ratnasingham & Hebert 2007). A cut off value of 1% is used for species identification, *i.e*. when no sequence of less than 1% divergence is found in the database, no species identification is given (Ratnasingham & Hebert 2007).

In a second step, all egg capsules sequences were included in a phylogenetic analyses to identify groups of closely related sequences (OTUs). All the neogastropod sequences available in GenBank were also included in the analysis in order to identify the sister groups of the OTUs. For the COI gene, several keywords were used in order to include all the sequences of the gene ("COI", "CO1", "COX1" "cytochrome oxidase 1" and "cytochrome oxidase I"). To limit the total number of sequences, only one sequence by species was conserved. An alignment with the egg capsules sequences was created and only the GenBank sequences that correspond to the COI, 16S and 12S fragments sequenced for the egg capsules were conserved. Finally, 159, 127 and 54 sequences from GenBank were used for the phylogenetic analyses (respectively for COI, 16S and 12S gene). A phylogenetic tree was built, using bayesian analysis, consisting of two Markov chains (2000000 generations each with a sampling frequency of one tree each hundred generations) run in four parallel analyses using Mr. Bayes (Huelsenbeck et al. 2001).

The three methods (BLAST, BOLD-IDS and phylogenetic reconstruction) were used for each gene, except for BOLD-IDS based identifications as BOLD only contains COI sequences.

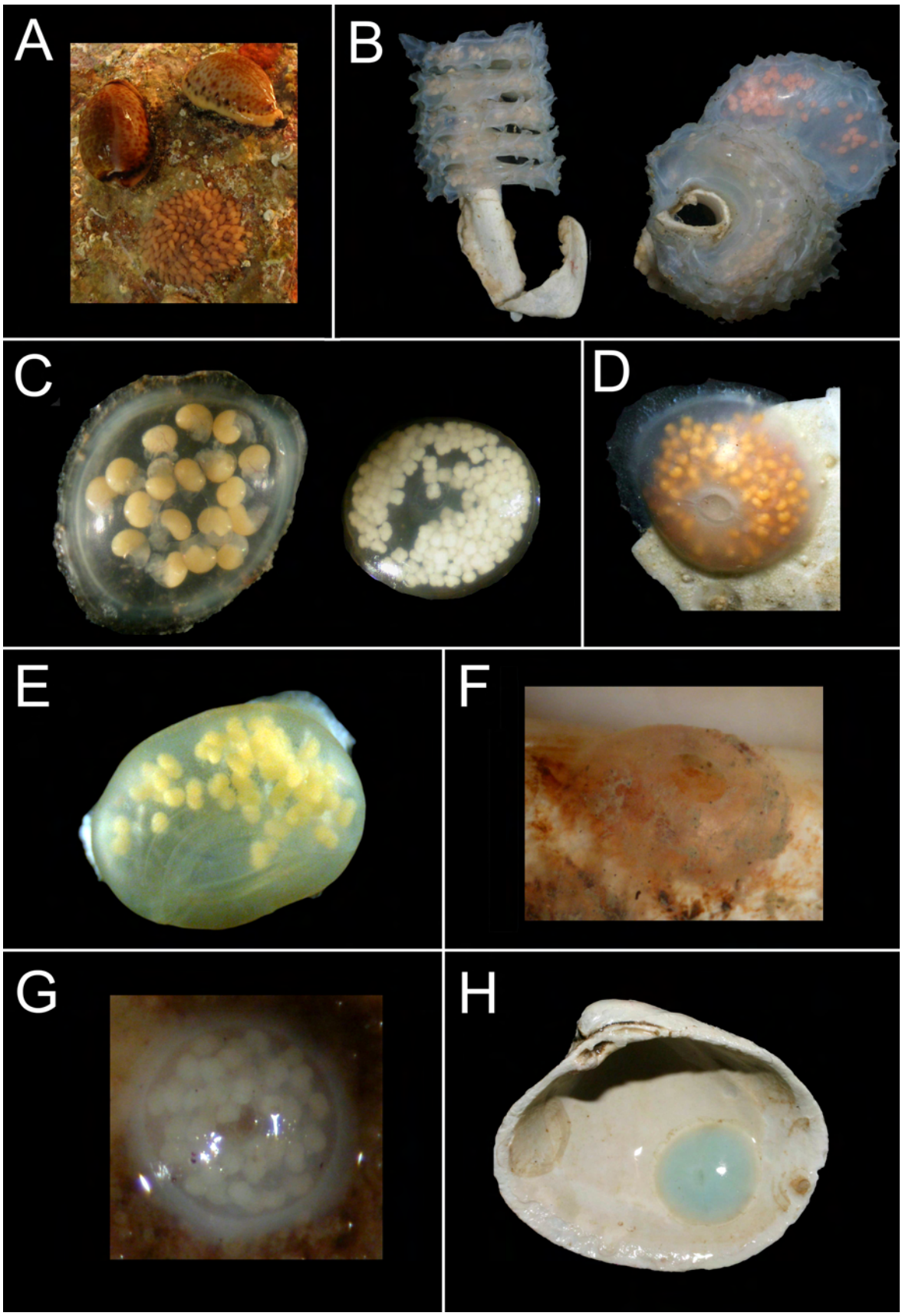


Figure 1: Illustration of some egg capsules. A. EC2 (*Erosaria spurca* – eggs and adults) B. EC3 (*Conus* sp.) C. EC9 (left) and EC11 (*Granulifusus* sp., at two different stages) D. EC8 (*Comitas* sp.) E. EC5 (*Belomitra* sp.) F. EC14 (not identified) G. EC23 (Conidae, Clathurellinae) H. EC26 (Conidae, Raphitominae) fixed on a bivalve shell. Pictures: André Hoareau (A) and Barbara Buge (B-H).

Table 1: identification for each egg capsules provided by COI, 16S and 12S genes.

		COI					16S			12S			Final ID
		BOLD		GenBank		Phylogenetic analyses	GenBank		Phylogenetic analyses	GenBank		Phylogenetic analyses	
N°	OTU	ID	%	ID	%		ID	%		ID	%		
EC1	1	x	x	*Pluminicola viras*	76		*Conus circumcisus*	82		*Coralliophila meyendorffii*	**99**		*C. meyendorffii*
EC2	2	*Erosaria spurca*	**99.51**	*Erosaria spurca*	**98**		*Erosaria spurca*	**99**		*Cypraea annulus*	94		*Erosaria spurca*
EC3	3	*Conus sulcatus*	88.94	*Conus venulatus*	90	*Conus sp.*	*Conus radiatus*	**95**	*Conus sp.*	*Conus textile*	90	*Conus*	*Conus sp.*
EC4	4	*Busycon carica*	87.19	*Busycon sinistrum*	86		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC5	5	*Batillaria multiformis*	87.63	*Busycon sinistrum*	86	*Belomitra sp.*	*Buccinum opisoplectum*	89		*Ilyanassa obsoleta*	84		*Belomitra sp.*
EC6	5	*Batillaria multiformis*	87.45	*Busycon sinistrum*	86	*Belomitra sp.*	*Buccinum opisoplectum*	89		*Ilyanassa obsoleta*	83		*Belomitra sp.*
EC7	5	*Busycon sinistrum*	85.74	*Nucella lapillus*	84	*Belomitra sp.*	*Burnupena cincta*	89		*Gemmula sogodensis*	83		*Belomitra sp.*
EC8	6	*Comitas*	**97.84**	*Pyrgulopsis gladulosa*	85	*Lophiotoma*	*Comitas kaderlyi*	**98**	*C. kaderlyi*	*Gemmula rosario*	84		*Comitas sp.*
EC9	4	*Busycon carica*	87.26	*Nucella lapillus*	87		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	87		*Granulifus sp.*
EC10	7	*Lacuna pallidula*	84.64	*Ilyanessa obsoleta*	84		*Buccinulum linea*	86		*Ilyanassa obsoleta*	79		?
EC11	4						*Granulifusus niponicus*	**95**	*G. niponicus*	*Ilyanassa obsoleta*	87		*Granulifus sp.*
EC12	8									*Gemmula rosario*	83	Turridae	Turridae
EC13	4	*Busycon carica*	87.61	*Ilyanessa obsoleta*	86		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	85		*Granulifus sp.*
EC14	7						*Buccinulum linea*	85					?
EC15	8						*Penion chathamensis*	85		*Lophiotoma unedo*	79	Turridae	Turridae
EC16	4	*Busycon carica*	87.09	*Busycon sinistrum*	86		*Granulifusus niponicus*	**97**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC17	4	*Busycon carica*	87.3	*Busycon sinistrum*	86		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC18	4	*Busycon carica*	86.99	*Busycon sinistrum*	85		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC19	4	*Busycon carica*	87.26	*Acanthinucella punctulata*	84		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC20	4	*Busycon carica*	87.26	*Acanthinucella punctulata*	84		*Granulifusus niponicus*	**98**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC21	9									*Lophiotoma polytropa*	84	Cypraeidae	?
EC22	4	*Busycon carica*	87.37	*Busycon sinistrum*	86		*Granulifusus niponicus*	**97**	*G. niponicus*	*Ilyanassa obsoleta*	86		*Granulifus sp.*
EC23	10	*Nannodiella*	87.4	*Gymnobela sp.*	86	Raphitominae	*Conus consors*	86	*Conidae*	*Ilyanassa obsoleta*	82		Conidae (Raphitominae)
EC24	4	*Busycon carica*	87.19	*Busycon sinistrum*	86		*Granulifusus niponicus*	**97**	*G. niponicus*	*Ilyanassa obsoleta*	87		*Granulifus sp.*
EC25	4	*Busycon carica*	87.19	*Busycon sinistrum*	86		*Granulifusus niponicus*	**97**	*G. niponicus*	*Ilyanassa obsoleta*	87		*Granulifus sp.*
EC26	11	*Gymnobela*	87.52	*Nannodiella*	87	Clathurellinae	*Raphitoma linearis*	87	*Conidae*	*Lophiotoma acuta*	81		Conidae (Clathurellinae)

Match scores are provided for BOLD identification with COI gene, and for BLAST identification for the three genes. For each egg capsules, a putative final identification is proposed, based on the results obtained. In bold: match scores superior (or equal) to 95 %.

Results

Similarity-based identifications

Only one capsule was identified at the species level by the BOLD identification engine (Egg Capsule EC2: 99.51% of identify with one sequence of Erosaria spurca), confirming the prior identification (Table 1). The EC8 matched with as a specimen of *Comitas sp.* (Turridae), but only with a similarity of 97.84%. All the others COI sequences match between 84 and 89% with one or several sequences of BOLD, far from the classical genetic distances encountered within one species (Hebert et al. 2003). Except for the specimen of *Erosaria spurca*, the first hit of the BLAST research for the COI gene is only matching with the query sequence with a matching percentage between 76 and 90.

Twelve 16S egg capsule sequences match with a 16S sequence of *Granulifusus niponicus* (between 95 and 98% of identity with BLASTn). The identification of the EC2 (*Erosaria spurca*) is also confirmed by the 16S gene. The 16S sequence of the EC3 matches with a sequence of *Conus radiatus* (95%). The 12S sequence of the specimen EC1 corresponds with a high identity score (99) to a sequence of *Coralliophila meyendorffii* in GenBank.

Tree-based identifications

Among the 26 egg capsules, eleven OTUs were recognized using phylogenetic reconstruction (Table 1). As shown in Fig. 2 for the 16S gene, these OTUs were positioned in a phylogenetic context, confirming several similarity-based identifications but also identifying potential sister-groups of unidentified egg capsules. For example, with the 16S gene, the twelve sequences similar to the sequence of *G. niponicus* form a clade (OTU 4, posterior probability PP = 0.98) that includes the GenBank sequence of this species (Fig. 2). The EC23 and EC26 are closely related to a sequence of *Raphitoma linearis* (Conidae, Raphitominae), with a PP of 1. With the 12S gene (results not shown), the OTU 8 clusters within a group composed exclusively of Turridae species (PP = 0.97), together with the EC8, confirming once again that it is a Turridae. With the COI gene, the OTU 5 (EC5, 6 and 7) is closely related to a sequence of *Belomitra sp.* (Buccinidae) (PP = 1). COI gene tree also identify EC23 as a Raphitominae (Conidae) and EC26 as a Clathurellinae (Conidae), thus confirming the 16S family identification for these two egg capsules.

Discussion

Eggs capsules identification

The species identification of the two egg capsules collected in France illustrates the capacity of the DNA to recover the same species identification as proposed by traditional taxonomy. However, only one of these two egg capsules was identified by the COI barcode. None of the egg capsules collected in Philippines is identified with the COI barcode, except one (EC8). Classical threshold for species identification based on the COI gene range from 1 to 2% (Hebert et al. 2004, Bichain et al. 2007, Ratnasingham & Hebert 2007), and except for the specimen of *Erosaria spurca*, match scores are clearly superior to this threshold. However, in several cases, we can reasonably propose a genus or family identification for several egg capsules. First, even if a threshold for genus level is more difficult to evaluate, as the level of diversity included can vary greatly between taxa (Holland et al. 2004), low genetic distances (inferior or equal to 5%) suggest a genus identification for several egg capsules (*Comitas sp.*, *Granulifusus sp.*, *Conus sp.*). Second, the positioning of the OTUs in the phylogenetic tree confirms the similarity based identifications, but also identifies the sister taxa of several egg

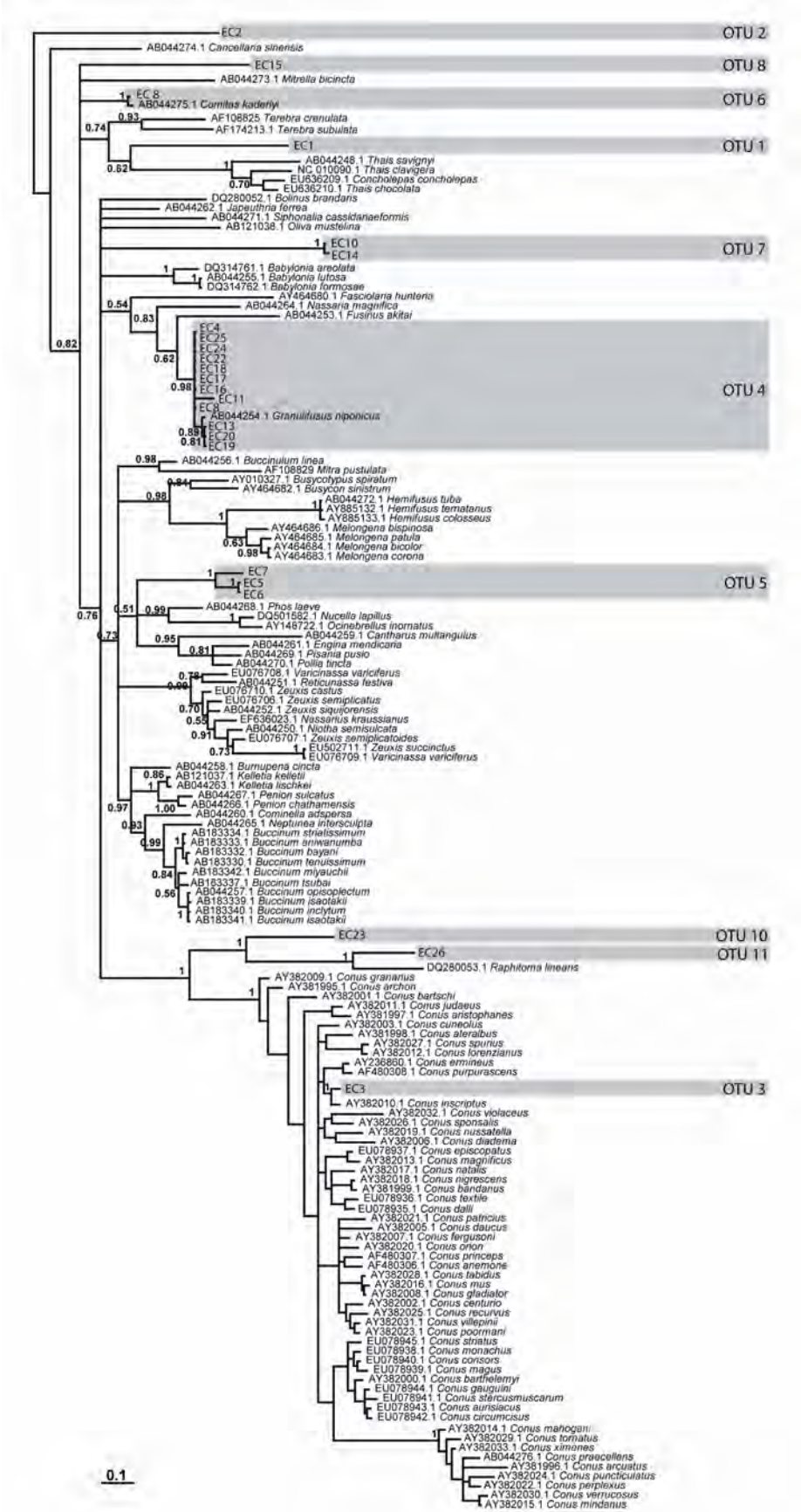


Figure 2: Phylogenetic tree (Bayesian analysis) obtained with all 16S neogastropods sequences from GenBank and the 23 eggs capsules 16S sequences. EC2, a non-neogastropod, is used as outgroup.

capsules (e.g. EC23 and EC26) and allow an identification at above-species rank (genus, subfamily or family).

Unidentified egg capsules

Three egg capsules (corresponding to OTUs 7 and 9) are not identified, whatever the gene or the method used: hit scores for these specimens with BLAST or BOLD are all inferior to 90%. For the three genes, BLAST and BOLD results indicate that when the first matching sequence displays a low similarity percentage (i.e <90%), the following matching sequences correspond (with lower or similar percentage) to completely different taxa. For example, BLAST results indicate that the COI sequence of the EC3 matches first with *Lacuna pallidula* (Littorinimorpha) with a sequence identity of 85%, but then with *Nucella lamellosa* (Neogastropoda, Muricidae) with the same score, *Ilyanassa obsoleta* (Neogastropoda, Nassaridae; 84%), *Urosalpinx perrugata* (Neogastropoda, Muricidae; 84%), and so on. This indicates that we cannot rely on such scores to propose a species, a genus or even a family name. However, this is not true in all cases. For example, EC23 and EC26 were identified with BLASTn as *Gymnobela* (Raphitominae, Conidae) and *Nannodiella* (Clathurellinae, Conidae), with respectively 86 and 87%. Indeed, these results are proven to be valid, as they are confirmed by the phylogenetic analyses.

These three unidentified capsules were furthermore not solidly positioned by phylogenetic analyses, as relationships with others taxa are not resolved.

Morphology of eggs capsules

Egg capsule of the EC3 is clearly that of a typical *Conus* (Fig. 1 B), and DNA analysis confirms this hypothesis. Egg capsule of the EC5, identified as a *Belomitra* (Buccinidae) is globular (Fig. 1 E). Even if *Belomitra* eggs morphology was never described, this result is consistent with the literature, as similar egg capsules shapes have been already reported for several (but not all) buccinid genera (Thorson 1935, 1940). All other egg-capsules examined in the present study share a similar cupola shape, with a hole at the apex of the capsule, and including more or less eggs. This morphology is traditionally supposed to correspond to Turrids or Muricids egg capsules (Pastorino 2007, Y. I. Kantor pers. com.), but our results indicate that this shape can be also found in Fasciolarids at least (specimens identified as *Granulifusus* sp., Fig. 1 D). DNA barcoding approach is thus not only an identification tool, as it allowed here the discovery of previously unreported egg capsules morphologies.

Databases representativeness

The accuracy of the DNA-based identification is clearly dependent on the quality of the databases. This quality can be evaluated by several criteria. First, the taxonomic coverage is low, as even a well-known species as *Coralliophila meyendorffi* is not represented in databases for COI and 16S gene. Actually, there is not a single Coralliophilidae COI or 16S sequence in these databases. Second, even if phylogenetic analyses seem to be a conclusive method in several cases where taxonomic coverage of databases does not allow a clear identification based on similarity, they cannot provide an accurate identification. The weak amount of information provided by each gene does not allow a correct resolution of the phylogenetic relationships within neogastropods. Third, it was impossible to concatenate several genes (in order to improve the resolution of the tree obtained with single-gene analyses), as taxa available for one gene did not correspond to the taxa of the other gene.

In order to test more thoroughly the taxonomic coverage of gene databases, we evaluated the number of sequences available for some of the most common molluscan families. We compared the number of species found in a tropical site in New-Caledonia (Koumac - Bouchet et al. 2002) with the number of COI sequences available in GenBank (Fig. 3). The

results illustrate the important bias in term of families represented in GenBank: among gastropods, five of the six speciose families in Koumac are each represented by less than three sequences. On the contrary, some families are over-represented in the database, regarding their number of species. The Cypraeidae comprise 232 species in the world (Lorenz 2002), and are represented by 682 COI sequences in GenBank; on the contrary, the Turridae are suspected to include more than 10000 species (Bouchet et al. 2002), and 92 COI sequences are available. This explains why the egg capsules of *Erosaria spurca* (Cypraeidae) was successfully identified at the species level and why the egg capsules of Turridae were not. A comparison of the number of sequences available in GenBank between 2006 and 2008 shows up a little increase of the number of sequence. The number of sequences clearly increases for some groups (e.g. *Conus*, Littorinidae, Neritidae). This increase is sometimes the consequence of the addition of a large number of specimens for a limited number of species (in the case of the Neritidae for example, where most of the sequences are the results of population genetics analyses). However, most of the speciose families still remain largely under-represented for the COI gene.

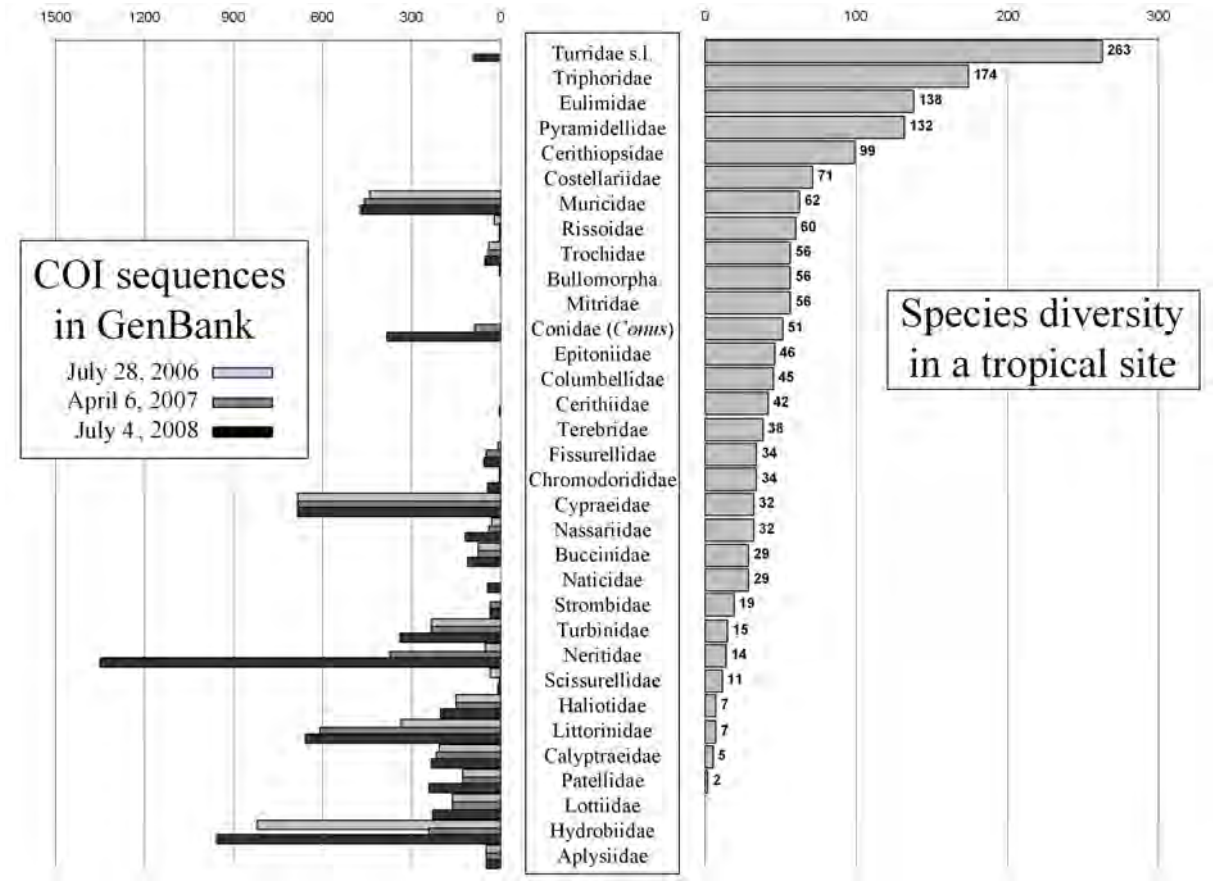


Figure 3: Comparison of the number of COI sequences found in GenBank in 2006, 2007 and 2008, with the number of species found in a tropical site (Koumac, Bouchet *et al*. 2002) for several gastropods families.

Conclusion

The COI gene, the universal barcode for animals, in association with two other mitochondrial genes, proved its ability to identify egg capsules, and thus constitute an interesting tool for taxonomic identification in various field of research. However, our results indicate that the

actual number of species, genera or even families represented in the database is too weak to provide a correct identification for most of the species, even for well-known European species. In such cases, when databases are not complete enough to provide an accurate identification of unknown specimens, phylogenetic analyses are appropriate to propose hypotheses of positioning at high taxonomic level. Barcoding efforts are now developed on emblematic or economically important groups, but similar projects have to be developed for several other groups such as molluscs, not only to facilitate taxonomic expertise but also to enhance species discovery in such poorly known groups.

Acknowledgments

The AURORA 2007 cruise (Principal Investigators: Marivene Manuel, NMP, and Philippe Bouchet, MNHN) on board M/V DA-BFAR was a joint project between the Philippines Bureau of Fisheries and Aquatic Resources (BFAR), National Museum of the Philippines (NMP), and Muséum National d'Histoire Naturelle (MNHN). It was made possible by a grant from the Lounsbery Foundation, and was carried under a *Census of Marine Life / Census of Margins* umbrella. We are grateful to André Hoareau and Jacques Pelorce for the two specimens collected in France. We are also pleased to thank the staff of MNHN's "Service de Systématique Moléculaire" for technical facilities, the Consortium National de Recherche en Génomique, Génoscope for the sequencing facilities, Barbara Buge for the pictures of the egg capsules and Yuri Kantor for constructive discussion about egg capsules morphology.

References

Ahrens D, Monaghan MT, Vogler AP (2007) DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae). *Molecular Phylogenetics and Evolution*, **44**, 436–449.

Bandel K (1976) Spawning, Development and ecology of some higher neogastropoda from the Carribean Sea of Colombia (South America). *The Veliger*, **19**, 176-193.

Barber P, Boyce SL (2006) Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proceedings of the Royal Society B*, **273**, 2053-2061.

Bichain JM, Boisselier MC, Bouchet P, Samadi S (2007) Delimiting species in the genus *Bythinella* (Mollusca: Caenogastropoda: Rissooidea): molecular and morphometric approachs. *Malacologia*, **49**, 291-311.

Blaxter ML (2004) The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London B*, **359**, 669-679.

Bouchet P, Lozouet P, Maestrati P, Héros V (2002) Assessing the magnitude of species richness in tropical marine environments: exceptionnally high numbers of molluscs at a new Caledonian site. *Biological Journal of the Linnean Society*, **75**, 421-436.

Bouchet, P. (2006). The magnitude of marine biodiversity. In F. BBVA (Eds) *The Exploration of Marine Biodiversity - Scientific and Technological Challenges* (pp.

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.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA Barcodes. *Proceedings of the Royal Society B*, **270**, 313-321.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. *PLOS Biology*, **2**, 1657-1663.

Holland BS, Dawson MN, Crow GL, Hofmann DK (2004) Global phylogeography of *Cassiopea* (Scyphozoa: Rhizostomeae): molecular evidence for cryptic species and multiple invasions of the Hawaiian Islands. *Marine Biology*, **145**, 1119-1128.

Huelsenbeck JP, Ronquist F, Hall B (2001) MrBayes: bayesian inference of phylogeny. *Bioinformatics*, **17**, 754-755.

Huelsken T, Marek T, Schreiber S, Schmidt I, Hollmann M (2008) The Naticidae (Mollusca: Gastropoda) of Giglio Island (Tuscany, Italy): Shell characters, live animals, and a molecular analysis of egg masses. *Zootaxa*, **1770**, 1-40.

Jones WJ, Preston CM, Marin R, Scholin CA, Vrijenhoek RC (2008) A robotic molecular method for in situ detection of marine invertebrate larvae. *Molecular Ecology Resources*, **8**, 540-550.

Le Goff-Vitry MC, Chipman AD, Comtet T (2007) *In situ* hybridization on whole larvae: a novel method for monitoring bivalve larvae. *Marine Ecology progress Series*, **343**, 161-172.

Livi S, Cordisco C, Damiani C, Romanelli M, Crosetti D (2006) Identification of bivalve species at an early developmental stage through PCR-SSCP and sequence analysis of partial 18S rDNA. *Marine Biology*, **149**, 1149-1161.

Lorenz F (2002) *New Worldwide cowries. Descriptions of new taxa and revisions of selected groups of living Cypraeidae (Mollusca: Gastropoda)* Conchbooks.

Pastorino G, Penchaszadeh PE, Scarabino F (2007) Egg capsules, eggs and embryos of the southwestern Atlantic gastropod *Coronium coronatum* (Gastropoda: Murricidae). *Journal of molluscan studies*, **73**, 61-65.

Pegg GG, Sinclair B, Briskey L, Aspden WJ (2006) MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia. *Scientia Marina*, **70**, 7-12.

Pradillon F, Schmidt A, Peplies J, Dubilier N (2007) Species identification of marine invertebrate early stages by whole-larvae *in situ* hybridisation of 18S ribosomal RNA. *Marine Ecology progress Series*, **333**, 103-116.

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System. *Molecular Ecology Notes*, **7**, 355-364.

Remigio EA, Hebert PDN (2003) Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. *Molecular Phylogenetics and Evolution*, **29**, 641-647.

Steinke D, Vences M, Salzburger W, Meyer A (2005) TaxI: a software tool for DNA barcoding using distance methods. *Proceedings of the Royal Society B*, **360**, 1975-1980.

Thomas M, Raharivololoniaina M, Glaw F, Vences M, Vireites DR (2005) Montane Tadpoles in Madagascar: Molecular Identification and Description of the Larval Stages of *Mantidactylus elegans*, *Mantidactylus madecassus*, and *Boophis laurenti* from the Andringitra Massif. *Copeia*, **1**, 174-183.

Thorson G (1935) Studies on the egg-capsules and development of Arctic marine prosobranchs. *Medd Om Gronland*, **100**, 20-73.

Thorson G (1940) Notes on the egg-capsules of some North-Atlantic prosobranchs of the genus *Troschelia*, *Chysodomus*, *Volutopsis*, *Sipho* and *Trophon*. *Videnskabelige Meddelelser fra Dansk naturhistorisk Forening i Kobenhavn*, **104**, 251-265.

Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B*, **360**, 1859-1868.

Victor BC (2007) *Coryphopterus kuna*, a new goby (Perciformes: Gobiidae: Gobiinae) from the western Caribbean, with the identification of the late larval stage and an estimate of the pelagic larval duration. *Zootaxa*, **1526**, 51-61.

4. Barcoding et α-taxonomie intégrative

Comme on l'a vu pour le genre *Eumunida*, le barcoding peut être utilisé comme un outil d'expertise taxonomique, mais également pour tester les anciennes hypothèses de délimitation d'espèces et en proposer de nouvelles. Cependant, par comparaison avec les Turridae, le genre *Eumunida* présente deux avantages : tout d'abord, les hypothèses proposées sur la base des caractères morphologiques se sont révélées généralement robustes ; ensuite, même s'il reste très certainement des espèces à découvrir dans ce genre, la part de cette diversité inconnue est certainement plus faible que dans le cas des Turridae. L'absence d'hypothèse de délimitation d'espèces robustes, liée aux difficultés d'utilisation des caractères morphologiques (Introduction, page 9), et le grand nombre d'espèces encore inconnues dans le cas des Turridae nécessitent la mise en place d'une approche barcoding de type exploratoire, c'est-à-dire qui ne consistent pas à discuter des hypothèses existantes mais à trouver une nouvelle stratégie pour proposer de nouvelles hypothèses de délimitation d'espèces. Pour cela, l'analyse du Barcode COI sera associée à d'autres caractères, dans une analyse d'α-taxonomie intégrative (Ebach & Holdredge 2005 ; Dayrat 2005 ; Will *et al*. 2005). La taxonomie intégrative consiste à analyser plusieurs caractères (morphologiques, moléculaires, caryotypiques, écologiques, géographiques, comportementaux…) en appliquant les différents critères de délimitation qui découlent du LSC (de ressemblance, biologique et phylogénétique). Ces différents aspects ont été abordés dans un grand nombre d'articles, dont voici quelques exemples : Lipscomb *et al*. 2003 ; Proudlove & Wood 2003 ; Sites & Marshall 2003 ; Pons *et al*. 2006 ; Rubinoff 2006 ; Sanders *et al*. 2006 ; Lanzone *et al*. 2007 ; Miller 2007 ; Payne & Sorenson 2007 ; Roe & Sperling 2007 ; Vogler & Monaghan 2007).

Le barcode COI, comme on l'a vu, présente plusieurs avantages (caractère héritable, facilement séquençable,… - C page 57). Sur ce caractère moléculaire, les critères de ressemblance et phylogénétique pourront être appliqués pour proposer des hypothèses primaires de délimitation d'espèces. Ensuite, d'autres caractères et d'autres critères seront analysés pour confirmer ou infirmer ces hypothèses. Les caractères morphologiques, écologiques ou géographiques pourront ainsi être analysés : la mise en évidence de différences sur ces caractères entre deux groupes sera un argument de plus en faveur de la présence de deux espèces différentes. Des méthodes empruntées à la phylogénie ou à la génétique des populations par exemple (Blaxter *et al*. 2005) permettront de tester la monophylie des groupes mise en évidence grâce au gène COI et d'estimer les flux de gènes. Enfin, l'analyse d'autres gènes nucléaires, indépendants du gène COI, permettra de tester grâce au critère généalogique décrit précédemment la présence ou l'absence d'échanges de gènes (recombinaison).

La taxonomie intégrative est donc un processus d'analyse itératif (Figure 8). Les résultats obtenus par l'analyse d'un nouveau caractère et l'application de nouveaux critères viennent s'ajouter aux précédents et permettent de discuter les hypothèses proposées, en les renforçant ou au contraire en les infirmant. Par exemple, dans le cas illustré sur la figure 8, l'analyse du gène COI ne permet pas de trancher entre deux hypothèses : les spécimens bleu clairs et bleu foncés appartiennent-ils à une seule ou à deux espèces différentes ? L'analyse d'autres caractères (comme figurés: morphologie, distribution géographique, mais également autre gène nucléaire,…) permet de rejeter la première hypothèse (une seule espèce), mais pas la seconde (deux espèces différentes).

Les analyses présentées dans le chapitre 3 suivent cette approche de taxonomie intégrative. Elles présentent les différents problèmes auxquels j'ai été confrontés au cours de

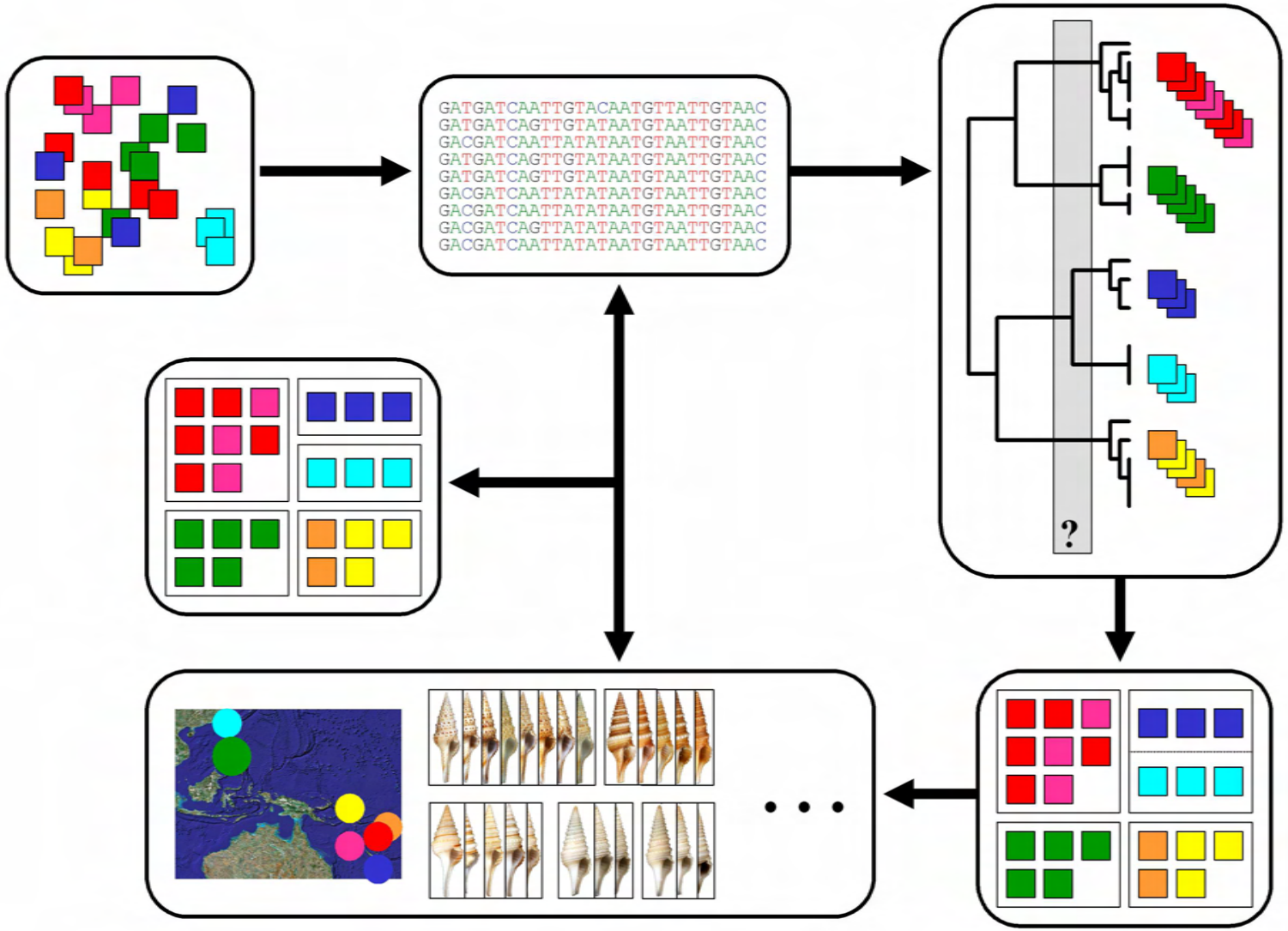


Figure 8 : Méthodologie d'α-taxonomie intégrative adoptée pour la délimitation d'espèces au sein des Conoidea. Les séquences COI sont utilisées en première approche pour proposer des hypothèses primaires de délimitation d'espèces, en utilisant des méthodes basées sur la reconstruction phylogénétique. Ces hypothèses primaires sont ensuite discutées à la lumière des autres caractères disponibles mais également en utilisant d'autres méthodes d'analyses, pour proposer de nouvelles hypothèses de délimitation d'espèces issues de l'analyse intégrative.

ma thèse, et les solutions qui ont été proposées pour les résoudre. J'aborderais ainsi les problèmes liés à l'échantillonnage des spécimens et à leur mise en collection. Cet échantillonnage doit en effet permettre d'estimer correctement la variabilité du groupe analysé. Je détaillerais aussi le choix des caractères analysés. Si le gène COI s'est imposé pour la première étape (hypothèses primaires de délimitation d'espèces), la difficulté a été d'identifier un gène indépendant qui permettrait notamment de mettre en évidence des phénomènes de recombinaison. De plus, l'analyse de la variabilité morphologique a nécessité la mise en place de méthodes formalisées et l'utilisation de caractères morphologiques nouveaux pour les Turridae. Enfin, l'accent sera mis plus particulièrement sur les méthodes qui ont été appliquées pour l'analyse de ces caractères, avec notamment la mise en place d'une nouvelle méthode exploratoire de délimitation d'espèces.

CHAPITRE III

APPLICATION DE L'APPROCHE DE TAXONOMIE INTÉGRATIVE CHEZ LES TURRIDAE

Investigators should not simply decide whether or where to draw lines of demarcation, but rather to present the full picture that research has revealed, and to do so in its full complexity rather than to reduce that complexity artificially.

Hey 2003

Send an unknown Turrid to 10 experts and you may end up with 11 names.

http://www.manandmollusc.net/

Dans ce chapitre sera détaillée la méthodologie de taxonomie intégrative suivie pour quatre groupes de Turridae : la sous-famille des Turrinae, et les genres *Bathytoma*, *Benthomangelia* et *Xenuroturris*. Je reviens tout d'abord sur la stratégie d'échantillonnage, aussi bien en terme d'individus que de caractères analysés, et sur son importance dans une telle approche.

A. Echantillonnages inter-spécifique, intra-spécifique et intra-individuel

1. De l'importance de l'échantillonnage

Dans les premières études barcoding qui ont été réalisées, Janzen *et al*. (2005) rappellent qu'il y « avait une forte tendance du barcoding à traiter quelques séquences comme si elles étaient représentatives d'un "type" ». Il s'agit là d'une des principales critiques qui a été faite à l'encontre du barcoding : pour certains auteurs (par exemple Moritz & Cicero 2004), la faible variabilité du gène COI au sein des espèces qui avaient été analysées (Hebert *et al*. 2004b), n'était qu'un artefact dû au fait que les spécimens analysés ne représentaient qu'une faible partie de la diversité des espèces. En effet, pour analyser la variabilité moléculaire, morphologique, écologique…, il est nécessaire d'estimer correctement la diversité en terme de nombre d'espèces mais également au sein de chaque espèce. La stratégie d'échantillonnage consiste donc à inclure :
(i) tous les variants géographiques d'une espèce, (notamment ceux en bordure d'aire de distribution – Eckert *et al*. 2008) : de cette façon, l'ensemble de la variabilité au sein d'une espèce sera analysée,
(ii) toutes les espèces d'un genre, ou au moins les espèces-sœurs : les spécimens de l'espèce-sœur sont en effet ceux qui sont le plus susceptibles de partager de la variabilité avec l'espèce analysée (Chapitre 2, B page 53). Ils doivent donc être intégrés à l'analyse pour tester si réellement les spécimens d'une même espèce se ressemblent plus entre eux qu'avec n'importe quel autre spécimen (Morando *et al*. 2003 ; Hebert *et al*. 2004b).

Pour mettre en place cette stratégie pour les Turridae, deux problèmes se sont posés. Tout d'abord, dans la plupart des cas, les spécimens collectés lors des missions réalisées avant 2004 étaient conservés dans le formol, puis séchés et mis en collection, pour permettre l'observation et la manipulation des spécimens. Cependant, ce traitement ne permet pas de conserver l'ADN, et ces spécimens ne peuvent donc pas être utilisés dans une analyse moléculaire. Ensuite, il n'existe pas pour les Turridae d'hypothèses robustes de délimitation d'espèces disponibles, et il est donc difficile de savoir à l'avance si l'échantillonnage réalisé est une bonne estimation de la diversité du groupe ou non.

Je présente donc ci-dessous les solutions à ces problèmes qui ont été appliquées lors des campagnes océanographiques organisées par le MNHN et l'IRD depuis 2004.

2. Conservation de l'ADN des spécimens collectés

Pour conserver l'ADN, le tissu doit être conservé dans l'alcool (entre 80 et 100°). Cependant, chez les gastéropodes, la rétractation de l'animal dans la coquille, associée à une fermeture hermétique de celle-ci grâce à un opercule, empêche la bonne fixation de l'ADN. De plus, la coquille de l'animal ne doit pas être endommagée, pour pouvoir être utilisée pour les analyses morphologiques. Pour extraire le tissu de la coquille tout en la conservant intacte, trois méthodes ont été utilisées :
- les spécimens sont plongés dans une solution de Chlorure de Magnésium, qui a pour effet d'anesthésier l'animal. Il peut alors en général être sorti plus facilement de la coquille.

- certains spécimens ne réagissent pas à cette anesthésie (si par exemple l'opercule ferme hermétiquement la coquille), et il est alors nécessaire de percer la coquille pour pouvoir extraire l'animal. Cette opération de perçage de la coquille est plus coûteuse en temps que l'anesthésie.
- les spécimens sont placés dans de l'eau bouillante pendant quelques minutes, ce qui a pour effet de « détacher » l'animal de la coquille. Il est alors facile de placer le tissu en alcool sans endommager la coquille. Cette technique n'a pour l'instant été utilisée que lors de la dernière mission sur le terrain, en 2008. Elle semble plus efficace que l'utilisation de Chlorure de Magnésium, où certains spécimens restaient à l'intérieur de la coquille, mais également plus rapide : quelques minutes dans l'eau bouillante suffisent, alors qu'une ou deux heures étaient nécessaires pour que l'anesthésiant agisse.

Lors des missions, ces trois méthodes permettent en règle générale d'extraire un morceau de tissu de la coquille, tissu dont l'ADN sera alors correctement conservé. Cependant, pour les petits spécimens (inférieure à 5mm), même s'ils ont été anesthésiés ou ébouillantés, l'extraction du tissu de la coquille peut être délicate, et cette étape est réalisée de retour au laboratoire. Dans certains cas, même au laboratoire, il est impossible d'extraire le tissu sans endommager la coquille. Dans ce cas, de nouvelles méthodes appliquées aux insectes pourraient être testées (Rowley *et al*. 2007 ; Hunter *et al*. 2008). Elles consistent à placer l'animal entièrement dans la solution d'extraction, qui ne détruit que l'intérieur de l'animal en laissant la cuticule intacte. Cette technique pourrait être utilisée pour extraire l'ADN du tissu de petits spécimens sans qu'ils aient besoin d'être extraits de la coquille, de la même façon qu'un insecte n'est pas extrait de sa cuticule.

3. Stratégie d'échantillonnage

Lors de la première mission où un échantillonnage en vue de constituer des collections « barcode » a été réalisé (Panglao 2004), lorsque plusieurs spécimens qui possédaient une morphologie similaire, seuls quelques spécimens ont été conservés en alcool. La stratégie suivie lors d'autres campagnes qui ont été réalisées par la suite consistait à l'inverse à conserver tous les spécimens, sans tenir compte de leur morphologie. Les résultats obtenus avec un tel échantillonnage montrent que plusieurs espèces peuvent être délimitées au sein d'un même morphe, indiquant que l'échantillonnage réalisé lors de la mission Panglao 2004 ne permet pas d'estimer correctement la diversité des Turridae.

Pour permettre une estimation correcte de la diversité spécifique, la stratégie qui a été mise en place consiste donc à maximiser le nombre de spécimens récoltés, sans tenir compte des hypothèses de délimitation d'espèces basées sur la morphologie qui pourraient être faites sur le terrain (voir par exemple Meyer & Paulay 2005).

Cependant, la préparation des spécimens pour conserver l'ADN qui doit être faite pendant la mission nécessite plus de temps et de moyens que ce qui était réalisé lors des missions précédentes (où les spécimens n'étaient pas préparés pour des analyses moléculaires). Même si, pour chaque mission, une ou plusieurs personnes sont dédiées à la préparation des spécimens pour les analyses barcoding, les conditions de travail lors des campagnes océanographiques ne sont pas toujours les mêmes. Lors des missions Panglao 2004 et 2005, Santo 2006 et Aurora 2007, l'équipe barcode était installée dans un laboratoire (soit à terre, soit dans le bateau), mais lors des autres missions, les conditions de travail sur le bateau (moins de place et équipe « barcode » restreinte) ne permettaient pas de préparer spécifiquement chaque spécimen. Il a fallu alors trouver un compromis entre « collecter un grand nombre de spécimens » et « préparer les spécimens » tel que je l'ai décrit. La stratégie

qui est suivie maintenant consiste à isoler un morceau de tissu de chaque spécimen lorsque les moyens le permettent (temps et nombre de personnes disponibles), et, dans le cas contraire, de placer tous les spécimens collectés en alcool sans isoler le tissu. De cette façon, même si l'ADN de certains spécimens qui se seraient rétractés dans la coquille ne sera pas correctement conservé, le grand nombre de spécimens disponibles permet d'obtenir un échantillonnage qui permet une estimation de la diversité des espèces.

Concernant le choix des sites géographiques où les spécimens ont été collectés, la stratégie qui a été adoptée est la suivante : la diversité des milieux et des profondeurs prospectées, et ainsi la diversité des espèces récoltées, est maximisée (Figure 9), même si certaines missions sont consacrées exclusivement à l'exploration des milieux profonds (Salomon 2 et 3 par exemple), et d'autres à un type d'environnement particulier (monts sous-marins – EBISCO et Norfolk 2). Cette stratégie a été suivie notamment lors de dix missions pendant lesquelles le matériel utilisé au cours de ma thèse a été récolté, et qui couvrent une aire géographique importante du Pacifique Sud-Ouest, des Philippines au Vanuatu (Figure 10). Ces différentes missions ont permis de collecter quelques 4700 échantillons de Conoidea (dont 4000 Turridae), spécifiquement préparés pour des analyses moléculaires (Annexe).

Figure 9 : Différentes techniques de récoltes utilisées lors de la mission Santo 06. De gauche à droite et de haut en bas : ramassage à vue en rivière, sur la zone de balancement des marées et en plongée sous-marine, aspirateur sous-marin, chalut à perche, drague à main (manipulée par Rudo Von Cosel), brossage de l'épifaune des fonds durs, drague Warén, filet posé sur le fond marin.

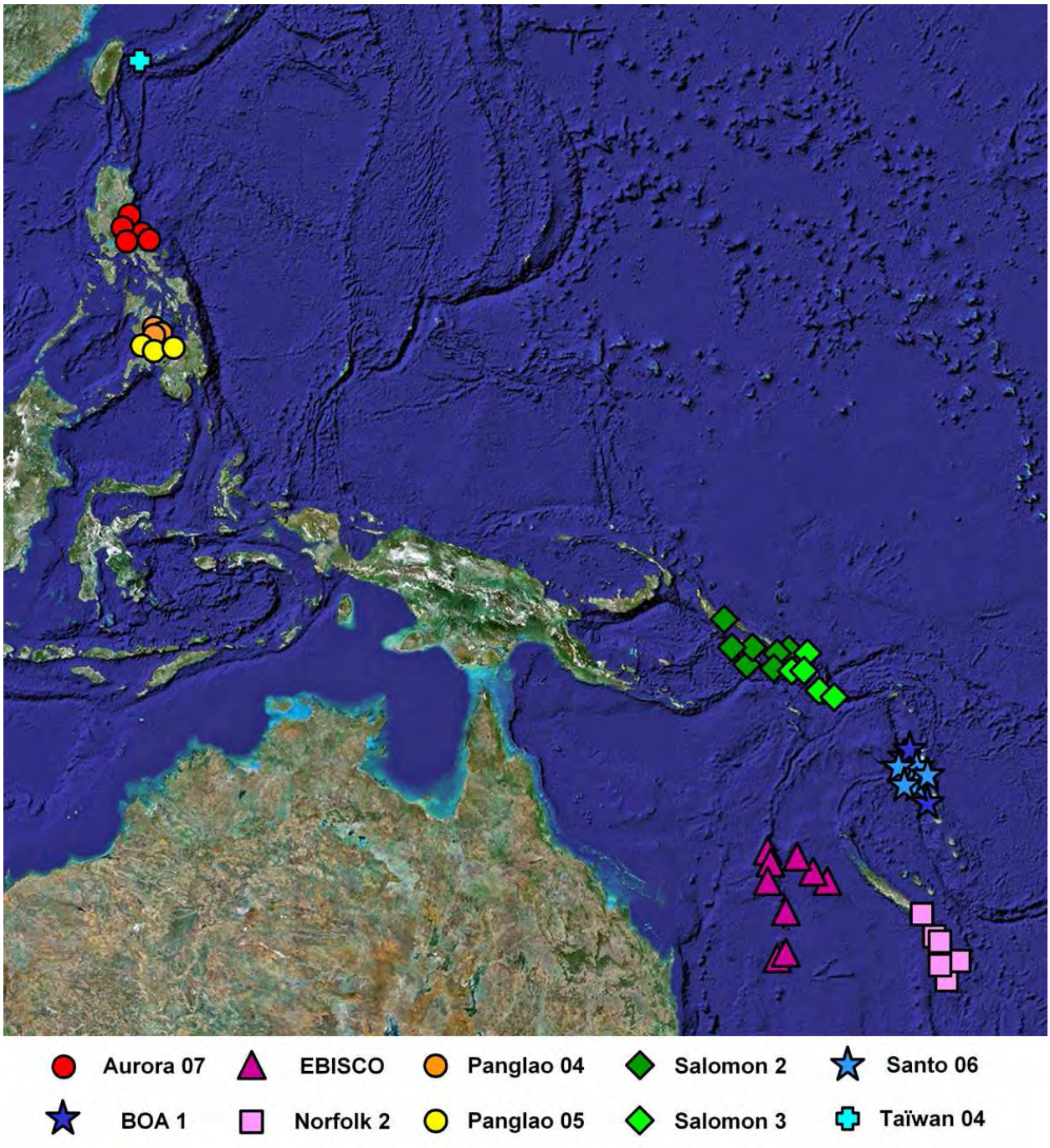


Figure 10 : Répartition géographique des différents points de collecte réalisés lors des 10 missions qui ont permis de constituer dans sa majorité l'échantillonnage utilisé (Annexe). Seules quelques stations sont figurées, la plupart se superposant les unes aux autres à cette échelle.

4. Choix des caractères

Principalement deux types de caractères ont été analysés. Tout d'abord, pour les caractères moléculaires, le barcode COI a été systématiquement séquencé pour tous les individus. Ensuite, pour la plupart des spécimens séquencés pour le gène COI (voir les détails dans les articles 5, 6, 7 et 8), un gène nucléaire a également été séquencé. Le gène 28S a été choisi dans la plupart des cas, d'abord parce qu'il présente des caractéristiques intéressantes, notamment un temps de coalescence rapide (légende de la Figure 6 page 55), mais également

parce qu'il existe, comme pour le COI, des amorces universelles pour ce gène. Cependant, ce gène ne présente pas systématiquement un niveau de variabilité comparable au COI, et plusieurs groupes différents pour le gène COI possèdent exactement la même séquence 28S (Article 5). Dans le cas du genre *Bathytoma*, un autre gène nucléaire, le gène ITS2, a permis de résoudre le problème : son niveau de variabilité est en effet similaire à celui du gène COI. De plus, dans ce genre, aucune variabilité intra-individuelle n'a pu être détectée lors du séquençage, contrairement à ce qui a été reporté pour d'autres taxa (Gandolfi *et al*. 2001). Cependant, le gène ITS2 n'a pas pu être utilisé pour d'autres groupes que le genre *Bathytoma* : c'est en effet le seul groupe taxonomique, parmi ceux que j'ai testé, où le gène ITS2 ne présentait pas de motif (tels que des microsatellites) qui empêchait son séquençage. Il apparaît donc difficile de trouver un gène nucléaire qui serait l'équivalent du gène COI, c'est-à-dire un gène qui présente de la variabilité au niveau intraspécifique et qui de plus est facilement séquençable pour tous les organismes : aucun des deux gènes testés ici (28S et ITS2) ne remplit totalement ces conditions. C'est également le problème rencontré dans plusieurs groupes taxonomiques, comme les plantes ou les champignons, avec le gène COI (champignons ou plantes par exemple ; Tableau 1 page 60)

Pour les analyses moléculaires, l'ADN de 2400 spécimens a été extrait, et 5047 séquences ont été obtenues, dont 1932 pour le gène COI, 1431 pour le gène 28S et 116 pour le gène ITS2 (Annexe).

Pour les caractères morphologiques, deux approches ont été utilisées. Dans le cas du genre *Bathytoma*, mais également pour la sous-famille des Turrinae, les caractères morphologiques ont été utilisés de manière traditionnelle, c'est-à-dire en utilisant des caractères de la coquille qui permettent de former des ensembles d'individus homogènes morphologiquement (morphoespèces). Ce travail a été réalisé à chaque fois par un spécialiste. L'expérience qu'il possède du groupe lui permet en effet de ne pas baser ces hypothèses taxonomiques uniquement sur l'échantillonnage considéré, mais d'intégrer également ce qu'il connaît de la variabilité du groupe à travers tous les spécimens qu'il a pu observer par le passé (Figure 2 page 9). Cependant, ces caractères ne sont pas définis de façon formelle et peuvent varier selon les taxonomistes. Ainsi, il est difficile pour deux taxonomistes qui travailleraient par exemple sur un même groupe d'espèces mais avec deux échantillonnages différents, d'utiliser la même série de caractères pour ensuite comparer leurs hypothèses de délimitation d'espèces. L'analyse morphométrique (EFA : Elliptic Fourier Analysis) réalisée dans le cas du genre *Benthomangelia* permet de formaliser les caractères, qui pourront ainsi être réutilisés par la suite sur un nouvel échantillonnage.

D'autres caractères, tel que la morphologie de la radula (Article 8) ou la distribution géographique et bathymétrique (Article 6 par exemple), ont également été utilisés. Ils sont à chaque fois détaillés dans les différents articles.

Pour l'ensemble des exemples qui seront présentés, la méthodologie d'analyse de la diversité spécifique peut-être divisée en deux étapes. La première étape est exploratoire et permettra de proposer des hypothèses primaires de délimitation d'espèces. Dans le cas de la sous-famille des Turrinae, un groupe particulièrement diversifié, une nouvelle méthode exploratoire est proposée pour permettre une application à un grand nombre d'espèces. La seconde étape permettra d'analyser en détail ces hypothèses primaires pour déterminer si elles correspondent ou non à des espèces. Je présenterai tout d'abord la première étape exploratoire qui a été suivie dans le cas des Turrinae, et je discuterai ensuite des différentes méthodes qui ont été utilisées lors de la deuxième étape, pour chacun des quatre exemples développés (Turrinae, *Xenuroturris*, *Bathytoma*, *Benthomangelia*).

B. 1ère étape : approche exploratoire

1. Choisir une méthode adaptée

L'objectif est ici de proposer des hypothèses primaires de délimitation d'espèces en analysant la variabilité du gène COI. Pour cela, plusieurs méthodes exploratoires existent, mais ne sont pas toutes applicables aux Turridae.

La méthode la plus couramment utilisée est l'approche phylogénétique (par exemple : Pfenninger *et al*. 2006 ; Fu & Zeng. 2008). Dans la plupart des cas, cette approche est utilisée pour tester des hypothèses primaires de délimitation d'espèces déjà existantes. Par exemple, l'analyse phylogénétique permet de tester si un groupe morphologique est monophylétique. Cependant, dans le cas des Turridae, l'objectif n'est pas de tester des hypothèses primaires, mais d'en proposer de nouvelles. Il faut donc dans ce cas que les hypothèses primaires soient définies à partir de l'arbre phylogénétique. La difficulté consiste alors à identifier dans l'arbre phylogénétique à quel niveau se situent les limites entre espèces. La méthode GMYC (General Mixed-Yule Coalescent model), proposée en 2006 par Pons *et al*., permet de placer cette limite dans un arbre phylogénétique, en se basant sur la modélisation des évènements de spéciation et de coalescence.

Les méthodes « Population Aggregation Analysis » (PAA) sont basées sur une délimitation *a priori* de populations, en regroupant les individus selon des critères géographiques et morphologiques (Sites et Marshall 2003 ; Marshall 2006). Ces populations sont ensuite utilisées comme postulat de départ, et sont ensuite regroupées en espèces. Ces méthodes ne seront donc efficaces que si les populations définies *a priori* n'incluent que des spécimens de la même espèce. Si plusieurs espèces sont placées dans une seule et même population, les méthodes PAA ne remettront pas en question cette hypothèse (Ahrens *et al*. 2007). Ces méthodes sont donc contre-indiquées dans le cas des Turridae, où l'objectif est de s'affranchir d'hypothèses *a priori*, notamment basées sur des caractères morphologiques. Les résultats obtenus au moins dans le cas des Turrinae indiquent en effet que plusieurs spécimens qui auraient été placés dans une même population (même site de collecte et morphologie similaire) appartiennent à des espèces différentes.

Il existe également plusieurs autres méthodes, utilisées par exemple en astronomie (Balastegui *et al*. 2001), mais également en sciences médicales (Chacon & Luci 2003) ou en génétique, pour regrouper des types de cellules, des classes de tumeurs ou des gènes selon leur profils d'expression (Eisen *et al*. 1998 ; Jansen *et al*. 2005 ; Madeira & Oliveira 2005 ; Statnikov *et al*. 2005). Ces méthodes peuvent être basées sur la construction d'arbres (phylogénétiques ou non – Eisen *et al*. 1998), sur la méthode des centres mobiles (ou méthode k-means - Steinley 2003 ; Steinley & Brusco 2007) ou sur la théorie des réseaux neuronaux (Kohonen 1995). Ces méthodes permettent de définir K groupes au sein d'un ensemble en se basant sur différents critères : par exemple, pour la méthode k-means, la variance au sein de chacun des groupes est minimisée. En revanche, il est dans la plupart des cas nécessaire de choisir (soit parce qu'il est connu, soit en testant plusieurs valeurs) la valeur de K. Dans le cas des Turridae, l'absence d'hypothèses de délimitation d'espèces ne permet pas de savoir *a priori* combien d'espèces sont présentes dans le jeu de données, et donc quelle est la valeur de K. Cependant, plusieurs critères statistiques ont été proposés pour déterminer le nombre « réel » de classes K (Calinski & Harabasz 1974 ; Fraley & Raftery 1998 ; Wicker *et al*. 2001, 2002), mais ils ne sont pas pour la plupart implémentés pour des approches de délimitation d'espèces.

Pour l'analyse de la sous-famille des Turrinae, deux méthodes exploratoires seront utilisées : la première, notée par la suite DBM (Distance Based Method), se base sur la comparaison des distances génétiques entre et au sein des espèces ; la seconde est la méthode GMYC. Elles permettent de proposer des hypothèses primaires de délimitation d'espèces, et ne nécessitent ni la définition *a priori* de population, ni de préciser le nombre de groupes K.

2. Méthode DBM (Distance Based Method)

La première méthode est basée sur l'utilisation des distances génétiques entre séquences. Une des prédictions du LSC est la ressemblance : deux individus de même espèce se ressembleront plus entre eux qu'avec aucun autre spécimen d'une autre espèce. Autrement dit, la variabilité intraspécifique (au sein d'une espèce) sera plus faible que la variabilité interspécifique (entre les espèces). Cette prédiction est à la base du principe du barcoding, et les analyses « barcoding » qui ont été faites sur différents groupes taxonomiques (Tableau 1) avaient pour objectif de le tester. Pour cela, la méthode la plus couramment utilisée (Hebert *et al*. 2004b) consiste à établir la distribution des distances génétiques mesurées sur le gène COI entre chaque paire de spécimens. Si le postulat est vérifié, la distribution comprend deux modes : le premier mode correspond à la distribution des distances génétiques faibles, c'est-à-dire calculées entre individus de même espèce, le second correspond à celle des distances génétiques élevées, c'est-à-dire calculées entre individus d'espèces différentes. De plus, il peut y avoir une zone de fréquence nulle entre ces deux distributions : le barcode gap.

Ce barcode gap peut être interprété comme une valeur seuil, en deçà de laquelle la distance génétique entre deux spécimens est considérée comme intraspécifique, et au-delà de laquelle la distance génétique entre deux spécimens est considérée comme interspécifique. Plusieurs seuils « standards », c'est-à-dire qui pourraient être applicable à n'importe quel groupe taxonomique, ont été proposés dans la littérature. Ces seuils peuvent correspondre à des valeurs absolues (3 % de divergence par exemple – Smith *et al*. 2005) ou relatives (règle des 10X : le seuil correspond à 10 fois la valeur moyenne de la variabilité intraspécifique – Hebert *et al*. 2004b). Plusieurs problèmes se sont alors posés. Tout d'abord, plusieurs études ont montrées que ce seuil pouvait varier d'un taxon à l'autre, ne permettant pas l'application d'un seuil universel (absolu ou relatif) (Ferguson 2002 ; Holland *et al*. 2004 ; Bichain 2007 ; Gomez *et al*. 2007). Ce résultat a par ailleurs été confirmé lors des analyses des mes différents jeux de données (Articles 5, 6, 7 et 8). Ensuite, dans plusieurs études, la prédiction du LSC n'est pas vérifiée, c'est-à-dire que les distances intraspécifiques ne sont pas toutes inférieures aux distances interspécifiques. Les distributions des distances intraspécifiques et interspécifiques peuvent ainsi se chevaucher, faisant disparaître le barcode gap (Article 5). Ce problème peut être soit lié au gène COI, dont le niveau de variabilité au sein des espèces peut dépasser le niveau de variabilité entre espèces, soit lié à une remise en cause des hypothèses de délimitation d'espèces utilisées au départ (Meyer & Paulay 2005). En effet, si par exemple des individus appartenant à deux espèces différentes ont été considérés comme conspécifiques, les distances qui seront considérées comme intraspécifiques seront en fait du même ordre de grandeur que des distances interspécifiques. Dans ces cas où il n'y a pas de barcode gap, une reconnaissance visuelle de la limite entre variabilités intra et interspécifique devient difficile.

Cependant, même s'il n'existe pas de barcode gap, parce que certaines distances intra et interspécifiques se chevauchent, nous avons montré en collaboration notamment avec Guillaume Achaz (Université Paris 6) qu'il est possible de définir une limite entre les deux

ENCADRE 4

Méthodes d'identification du barcode gap

Afin de formaliser la détermination de la limite entre les distances génétiques intra et interspécifiques, notamment lorsque celles-ci se superposent, nous avons proposé en collaboration avec Guillaume Achaz deux nouvelles méthodes. Elles ont été testées sur le jeu de données correspondant à la sous-famille des Turrinae, et feront l'objet d'une analyse plus complète qui sera publiée prochainement et qui permettra de proposer en ligne un programme d'identification du barcode Gap.

1. Méthode KMGD (K-Means Gap Definition)

La méthode k-means (Figure 11) est une méthode de regroupement largement utilisée, notamment en biologie. Elle peut être appliquée dans notre cas pour délimiter deux groupes au sein de l'ensemble des distances entre chaque paire de séquences, l'un correspondant à des distances génétiques faibles (et donc à des distances entre spécimens de la même espèce) et l'autre à des distances génétiques élevées (et donc à des distances entre spécimens d'espèces différentes). Ces deux groupes peuvent cependant être de variances très différentes, l'un incluant par exemple des distances génétiques comprises entre 0 et 3 % (valeurs observées pour plusieurs taxa – Smith *et al.* 2005), l'autre des distances génétiques supérieures à 3 % (et jusqu'à 25 % dans le cas des Turrinae par exemple). Ce déséquilibre entre les variances peut entraîner un biais, car la méthode a tendance à déplacer la limite entre les deux groupes de façon à réduire la variabilité au sein du groupe le plus variable (Relative Cluster Density bias - Steinley & Brusco 2007). Pour éviter ce biais, la variance du second groupe a été limitée en utilisant la racine carrée des distances génétiques. Nous avons aussi utilisé une méthode k-means peu sensible à ce biais (Steinley 2003 ; Steinley & Brusco 2007). Un intervalle de confiance à 99 % a de plus été déterminé par la méthode des bootstraps.

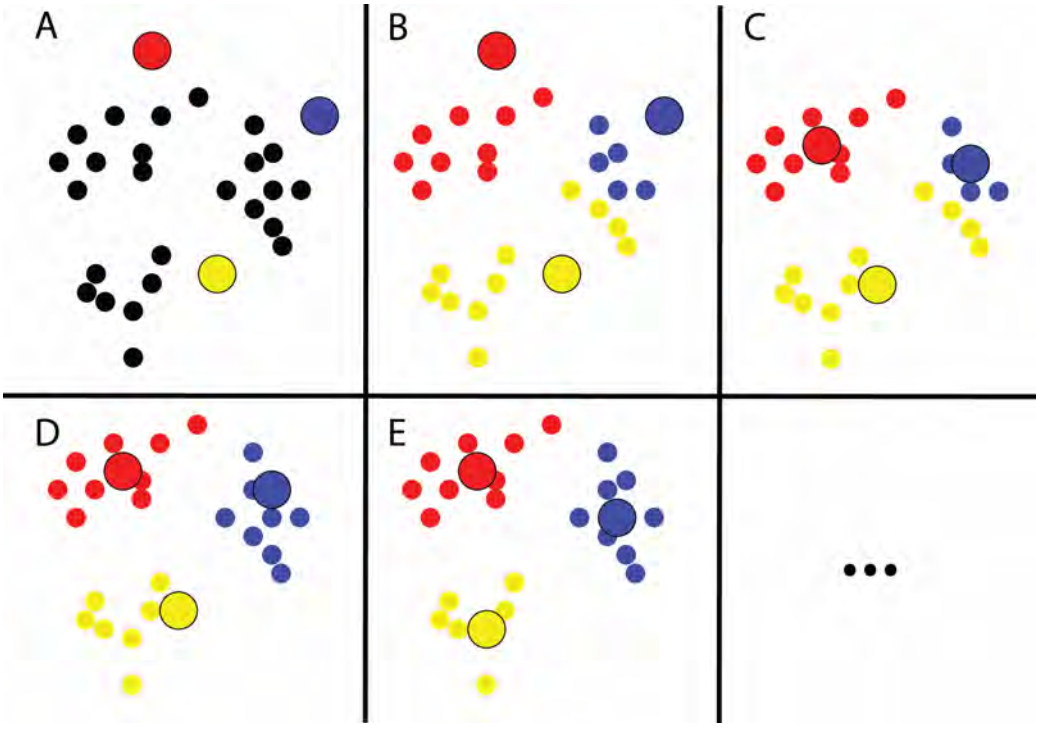


Figure 11 : Principe de la méthode K-means. Pour un jeu de données constitué de trois groupes, trois points sont choisis au hasard et constituent les trois centres initiaux (A). La distance entre chacun des points et les trois centres est calculée, et les points sont affiliés au centre dont ils sont le plus proches pour former trois groupes (B). Un nouveau centre, correspondant au centre de gravité, est défini pour chacun des groupes (C), et les distances entre chaque point et ces nouveaux centres sont calculées. De la même façon les points sont affiliés à trois groupes (D), et le processus (définition d'un centre, calcul des distances, affiliation - E) est répété jusqu'à ce que les centres ne bougent plus.

2. Méthode PGD (Peak Gap Definition)

Plutôt que d'établir la distribution des distances génétiques calculées entre chaque paire de spécimens (Figure 12A), ces mêmes distances génétiques peuvent être classées par ordre croissant (Figure 12B). Cela revient donc à ranger les distances de valeur K et à déterminer leur rang (i) L'augmentation brusque des distances génétiques sur la figure 12B correspond alors à la zone où les fréquences sont les plus faibles sur la figure 12A, autrement dit le barcode gap. Sur la courbe suivante (Figure 12C), qui correspond à la dérivée de la courbe précédente (Figure 12B), cette augmentation brusque (point d'inflexion de la courbe) correspond à la valeur maximale de la dérivée.

Dans le cas illustré ci-dessous, la méthode permet d'estimer que la distance génétique correspondant à la valeur maximale de la dérivée (le « pic » de la dérivée) est 0.032. Comme pour la méthode KMGD, un intervalle de confiance à 99 % a été déterminé par la méthode des bootstraps.

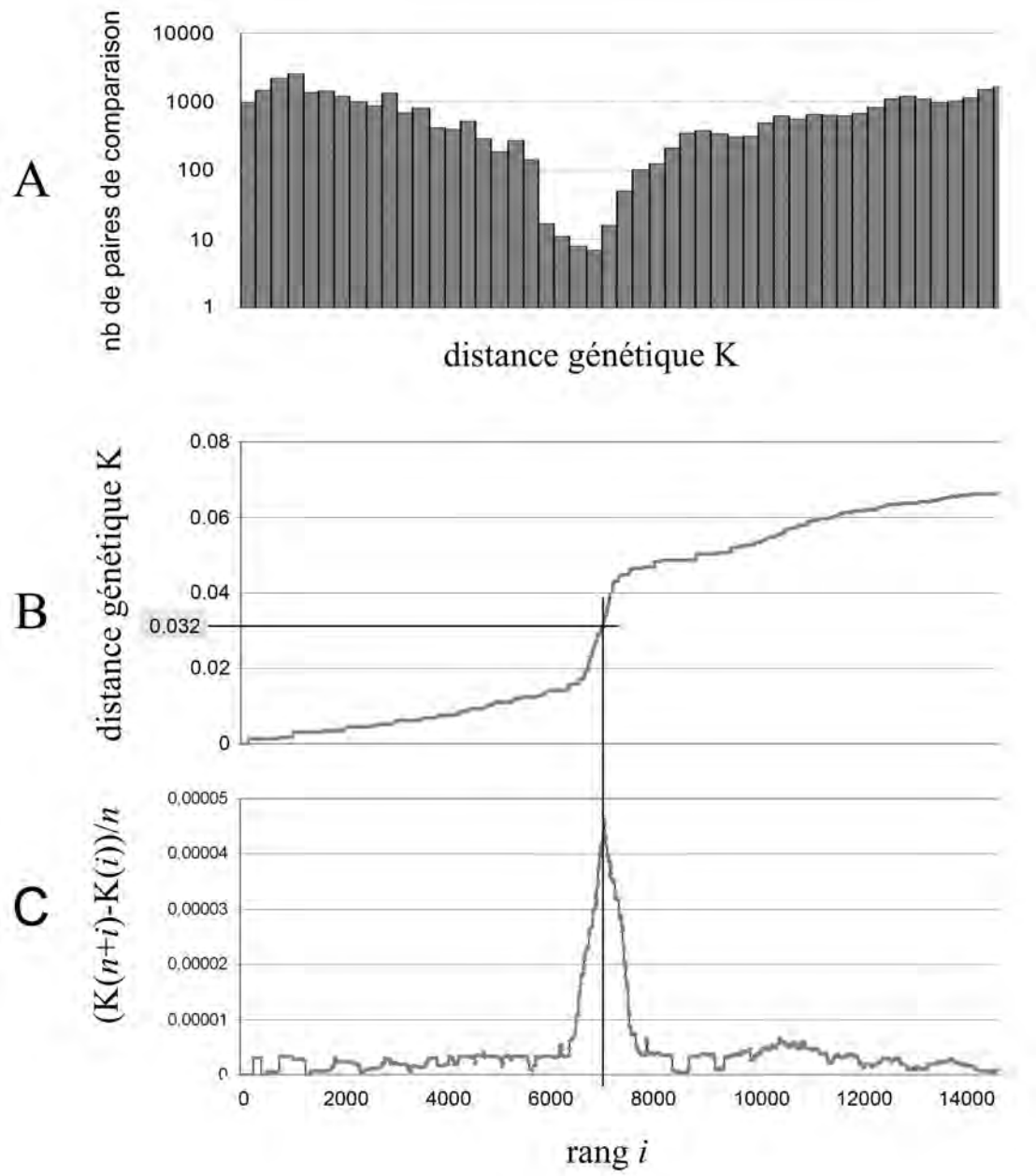


Figure 12 : A. Distribution des distances génétiques. B. Chaque distance génétique K est associée à un rang i. C. La dérivée de K est calculée en utilisant une valeur de n suffisamment grande pour éviter de mettre en évidence des pics locaux, mais suffisamment petite pour ne pas lisser le pic correspondant au barcode gap.

distributions (distances faibles intraspécifiques et distances élevées interspécifiques). Nous avons développé deux méthodes qui permettent de rationaliser la recherche de cette limite entre les deux distributions (Encadré 4). L'identification de cette limite permet ensuite d'attribuer facilement l'ensemble des distances génétiques calculées sur le jeu de données à l'une des deux catégories : distances intra ou interspécifiques.

3. Méthode GMYC (General Mixed-Yule Coalescent model)

La seconde méthode (GMYC – General Mixed-Yule Coalescent model) a été proposée récemment par Pons *et al*. (2006). Elle repose sur l'observation qui a été faite par Nee *et al*. (1992, 1994) à partir des résultats obtenus par analyses phylogénétiques : le rythme de branchements dans un arbre phylogénétique accélère lors de la transition entre les branchements de type « évènements de spéciations » (entre les espèces) et les branchements de type « coalescence » (au sein des espèces). La méthode GMYC repose sur la construction d'un arbre phylogénétique et sur la modélisation de ces deux catégories d'évènements (modèle de Yule entre les espèces et modèle de coalescence au sein des espèces). Les conditions d'application de ces deux modèles (par exemple, générations non chevauchantes pour le modèle de Yule, ou absence de sélection pour le modèle de coalescence) ne sont pas obligatoirement respectées, et les auteurs de la méthode travaillent actuellement à l'amélioration de la méthode en intégrant des modèles plus réalistes (Barraclough, com. pers.).

Ensuite, deux hypothèses alternatives sont testées en comparant leurs vraisemblances : soit il n'existe qu'un seul type de branchement dans l'arbre, soit il existe deux types de branchement. Si la vraisemblance de la seconde hypothèse est la plus élevée, la limite entre les deux types d'évènements est déterminée et positionnée sur l'arbre phylogénétique. Cette limite permet ainsi de définir une partition de l'arbre en espèces.

4. Biais potentiels pour ces deux méthodes

Ces deux méthodes (DBM et GMYC) présentent deux principales limites. Tout d'abord, elles se basent toutes deux sur une évaluation de la divergence entre les spécimens analysés. Pour estimer cette divergence, il existe plusieurs modèles de mutations, du plus simple, où toutes les mutations sont équiprobables, au plus compliqué, où chaque mutation a une probabilité différente des autres, et où d'autres paramètres (nombre de sites nucléotidiques invariants et hétérogénéité des taux de mutation selon les sites nucléotidiques) entrent également en jeu. Les modèles de mutation plus compliqués permettent de tenir compte des phénomènes d'homoplasie et d'estimer le nombre réel de différences entre deux spécimens. Dans la plupart des analyses de barcoding qui ont été réalisées jusqu'à présent (par exemple : Hajibabaei *et al*. 2006), un modèle simple (Kimura 2 Parameter) est utilisé. Pour l'analyse des Turrinae, la comparaison de trois modèles a permis de montrer que l'utilisation d'un modèle adapté au jeu de données analysé (Article 5) permet de prendre en compte les biais dus à l'homoplasie : avec ce modèle, la distribution des distances génétiques dans une lignée est déterminée de manière plus précise. Les méthodes présentées précédemment (KMGD et PGD) permettront donc une meilleure identification du barcode gap.

Ensuite, les deux méthodes assument que la limite entre distances intra et interspécifique pour la méthode DBM ou entre évènements de spéciations et de coalescence pour la méthode GMYC est unique pour l'ensemble du jeu de données. Cette limite peut cependant varier entre les espèces d'un même groupe (Ahrens *et al*. 2007 ; De Queiroz 2007). Considérer que cette limite est identique pour l'ensemble du groupe analysé suppose que les

espèces ont à peu près toutes le même âge, et qu'elles présentent toutes une vitesse de coalescence (et donc une taille efficace et des pressions de sélection) équivalente. Ces deux hypothèses sont peu susceptibles d'être vérifiées (Tautz *et al*. 2003 ; Moritz & Cicero 2004 ; Mallet *et al*. 2005 ; Hickerson *et al*. 2006). Ce biais peut être modélisé, et les résultats qui ont été obtenus sur le jeu de données des Turrinae sont en accord avec le modèle (Encadré 5). La limite « distances intra/interspécifiques » obtenue en analysant l'ensemble de la sous-famille des Turrinae a été comparée à la limite obtenue sur différents groupes au sein de la sous-famille. Les limites sont différentes entre les groupes, suggérant que l'âge des espèces et/ou les vitesses de coalescence n'étaient pas équivalentes pour l'ensemble des espèces incluses dans cette analyse (Encadré 5).

Prendre en compte les contraintes phylogénétiques liées à l'homoplasie et à l'hétérogénéité des temps de spéciation et des vitesses de coalescence s'avère donc nécessaire pour estimer la limite entre variabilité intra et interspécifique. Comme il a été montré, une solution serait d'analyser des groupes d'espèces homogènes, pour lesquels l'âge et les vitesses d'évolution sont équivalents d'une espèce à l'autre. Limiter la diversité du groupe analysé, en travaillant par exemple sur un genre monophylétique, permettrait de limiter l'hétérogénéité des valeurs des paramètres au sein de ce groupe. De plus, l'utilisation de modèles de mutation non plus adaptés à l'ensemble du groupe mais à chacun des sous-groupes analysés permet également d'améliorer l'estimation des divergences entre spécimens.

Cette approche exploratoire a été appliquée à la sous-famille des Turrinae (Article 5). 1000 spécimens attribués aux genres *Turris*, *Xenuroturris*, *Lophiotoma*, *Gemmula* et *Turridrupa*, reconnus comme inclus dans la sous-famille des Turrinae (Chapitre 1) ont été séquencés pour le gène COI. Les deux genres *Gemmula* et *Lophiotoma*, qui regroupent 90% de la diversité de ces cinq genres, ne sont pas monophylétiques, et incluent vraisemblablement les trois autres (Heralde *et al*. 2007). L'analyse sera donc réalisée sur l'ensemble de la sous-famille, et non sur chacun des cinq genres. Selon les méthodes et selon les jeux de données utilisés, entre 78 et 102 entités sont reconnues. Ces hypothèses primaires de délimitation d'espèces ont ensuite été testées à l'aide d'autres caractères et méthodes décrits ci-dessous.

ENCADRE 5

Influence de l'hétérogénéité de l'évolution des espèces sur le barcode gap

Comme modélisé sur la figure 13A, les distances génétiques intra et interspécifiques sont influencées par le temps de spéciation (Tautz *et al.* 2003 ; Hickerson *et al.* 2006) et la taille efficace des espèces (Moritz & Cicero 2004 ; Mallet *et al.* 2005). Un temps de spéciation plus ou moins court aura pour conséquence des distances entre les espèces plus ou moins importantes. Une taille efficace plus ou moins faible influera sur le temps de coalescence, ce qui conduira, pour un taux de mutation donné, à un niveau de variabilité faible si le temps de coalescence est rapide ou fort s'il est plus lent. De plus, une augmentation ou une diminution des distances intraspécifiques entraînera respectivement une diminution ou une augmentation des distances interspécifiques.

Par conséquent, les distances interspécifiques d'espèces récentes peuvent être confondues dans la distribution des distances intraspécifiques d'espèces plus vieilles et de plus grande taille ; inversement, les distances intraspécifiques d'espèces anciennes et de grande taille efficace peuvent être confondues dans la distribution des distances interspécifiques d'espèces récentes (De Queiroz 2007, Ahrens *et al.* 2007).

Certains résultats obtenus à partir de l'analyse de la sous-famille des Turrinae (Figure 13B) sont en accord avec le modèle présenté sur la figure 13A. L'histogramme des distances entre tous les individus d'un clade de Turrinae ne présentait pas de barcode gap. L'analyse séparée de deux sous-clades a permis de détecter un barcode gap pour chacun d'entre eux, mais à des positions différentes dans la distribution des distances génétiques. Selon le modèle (Figure 13A), la présence d'un barcode gap est compatible avec l'hypothèse que ce sous-clade inclut des espèces dont les âges et les vitesses de coalescence sont homogènes.

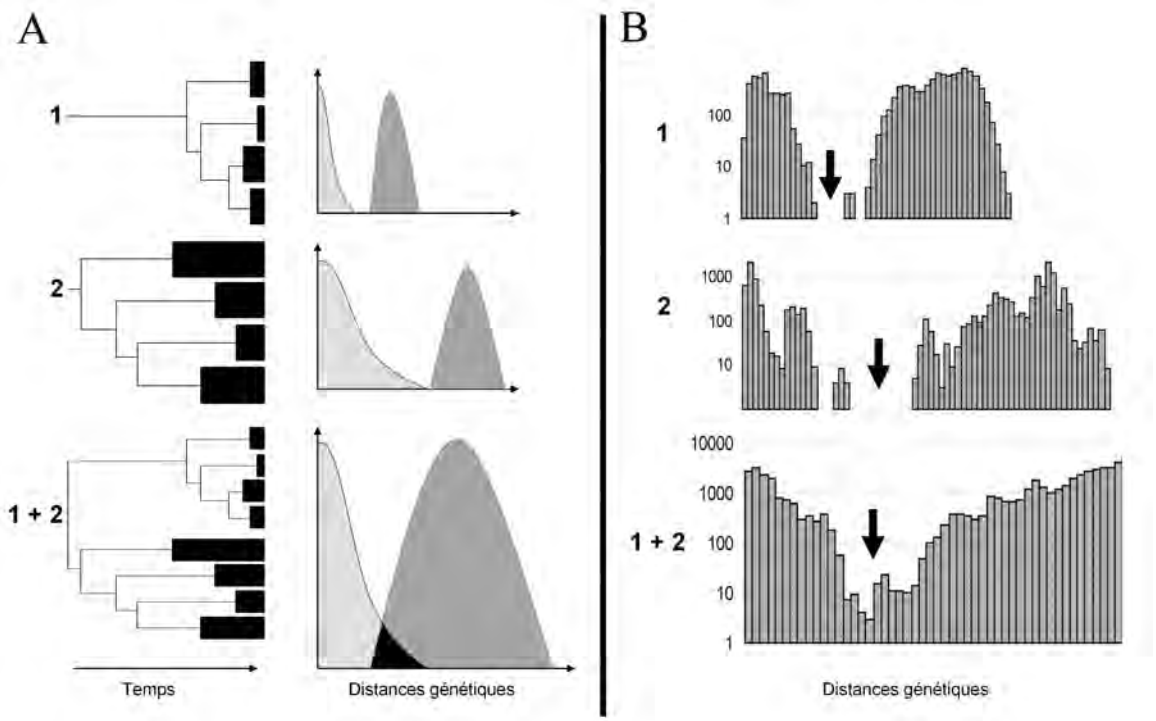


Figure 13 : A. 2 modèles de divergence entre espèces. Le premier modèle correspond à des espèces récentes de taille efficace faible. L'histogramme des distances génétiques est donc composé de distances intraspécifiques faibles séparées par un barcode gap des distances interspécifiques plus élevées. Le second modèle correspond à des espèces plus anciennes de taille efficace élevée. Par rapport au jeu de données 1, les distances intraspécifiques et interspécifiques sont plus élevées, mais restent séparées par un barcode gap. Lorsque les deux modèles sont analysés ensemble, les distances intra et interspécifiques se superposent en partie (en noir).
B. Exemple extrait du jeu de données « Turrinae ». Les seuils entre distances intra et interspécifiques, déterminées par les méthodes KMGD et PGD, sont indiqués par la flèche noire. Les histogrammes B1 et B2 présentent chacun un barcode gap, mais ils correspondent à deux valeurs de distances génétiques différentes. Cet exemple est issu de l'article 5 (clade I et sous-clades IA et IB).

An integrative framework for species delimitation: a case studies in a hyperdiverse group of marine molluscs, the Turrinae (Gastropoda, Conoidea)

N. Puillandre[1,2,*], M.-C. Boisselier[1], Philippe Bouchet[2], C. Cruaud[3], S. Samadi[1]

[1] Université Pierre et Marie Curie (UPMC), UMR 7138, Systématique, adaptation, évolution (UPMC/IRD/MNHN/CNRS), Paris, France.

[2] Muséum National d'Histoire Naturelle (MNHN), Departement Systematique et Evolution, UMR 7138 Systématique, Adaptation, Evolution, 55 rue Buffon, 75231 Paris Cedex 05, France

[3] GENOSCOPE, Centre National de Séquençage, Evry, France

* Corresponding author: UMR 7138, Systématique, adaptation, évolution, UPMC; IRD; MNHN; CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.; puillandre@mnhn.fr; +33 1 40 79 37 43

En préparation ; pour soumission à Systematic Biology

Abstract

Integrative taxonomy, combining molecular, morphological ecological and geographical data is now advocated to propose robust hypotheses of species delimitation. The classical methodology consists in proposing primary hypotheses of species delimitation, then tested with several characters and criteria. However, in hyperdiverse groups, such as the subfamily Turrinae (Mollusca: Conoidea), defining primary hypotheses of species delimitation can be difficult. We propose here to use the animal DNA Barcode, the COI gene, to propose primary hypotheses of species delimitation using a dataset of 1000 specimens collected in the South-West Pacific. For this, two exploratory methods are applied: the one proposed in 2006 par Pons et al., and a new one, based on the recognition of the limit between intra and interspecific distances in the genetic pairwise distances distribution. We also propose to take into account the phylogenetic constraints, by using models of mutation adapted to the dataset, but also by comparing results obtained with the whole subfamily vs results obtained with clades that include a more limited and a more closely related diversity. The primary hypotheses of species delimitation obtained are then tested with other characters (a nuclear gene, the rRNA 28S, but also morphological characters, geographical and bathymetrical distribution) and criteria (monophyly and gene flow). Eighty-eight species are delimited, of which only 28 are linked to a species name, thus validating the use of exploratory methods in delimiting species

Introduction

Species delimitation problematic are back on the front of the stage thanks to the new impetus provided by recent advances in taxonomy (Agapow 2004; Wiens 2007). For example, the global COI sequencing effort currently held through the Barcode initiatives contributes to the detection of new taxa, even in well-studied groups. Indeed, even in taxa such as birds or Lepidoptera hidden diversity is discovered (Hebert et al. 2004; Hajibabaei et al. 2006). As repeatedly quoted, COI sequences are of course not sufficient to provide satisfactory species delimitation (Moritz & Cicero 2004), but is a powerful tool to quickly identify species complex (Smith et al. 2005; Hebert et al. 2004, but see Brower 2006; Savolainen 2005). Consequently, the taxonomists' community is now taking advantage of these data, often representing huge datasets, to enhance its field of research and to develop new tools of species delimitation, including molecular data (Vogler & Monaghan 2007). This is particularly true for the underexplored parts of biodiversity such as marine environments. Indeed, marine realm is since the few last decades the main source of taxonomic discoveries and marine biologists have repeatedly emphasized that the barcoding effort will enhance biodiversity exploration (Schander 2005, Meyer & Paulay 2005, Gomez et al. 2007).

Alpha-taxonomy also benefits from the clarification of the apparently endless debates about the species concepts and their apparent incompatibility. A consensus seems to be reached: the conceptual framework first proposed by de Quieroz (1998), that is now largely accepted (Samadi & Barberouse 2006, Knowles and Carstens 2007, Rissler and Apodaca 2007 or Giraud et al. 2008), emphasized the distinction between the species definition, in which species are viewed as definitively diverging lineages, and the operational criteria that are used to delimit species. An important point that worth to be underlined is that species should be considered hypotheses that may be validated or rejected using additional evidence. The taxonomist is thus engaged in a process of modification/validation as new data are available and/or new criteria are applied. This practice emphasizing the need to use several criteria, each appealing a different species property (Marshall 2006), thus justifying the integrative perspective.

In this conceptual framework, many authors advocate that DNA barcodes data must be associated with all available source of evidence to propose robust hypotheses of species delimitation Will 2005, (Rubinoff 2006, Vogler & Monaghan 2007; Monaghan et al. 2006; De Salle 2006; Miller 2007; Wiens 2007). The methodology of integrative taxonomy generally consists on two steps, even if not always clearly stated: firstly, hypotheses of species delimitation are proposed (e.g. Fritz et al. 2008, Rezac et al. 2008, Mariaux et al. 2008, Savage & Watling 2008), and secondly, these primary hypotheses are validated by other criteria and characters. In this second step, multiple gene genealogies are encouraged, as a safeguard against the frequent well-known bias encountered when using a single gene (Nichols 2001; Funk & Omland 2004).

The primary hypotheses proposed in the first step can correspond to species that are already described in literature, or to morphospecies, geographic populations, and even clades recognized in a phylogenetic tree. However, in some groups, defining such primary hypotheses of species delimitation is not trivial. This is for example the case in the Turrinae (Mollusca, Conoidea), a group of predatory marine molluscs, present from shallow waters to abysses, and worldwide distributed. Firstly, compared to the total diversity of the group, only a few hypotheses of species delimitation are available in literature. Secondly, morphological variability is difficult to use because determinism of morphological characters (i.e. shape and ornamentation of the shell) is generally not known. Thus, it may lead to the recognition of erroneous species limits, based on characters covarying with the environment and not with species delimitations (Bouchet & Warén 1980, Pfenninger et al. 2006, Kantor et al.

unpublished results). Thirdly, the Turrinae is a hyperdiverse groups, and cryptic species are supposed to co-occur sympatrically (Kantor et al. unpublished): defining geographic populations will lead to inconsistent primary hypotheses. To overcome such difficulties, molecular characters, such as the animal barcode (COI gene), associated to phylogenetic approach are commonly used (e.g. Fu et al. 2008). However, the problem consists in that case in identifying the limit between the species in the tree without any other information (names, morphology, geography…) than the molecular character itself.
Thus, for groups such as the Turrinae, the characters that should be used, at least to define primary hypotheses of species delimitation, is the DNA, and the methods that should be employed must be exploratory, *i.e.* based only on DNA data and freed of any morphologically- or geographically-based postulates. For example, methods such as Population Aggregation Analysis (PAA - reviewed in Sites & Marshall 2003; see also Marshall 2006), is clearly not indicated, as they require a prior delimitation of populations, consisting on morphologically similar specimens collected in the same place (Ahrens et al. 2007).
Our goal is thus to use methods that allow proposing robust primary hypotheses of species delimitation without any prior assumptions. We propose here a new formalized method (DBM, for Distance-Based Method) where each COI sequence obtained for 1000 specimens of Turrinae was considered independently from all others data linked to the specimen, such as geographic localisation or shell morphology. In this way, the methodology that we followed can be considered exploratory, *i.e.* primary hypotheses of species delimitation will be proposed by considering only the COI variability. This method is based on the similarity criterion (two specimens from the same species are more similar than two specimens from different species), and relies on the recognition of the limit between intra and interspecific variability in the genetic pairwise distance distribution (Hebert et al. 2003). This method will also be compared to the exploratory method proposed by Pons et al. (2006), referred afterward as the GMYC method (General Mixed Yule Coalescent model). In a second step, the primary hypotheses of species delimitation obtained with the DBM and GMYC methods will be tested with a nuclear gene, the 28S, to check for incongruence between these two independent markers. Several other data, such as morphology or geography, will also participate in the elaboration of robust hypotheses of species delimitation in this integrative framework.

Material and Methods

Sampling

The specimens of Turrinae were collected between 2004 and 2007 in several deep-sea cruises conducted by the Muséum National d'Histoire Naturelle (MNHN) and the Institut de Recherche pour le Développement (IRD) in Taiwan, Philippines, Solomon Islands, Vanuatu, Chesterfield Islands and Norfolk ridge (South of New-Caledonia). Material for molecular studies had been preserved in 95% ethanol by clipping pieces of the head-foot from anaesthetized specimens, and shells were kept intact for morphological characters analysis. All specimens susceptible to belong to the subfamily Turrinae were collected, without taking into account any kind of *a priori* species or population delimitation (Tautz 2003; Vogler & Monaghan 2007). This should lead to the inclusion of several specimens for each species (depending on their abundance, but not on an *a priori* sorting), and also on the inclusion of potential cryptic species. One thousand specimens were used for molecular analyses, and for each of them data corresponding to their sampling site (geographic coordinates, depth of

collection) were databased (BOLD projects "Conoidea barcodes and taxonomy"). All specimens are conserved in MNHN collection.
The Turrinae are particularly diverse in the sampled area. In the Indo-Pacific, Powell (1964) listed 76 Recent species, allocated to 11 genera. Only for the genus *Gemmula*, Tucker (2004) recognized 49 valid species, of which 20 are present in the Indo-Pacific (Powell 1964). However, this genus is supposed to include much more species (pers. com. A. Sysoev). When possible, specimens were linked to an available name in the literature. These name attributions were made by several specialists of the group (Alexander Sysoev, Baldomero Olivera, Yuri Kantor) on the basis of morphological characters but also on the geographic distributions. For several specimens, only a genus name associated to a number is provided. In some of these cases, a putative species identification is proposed (noted "cf.").

Sequencing

DNA was extracted from a piece of foot, using 6100 Nucleic Acid Prepstation system (Applied Biosystem). Two gene fragments were amplified: (i) a fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene using universal primers LCO1490 and HCO2198 (Folmer et al., 1994) and (ii) a fragment of 900 bp of the rDNA 28S gene, involving D1, D2 and D3 domains (Hassouna et al., 1984), using the primers C1 and D3 (Jovelin & Justine, 2001). For COI gene, the primer LCO1490 was also used in combination with a newly designed primer (COIH615: CGAAATYTNAATACNGCYTTTTTTGA) when PCR were negative with HCO2198. All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq (MPBiomedicals) for all genes. COI gene amplifications are performed according to Hebert et al. (2003); for 28S gene, they consist in an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 52°C and extension at 72°C for 1'. The final extension was at 72°C for 10'. PCR products were purified and sequenced by the Genoscope (genbank accession numbers EU015659, EU015661, EU015664, EU015677, EU015681, EU015682, EU015684, EU015724, EU127874-EU127882, EU820248-EU821230 for COI gene and EU015543, EU015545, EU015548, EU015562, EU015566, EU015567, EU015569, EU015609, EU127883-EU127891, EU819556-EU820247 for 28S gene). In all cases, both directions were sequenced to confirm accuracy of each haplotype sequence.

Exploratory methods: primary hypotheses of species delimitation

Distance Based Method (DBM)

Genetic distances will be calculated between each pair of COI sequences. In order to evaluate the effect of multiple changes on the distance between DNA sequences (Steinke et al. 2005), three genetic distances will be compared: (i) the uncorrected p-distance, (ii) the K2P distance, a simple model that slightly correct for multiple changes but frequently used in DNA barcode analyses (Hajibabaei et al. 2007) and (iii) the best-fitting distance (*i.e*. that corrects optimally for multiple changes) as determined by Modeltest (Posada & Crandall, 2001), in conjunction with PAUP 4.0b10 (Swofford, 2002), following the AIC criterion. Results are visualized on a genetic pairwise distances histogram. Following the similarity criterion, genetic distances between specimens from the same species are supposed to be inferior to genetic distances between specimens from different species, revealing a non-continuous distribution (Hebert et al. 2004; Ward et al. 2005; Barret & Hebert 2005; Hajibabaei et al. 2006, Sites & Marshall 2004). This barcode gap, *i.e.* the range of genetic distances not represented in the matrix of pairwise comparisons, can be used as a threshold offering primary species delimitation under

the assumption that individuals within species are more similar that between species (Genotyping clustering - Mallet 1995).
However, in some cases, this barcode gap does not correspond to a real discontinuity in the distribution, but only to a decrease of the distance frequency between the two modes of the distribution, *i.e.* the intra and interspecific distances overlap. This can be due to intrinsic problems such as incomplete lineage sorting, where the COI sequence of a specimen is more similar to the sequence of another species than to sequence of the same species (Salter Kubatko & Degnan 2007; Rosenberg & Tao 2008) or methodology-based ones, such as underestimation of genetic distances because of homoplasy. To overcome this difficulty, we propose two new formalized methods of "gap definition" that allow an estimation of the limit between both distributions, even if they are not clearly separated by a gap. The first method is based on the k-means methods (KMGD for K-Means Gap Definition), and the second one will be referred as the PGD (Peak Gap Definition):
- KMGD: the whole set of genetic distances can be considered a group of pairs of specimens characterized by one variable, the genetic distance between both specimens. Low genetic distances (intraspecific distances) and high genetic distances (interspecific distances) constitute two groups that can be identified by the k-means method. However, these two groups can display different variances, one including for example genetic distances between 0 and 3% (e.g. Smith et al. 2005), the other genetic distances from 3 to up to 20% between species from different genera. This disequilibrium can induce a bias in the result of the k-means method (Relative Cluster Density bias, Steinley & Brusco 2007). To correct it, we limited the variance within the group of high genetic distances by using the square root of the genetic distances. We also used an initialization strategy of the k-means method implemented by Steinley (2003), supposed to be weakly sensible to the Relative Cluster Density bias (Steinley & Brusco 2007). We further performed 1000 bootstraps on the whole dataset to compute a 99% confidence interval for the cut-off distance.
- PGD: Each of the ordered pairwise genetic distances $K(i)$ is associated with a rank i. The curve of the $K(i)$ values show a sudden increase in the vicinity of the limit between the two modes of the genetic pairwise distances distribution (even if this does not correspond to a real gap). This method detects this sudden increase by finding the maximum of the derivative $dK(i)/di$. We estimated the derivative by computing the difference $(K(n+i)-K(i))/n$ for a given smoothing windows size n. The estimated derivative typically peaks at a rank that corresponds to the limit between the two modes of the genetic pairwise distances distribution. Too small window size n makes the derivative too rugged to capture the summit of the peak properly, whereas too large n leads to a large flat peak that encompasses many local peaks. Therefore, we tested increasing values of n until the estimated summit of the peak is identical for several successive n. We report the "cut-off" distance that is associated with this peak that split the distances between intra- and inter-species. As for the KMGD method, 1000 bootstraps were performed using the chosen n. The implementation of the method is available upon request to the authors.

General Mixed Yule Coalescent model (GMYC)

The GMYC method is described by Pons et al. (2006). It is based on the difference in branching rates between speciation branching events (interspecific relationships) and coalescence branching events (intraspecific relationships) in the tree. This difference can be visualized as a switch between slow and fast rate of branching events in a lineage-through-time plot, as described in Nee et al. (1992). A phylogenetic tree is obtained using bayesian analysis, consisting of two Markov chains (20000000 generations, with a sampling frequency of one tree each thousand generations and four swaps at each sampling, and a temperature of 0.5) run in six parallel analyses each using Mr. Bayes (Huelsenbeck et al., 2001). When the

log-likelihood scores is found to stabilize, a consensus tree is calculated after omitting the first 25% trees as burn-in. As the tree used for the analysis must be a completely resolved and ultrametric tree, all multifurcations are artificially resolved with 0-branch length and the tree is ultrametricized using the Mesquite V2.01 software (Maddison & Maddison). The GMYC (General Mixed Yule Coalescent) model, implemented in R and provided by the authors, is run on the tree. The first step is to compare the likelihood of the tree assuming a single branching process to the likelihood of the tree assuming a switch of branching rates between the two types of events. If such a switch is detected, its position is determined and placed in the tree, allowing the delineation of clusters.

Potential bias of the exploratory methods

Population size, rate of speciation and time of divergence may vary from a species to another, leading to heterogeneity of the "speciation threshold" through the entire dataset. We thus applied both exploratory methods on subgroups drawn from the entire dataset. The presence of heterogeneous time of speciation of effective size would lead to the recognition of different limits with the DBM and GMYC methods among these subgroups. These subgroups correspond to clades defined by the phylogenetic analysis. First the total dataset was separated in two clades (Group I and Group II). Second, Group I was separated in Group IA, Group IB and Group IC, and Group II was separated in Group IID and Group IIE. Groups I and II correspond to well supported clades, as well as Groups IA and IE. Group IB correspond to a weakly supported clade. Group IC include all the specimens from Group I that are not included in Groups IA or IB. Similarly, Group ID includes all the specimens from Group II that are not included in Groups IIE.

Discussion of the primary hypotheses

In this second step of the analysis, the question is to determine if the clusters defined with the DBM and the GMYC corresponds to species. To test these primary hypotheses of species delimitation, several other characters and criteria are used. If the results are congruent, the primary hypotheses are validated; otherwise, the primary hypotheses will be interpreted as false positive (several genetic clusters that correspond to one species) or false negatives (one cluster that corresponds to several species) are detected.

Monophyly of the clusters will be tested by a phylogenetic reconstruction (methods are described in the "General Mixed Yule Coalescent model (GMYC)" section). A phylogenetic tree will also be constructed with the 28S gene, a nuclear marker not link with the COI gene: the two trees (COI and 28S) are expected to be identical between species but discordant within species (Andjic et al. 2007; Arzanlou et al. 2008). Morphological characters and geographical and bathymetrical distributions will also be analysed in order to detect differences between the genetic clusters. Genetic analyses were performed, using Arlequin 3.1 (AMOVA with a 10000 permutations test) to estimate structure between geographic populations within entities, and Network 4.5 (Median-Joining option) to construct haplotypes networks

Results

COI gene variability and primary hypotheses of species delimitation

One thousand specimens of Turrinae were sequenced for a fragment of 658 bp of the COI gene; 648 haplotypes are found, displaying 477 polymorphic sites, and a high haplotypic diversity (0.995). The Tamura-Nei model (TN+I+G, with I = 0.541 and G = 1.014) model was defined as the best fitting model. Genetic pairwise distances for COI gene were calculated with three different methods: the p-distances, the K2P distances, and the Tamura-Nei (TN) distances. Although there is no clear barcode gap, the three methods detect a threshold between low and high genetic distances, and the position of this limit (approximately between 0.02 and 0.04 %) is not modified depending on the mutation model used (Figure 1). However, as shown in figure 1B, the number of pairwise comparison for the genetic distances that correspond to this threshold is weaker for the TN distance than for the K2P distance and the p-distance.

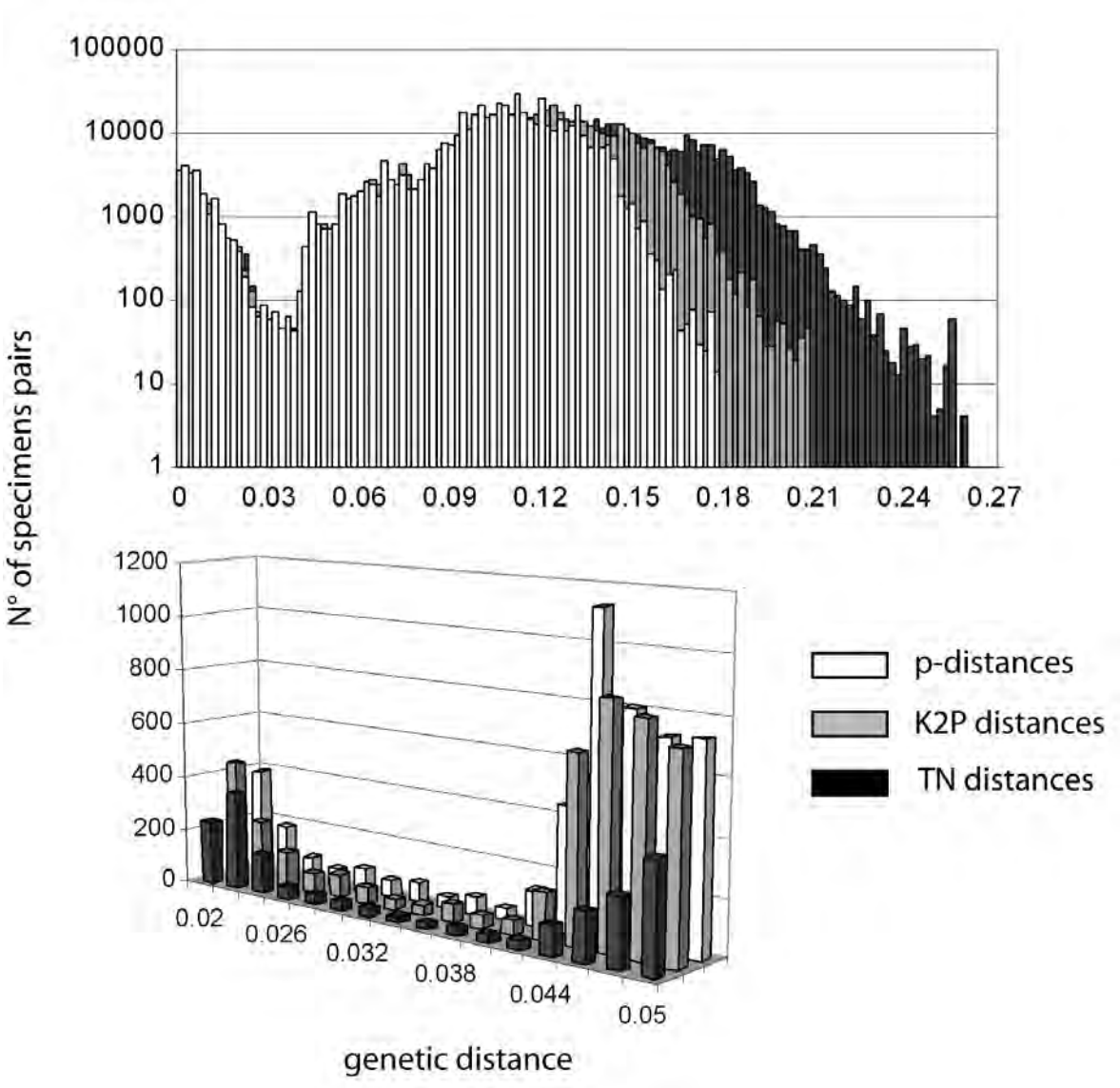


Fig. 1: A. Distributions of p-distance, K2P distances and TN distances between each pair of specimens for the COI gene. B. Same results, but focusing on the barcode gap zone (between 0.02 and 0.05).

Both methods of gap definition (KMGD and PGD) provide similar values of threshold between low and high genetic distances (Table 1, Fig. 2). However, this value is not always identical between both methods. For example, for the total dataset and for the dataset II, the threshold defined with the KMGD method is very different from the threshold defined with the PGD method, and does not correspond visually to the limit between the two modes. Thresholds values are also different when using different subsets (Figure 3): it varies from 0.0295 to 0.0513 (for the PGD method) between the different datasets analysed. For three datasets (IA, IB and IID), the two distributions (low and high genetic distances) are separated by a barcode gap. This discontinuity between the two distributions was hidden when all the sequences were analysed together (total dataset). These different thresholds obtained with different methods or datasets result in different primary partitions (Figure 4). At least 78 entities (isolated specimen or groups of specimens) and maximum 99 are recognized (see Table 2 for details). Sixty-one of them are identical whatever the method or the dataset used. In all other cases, results are different depending on the dataset or the method used: one group of specimens defined with one method (or dataset) is split in two, three or four entities with another method (or dataset).

Table 1: Threshold (with 99% confidence interval) between the two groups of genetic distances (low and high) obtained with the two methods of Gap Definition and for each dataset.

Dataset	KMGD			PGD		
	99%	Threshold	99%	99%	Threshold	99%
Total	0.053	0.0545	0.056	0.0325	0.0335	0.0335
I	0.0218	0.0443	0.0527	0.0362	0.0373	0.0383
IA	0.021	0.0366	0.0692	0.0349	0.0501	0.0501
IB	0.158	0.0297	0.0505	0.0293	0.0356	0.0387
IC	0.0212	0.0451	0.0675	0.0487	0.0513	0.0527
II	0.0846	0.0852	0.0867	0.0289	0.0295	0.0306
IID	0.0122	0.0516	0.0604	0.0356	0.0373	0.0518
IIE	0.0283	0.0355	0.043	0.0289	0.0295	0.0306

Table 2: Numbers of entities among the 1000 specimens defined by the MYC method and the Gap definition method (KMGD and PGD) for each dataset.

Dataset	MYC			KMGD	PGD
	99%	Threshold	99%	Threshold	Threshold
Total	84	84	87	?	87
I	43	45	48	45	46
IA	7	7	7	8	7
IB	7	8	10	8	8
IC	37	38	38	34	33
II	41	42	46	?	42
IID	23	24	26	23	23
IIE	20	20	20	14	19

Each of the 99 entities recognized includes a mean of 10.1 specimens: 30 entities are represented by only 1 specimen, and 56 are represented by less than 5 specimens; 74 are represented in only 1 geographic region (Taiwan, Philippines, Solomon Islands, Vanuatu, Chesterfield islands or Norfolk Ridge); only 10 entities are present in shallow (inferior to 100 m deep) and deep waters (superior to 100 m); 28 species names were attributed to a unique entity and 9 others species names were attributed to specimens dispatched in different entities

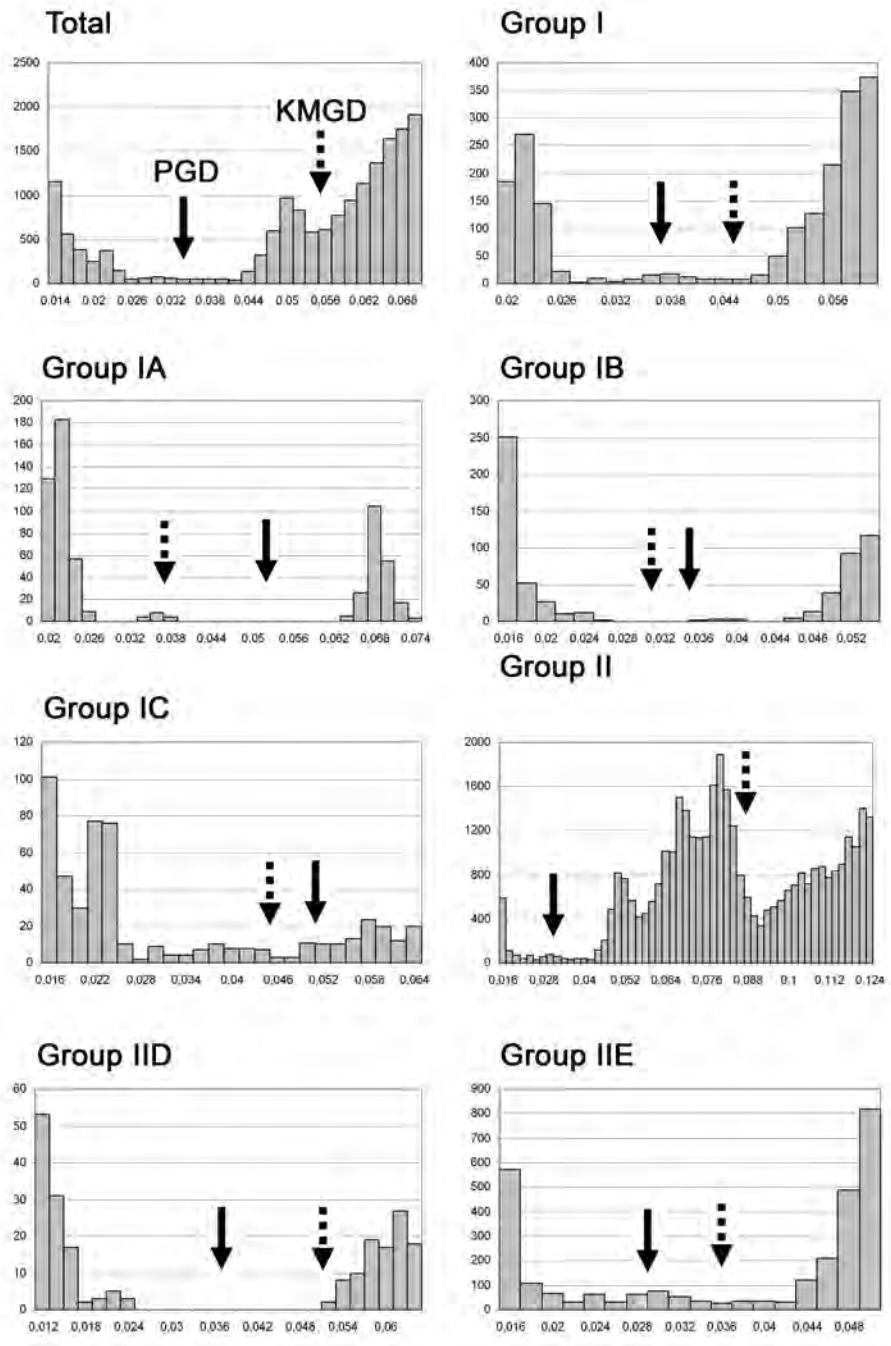


Fig. 2: Distributions of TN distances for the barcode gap zone for each dataset. Threshold between low and high genetic distances for PGD and KMGD methods are visualized.

(for example, the name *Lophiotoma jickelli* was attributed to two entities, named "*L. cf. jickelli* 1" and "*L. cf. jickelli* 2"); other entities are identified at the genus level, and are labelled within each genus from 1 to n.

Monophyly of all recognized entities is highly supported (Posterior Probability PP = 1), except in six cases (see caption of the Fig. 4 for details). In three cases where one entity defined with one method is split in two entities with another method, monophyly of one or both entities is not supported. For example, the clade that groups *Lophiotoma cf. indica 7* and *L. cf. indica 8* is highly supported (PP = 1), but each entity is not supported (PP *L. cf. indica 7* = 0.97; PP *L. cf. indica 8* = 0.76).

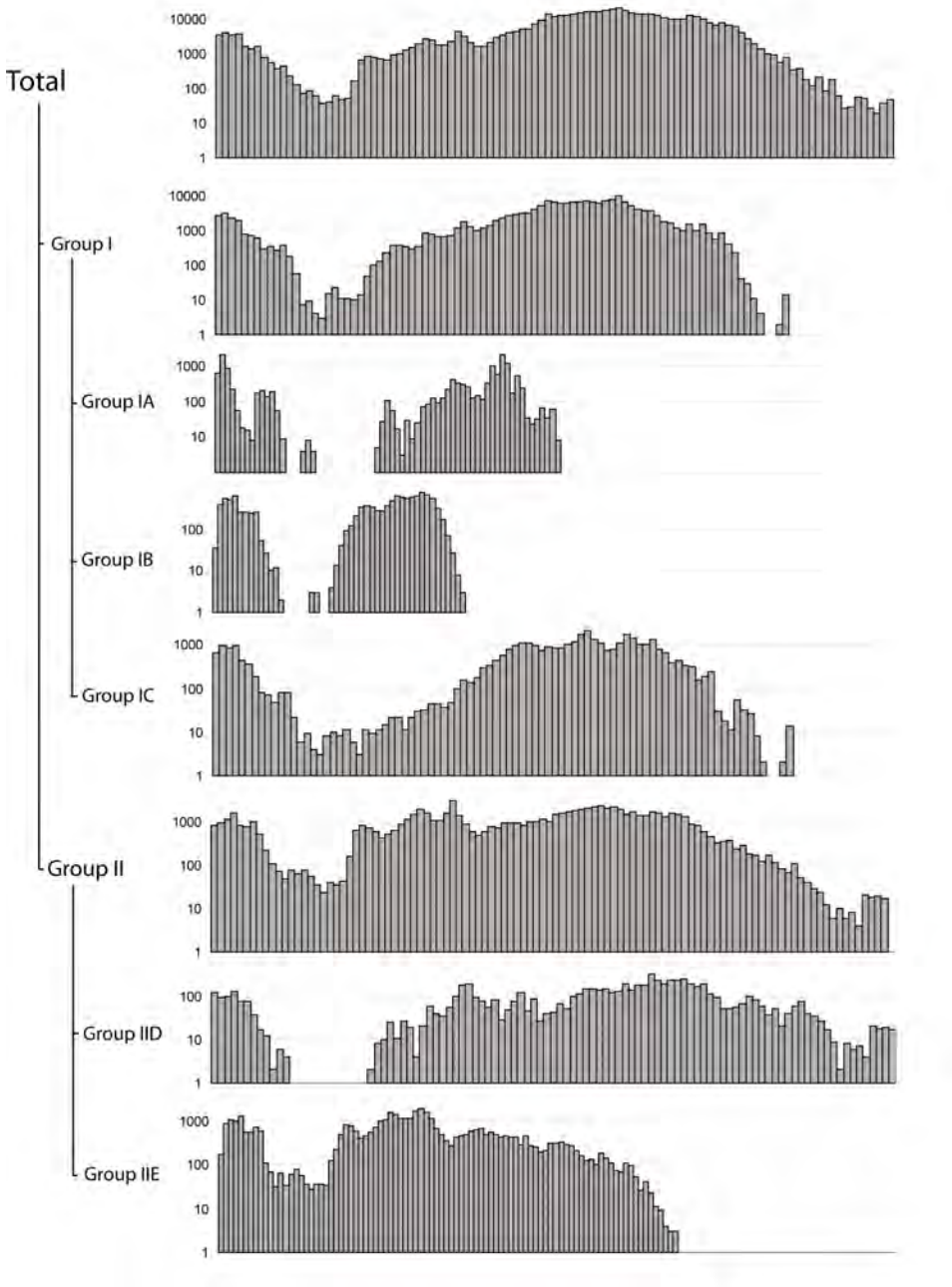
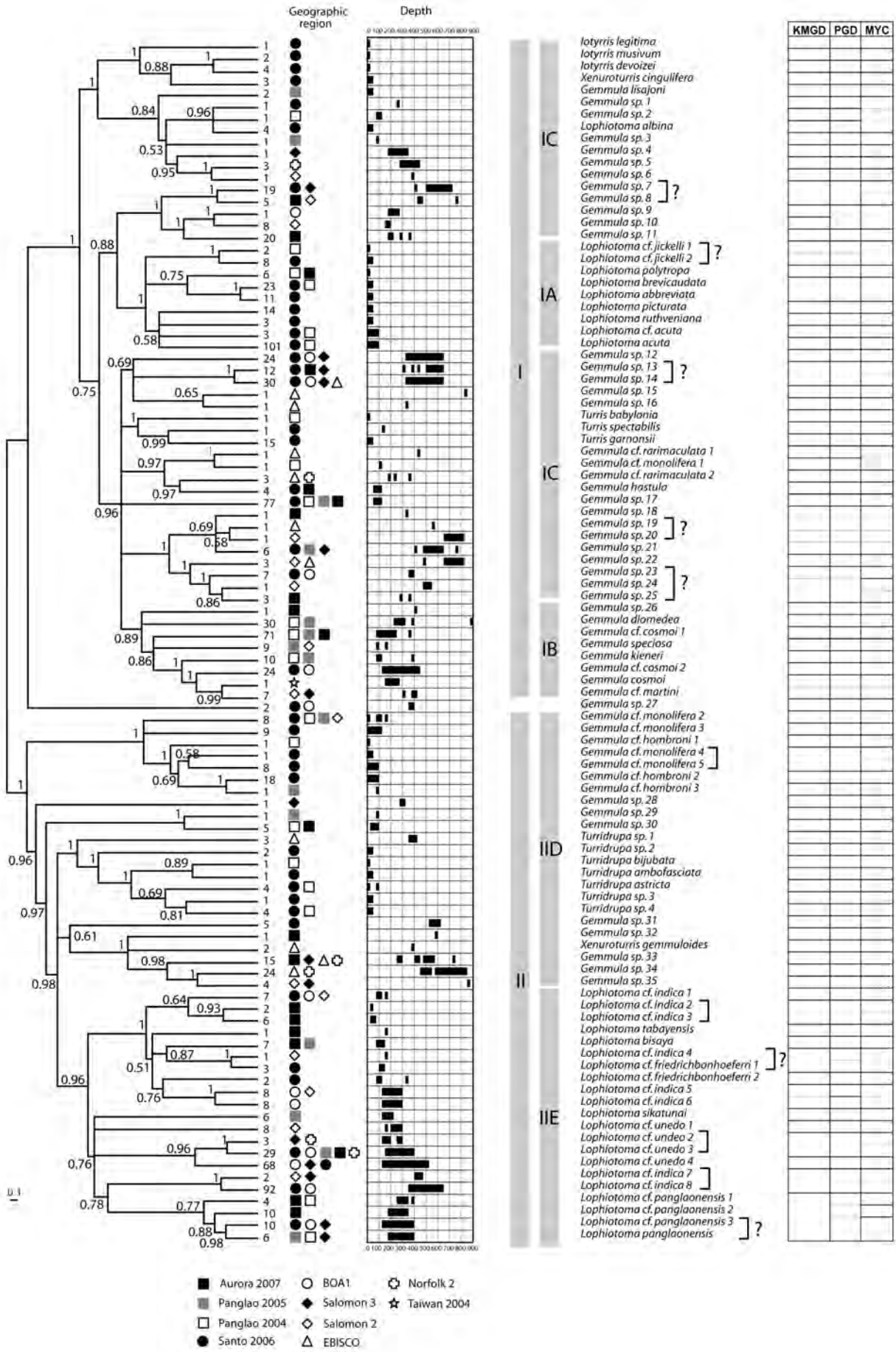


Fig. 3: Distributions of TN distances for the different COI datasets used.

Fig. 4 (next page): COI Bayesian tree obtained for the total dataset. Posterior probability is given for each node. Each branch represents an entity delimited with at least one exploratory method (except *L. abbreviata* and *L. brevicaudata* – details in the text). Each entity has a PP of 1, except *Gemmula cf. monolifera 2* (0.97), *G. cf. monolifera 5* (0.78), *Lophiotoma bisaya* (0.99), *L. unedo 3* (0.72), *L. cf. indica 7* (0.97) and *L. cf. indica 8* (0.76). Entities *G. sp. 7*, *G. sp. 8*, *G. sp. 13* and *G. sp. 14* were not found monophyletic with the total dataset (although represented as unique branches here) but highly supported (PP = 1) with other datasets. At the tip of each branch is given the number of specimens included in the corresponding entity, the geographic localization and the depth of collection (black bars represent the range of depth where specimens were collected). Groups I, II, A, B, C, D and E used for the analyses are represented as grey boxes. When possible, entities are linked to a species name; in several cases, they are linked only to a genus name associated to a number; several entities could be linked to one species name, and are thus numbered from 1 to n. Results of exploratory methods (MYC, PGD and KMGD) are given as cells. A lack of separation between two cells of two entities indicates that these two entities are recognized as one by the corresponding method. A dotted line between two cells indicates that two entities were recognized with some (but not all) datasets for the corresponding method. Brackets are used to indicate that the two entities certainly correspond to one species; doubtful statuses for seven pairs or triplets of entities are indicated by an interrogation point.


Geographic region
Depth
KMGD
PGD
MYC
I
IA
IB
IC
II
IID
IIE
Iotyrris legitima
Iotyrris musivum
Iotyrris devoizei
Xenuroturris cingulifera
Gemmula lisajoni
Gemmula sp. 1
Gemmula sp. 2
Lophiotoma albina
Gemmula sp. 3
Gemmula sp. 4
Gemmula sp. 5
Gemmula sp. 6
Gemmula sp. 7
Gemmula sp. 8
Gemmula sp. 9
Gemmula sp. 10
Gemmula sp. 11
Lophiotoma cf. jickelli 1
Lophiotoma cf. jickelli 2
Lophiotoma polytropa
Lophiotoma brevicaudata
Lophiotoma abbreviata
Lophiotoma picturata
Lophiotoma ruthveniana
Lophiotoma cf. acuta
Lophiotoma acuta
Gemmula sp. 12
Gemmula sp. 13
Gemmula sp. 14
Gemmula sp. 15
Gemmula sp. 16
Turris babylonia
Turris spectabilis
Turris garnonsii
Gemmula cf. rarimaculata 1
Gemmula cf. monolifera 1
Gemmula cf. rarimaculata 2
Gemmula hastula
Gemmula sp. 17
Gemmula sp. 18
Gemmula sp. 19
Gemmula sp. 20
Gemmula sp. 21
Gemmula sp. 22
Gemmula sp. 23
Gemmula sp. 24
Gemmula sp. 25
Gemmula sp. 26
Gemmula dlomedea
Gemmula cf. cosmoi 1
Gemmula speciosa
Gemmula kieneri
Gemmula cf. cosmoi 2
Gemmula cosmoi
Gemmula cf. martini
Gemmula sp. 27
Gemmula cf. monolifera 2
Gemmula cf. monolifera 3
Gemmula cf. hombroni 1
Gemmula cf. monolifera 4
Gemmula cf. monolifera 5
Gemmula cf. hombroni 2
Gemmula cf. hombroni 3
Gemmula sp. 28
Gemmula sp. 29
Gemmula sp. 30
Turridrupa sp. 1
Turridrupa sp. 2
Turridrupa bijubata
Turridrupa ambofasciata
Turridrupa astricta
Turridrupa sp. 3
Turridrupa sp. 4
Gemmula sp. 31
Gemmula sp. 32
Xenuroturris gemmuloides
Gemmula sp. 33
Gemmula sp. 34
Gemmula sp. 35
Lophiotoma cf. indica 1
Lophiotoma cf. indica 2
Lophiotoma cf. indica 3
Lophiotoma tabayensis
Lophiotoma bisaya
Lophiotoma cf. indica 4
Lophiotoma cf. friedrichbonhoeferi 1
Lophiotoma cf. friedrichbonhoeferi 2
Lophiotoma cf. indica 5
Lophiotoma cf. indica 6
Lophiotoma sikatunai
Lophiotoma cf. unedo 1
Lophiotoma cf. undeo 2
Lophiotoma cf. unedo 3
Lophiotoma cf. unedo 4
Lophiotoma cf. indica 7
Lophiotoma cf. indica 8
Lophiotoma cf. panglaonensis 1
Lophiotoma cf. panglaonensis 2
Lophiotoma cf. panglaonensis 3
Lophiotoma panglaonensis
Aurora 2007
Panglao 2005
Panglao 2004
Santo 2006
BOA1
Salomon 3
Salomon 2
EBISCO
Norfolk 2
Taiwan 2004

Discussions of the primary hypotheses

Each entity defined in the primary partition using the COI gene were analysed 28S gene. A fragment of 908 bp (after alignment) of the 28S gene was sequenced for 709 specimens, displaying 228 haplotypes with 359 polymorphic sites, and a haplotypic diversity of 0.979. The General Time Reversible model (GTR+I+G, with I = 0.817 and G = 0.651) model for the 28S gene was defined as the best fitting model. Morphology, geographic and bathymetric distribution were also used to test the primary hypotheses. Finally, 81 entities can be retained as secondary species hypotheses, as they are congruently recognized by most criteria and characters. In addition, the status of seven pairs or triplets of entities is doubtful (Fig. 4). We will not discuss all these cases, but detail only a few examples.

Comparison with 28S gene

Over the 99 entities defined by at least one method with COI gene, 47 include specimens that share a common 28S sequence. However, in 17 cases, and especially for the Group IIE, 28S gene displays almost no variability, and several pairs of groups recognized with the COI gene share common 28S haplotypes. In 10 cases, the reciprocal monophyly of the two groups recognized with COI gene is not supported with the 28S gene although they do not share common 28S haplotypes. For example, *L. abbreviata* is monophyletic and included within the paraphyletic group *L. brevicaudata*. Finally, 11 entities are represented by only one specimen. Finally, 14 were not successfully sequenced for 28S gene. Especially for the Group IIE, 28S is not enough variable and thus several entities defined with COI gene possess the same 28S sequence.

Complete congruency

In the case illustrated on figure 5, five entities are recognized with the COI gene, of which four include several specimens. Results obtained with the 28S gene are completely congruent, as specimens grouped together with the COI gene are also grouped with the 28S gene. Furthermore, the monophyly of the four entities that include several specimens is highly supported with COI gene (Posterior Probabilities PP = 1). Haplotypes network illustrates also the differences between the five entities. No apparent structure can be detected within each of them, even for the *Gemmula sp. 34* that includes specimens from very distant geographic regions (Philippines, Solomon Islands, Chesterfield Islands and Norfolk ridge). However, the sampling is too weak in each geographic region within the *G. sp. 34* to estimate potential genetic structure. Finally, each of the five entities possesses different shell morphology.

Incongruence between DBM and GMYC

In most cases of incongruence between both methods, only few specimens are implied. We will discuss here one of the three cases that include enough specimens to discuss intragroups variability. Forty-two specimens are placed in the same group with DBM, but in two groups with GMYC (only when using the Group IC). Theses two groups recognized with the GMYC method (*Gemmula sp. 13* and *G. sp. 14* – Fig. 4), are very well supported (PP = 1). The haplotypes network also shows up a structure between the two groups, and several specimens of *G. sp. 13* were collected in the same site as specimens from the group *G. sp. 14*. However, in this case, 28S gene does not allow the recognition of these two groups. Several specimens of each group possess the same 28S sequence. Furthermore, shell morphology appears similar between *G. sp. 13* and *G. sp. 14*, as illustrated in Figure 6. Furthermore, this shell morphology is also found in several other well-defined groups.

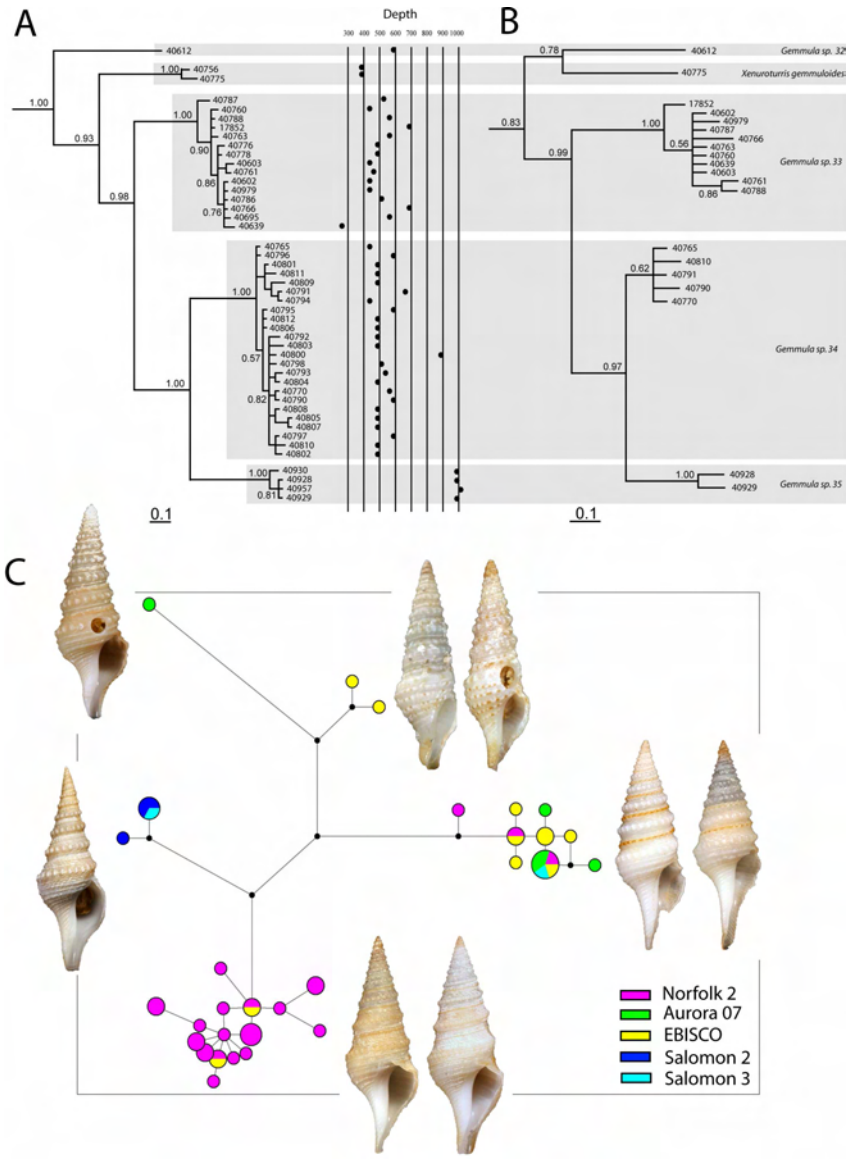


Fig. 5: A. COI tree for the entities *X. gemmuloides*, *Gemmula sp. 32*, *G. sp. 33*, *G. sp. 34* and *G. sp. 35*. Corresponding depth of collection for each specimen is given. B. 28S tree for the corresponding entities. C. COI haplotypes network for the same entities. Some specimens are illustrated for each entity by their shells.

Incongruence with shell morphology

Even when exploratory methods are congruent, the hypothesis of species delimitation provided can be rejected by other characters. For example, one entity recognized with all exploratory methods and with all datasets includes specimens corresponding to two different shell morphologies (Fig. 7). Furthermore, we associated theses morphologies to two available names in the literature: *Lophiotoma abbreviata* and *L. brevicaudata*. This two species hypotheses correspond to two clades with the COI gene (PP = 1 for both groups), and the genetic distance between the two groups does not exceed 0.025.

Genetic structure within a recognized entity

For all groups that include more than six specimens from different geographic regions, potential genetic structure was analysed. In all cases, the genetic structure detected was weak (data not shown). To illustrate these results, the haplotypes network for the species *Gemmula sp. 17* is shown in Figure 8. Genetic structure on COI gene between specimens from Philippines and specimens from Vanuatu is weak (F_{st} = 0.04; p-value = 0.043). However, a strong genetic structure is found with the 28S gene (F_{st} = 0.197; p-value = 0.0025).

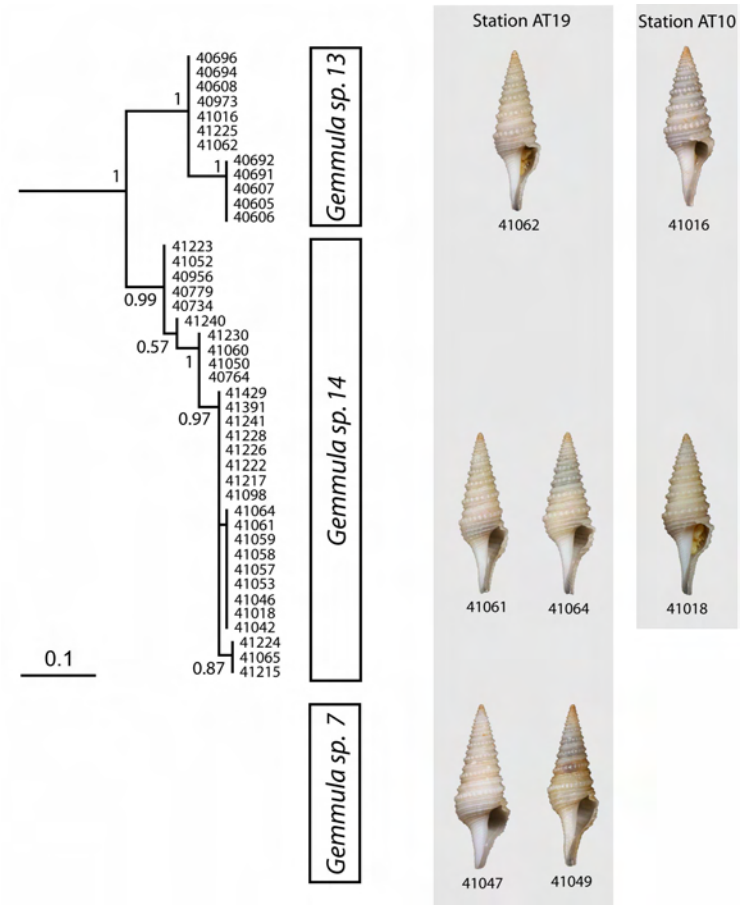


Fig. 6: COI tree for entities *Gemmula sp. 13* and *G. sp. 14*. Three specimens (one of the entity *G. sp. 13*, and two of *G. sp. 14*) collected at the sampling site AT19, and two (one of *G. sp. 13* and one of *G. sp. 14*) at the sampling site AT10 are illustrated. Two other specimens from another entity (*Gemmula sp. 7*) collected also at sampling site AT19 are shown.

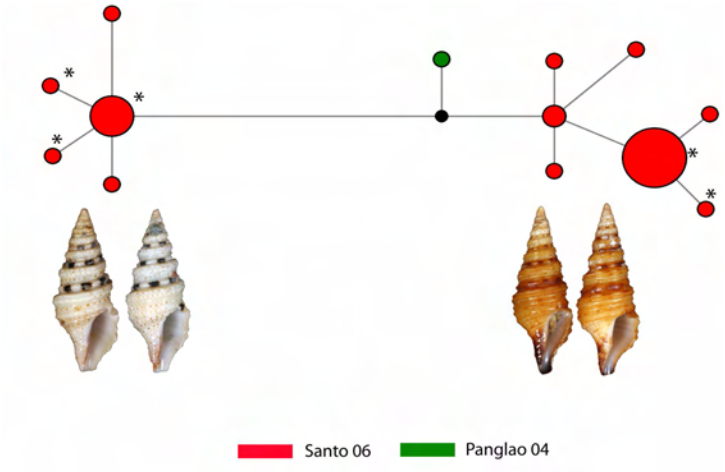


Fig. 7: COI haplotypes network for species *Lophiotoma abbreviata* and *L. brevicaudata*. *: haplotypes collected in the same sampling site (Santo 06, FS79). Species are illustrated by 2 shells each.

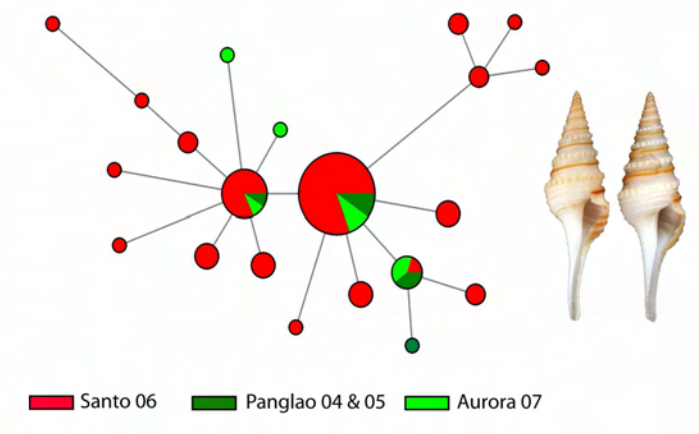


Fig. 8: COI haplotypes network for entity *Gemmula* sp. 17. One specimen from Vanuatu and one from Philippines are illustrated by their shells.

Discussion

Species delimitation topic is now developing quickly to answer the growing concern over biodiversity (Wiens 2007). In this context, we proposed here a formalized methodology of species delimitation, adapted to hyperdiverse groups. To manage such diversity, two exploratory methods were used, of which one (DBM) have been formalized in this study. As revealed by our results, these methods are effective in producing robust primary hypotheses of species delimitation. Most of these hypotheses were not rejected by the analyses of other characters and criteria.
Within DBM, both KMGD and PGD allow an estimation of the barcode gap, and constitute reproducible methods that will have to be tested with other datasets. However, even taking into account the relative cluster density bias, the KMGD method provide artefactual values of barcode gap when the variance of the "high distances" group is far higher than the variance of the "low distances" group (Fig. 3). In such cases, the PGD should be employed. It is also important to note that these two gap definition methods just attempt to identify the limit between the two distributions, and can provide in some cases hypotheses of species delimitation that are secondarily rejected.

Furthermore, we attempt to integrate the methodology in an evolutionary context, taking into account the sequences' evolution and the species genealogy that we are studying. Firstly, different model of substitution were compared. As already supposed by Nielsen & Matz (2006), using an elaborated model of substitution seems to make little difference at this taxonomic level. However, a deeper gap has been observed for genetic distances that correct homoplasy. We thus recommend using a well-adapted model of substitution as it should enhance the accuracy of the gap definition.
Secondly, the analysis of sub-datasets has revealed that the threshold between intra and interspecific distances and the limit between speciation and coalescent branching events can be different from one species to another (Table 2 and Figure 3). As evidenced by the different thresholds obtained with different datasets, a unique threshold defined for a whole dataset could conceal discrepancies in rate of evolution between species. Indeed, species are characterized by different times of speciation (Tautz 2003; Hickerson et al. 2006), leading to variable limits between intra and interspecific varability and between speciation and coalescent branching events among species. Furthermore, population size can vary greatly from one species to another, leading to different coalescence times for a given gene. This will lead again to differences between species in the limit between intra and interspecific variability and in the position of the branching events in the tree. Thus, the comparison of different subsets allows circumventing such potential bias by identifying groups of species that potentially share similar speciation and coalescence times.

If PAA methods were used, specimens collected in the same site and displaying a highly similar shell would have been placed placed in the same population. However, several species delimited in this study are morphologically highly similar, and some of them were sampled at the same site (*e.g. G. sp. 13*, *G. sp. 14* and *G. sp. 7* – Fig. 6). These results prevent us to group specimens as required in PAA methods, and exploratory methods based on genetic data should be preferred. Thus, this "pseudo-cryptic" diversity (Knowlton 2000; Saez 2005) was revealed only because all specimens were included in the analysis, and not only some representatives by morphospecies (Janzen 2005, Meyer & Paulay 2005). Furthermore, both exploratory methods used rely on a correct estimation of intra and interspecific variabilities

(Hebert et al. 2004; Pons et al. 2006; Arhens et al. 2007), and will consequently be particularly sensitive on sampling bias.

The second step is a necessary discussion of the entities defined by the exploratory methods. One entity can correspond to a species, but also to several species (false negative) or only to a population within a species (false positive). Several species recognized in our analysis were missed with one or the other exploratory methods; one pair of species, recognized by several secondary criteria (reciprocal monophyly, morphological differences) was missed by both exploratory methods (*L. abbreviata* and *L. brevicaudata*). Genetic distances between these two species are included in the part of the distribution that corresponds to intraspecific distances for other species.
Monophyly of the COI gene, variability and monophyly of the 28S gene, genetic structure, morphological variability and geographic distribution have been authoritative in several cases. Actually, these characters and criteria have permit to choose between the alternative hypotheses proposed by the exploratory methods. For example, the fact that two specimens of two different entities are present in the same sampling site, as for the species *G. sp*. 33 and 34, suggests that they cannot interbreed (Knowlton 2000). Similarly, the presence of two highly different morphotypes, moreover reciprocally monophyletic with the COI gene, within one entity is a strong indication that two different species can be recognized (*L. abbreviata* and *L. brevicaudata*).
Our results are also an illustration that the different properties used as criteria of species delimitation are not acquired simultaneously. For too recent species, one criteria can be verified, and not the other. This is for example the case of the 28S gene in our analysis, which mutation rate is certainly too slow to allow differences to appear between pairs of species that are reciprocally monophyletic for the COI gene. Identifying a nuclear gene that could be as variable and that fix differences as quickly as the COI gene is not trivial. ITS genes have also been proposed, but problems of alignment are frequent when the dataset includes not closely related species as it is the case in hyperdiverse groups such as the Turrinae.

If all the species delimited here and not linked to an existing name were associated with a new name, this could constitute a great increase in the number of species described: in the past 30 years, only 13 new names have been proposed in the genus *Gemmula*. Twelve new species delimited here, each of them represented by more than five specimens, can be attributed to the genus *Gemmula* and need new names. More generally, the genera *Lophiotoma* and *Gemmula* (besides certainly paraphyletic as already found by Heralde 2007) include 137 species, of which around 100 are considered valid (Tucker 2004). Comparatively, we recognized 67 entities within these genera only in the South-West Pacific.

Conclusion

DNA taxonomy and integrative approach are now currently held in alpha-taxonomy, but are sometimes difficult to apply in hyperdiverse groups. When species diversity is weaker, morphological characters easier to use and geographic distributions assessable, the taxonomists can quickly propose primary hypotheses of species delimitation. But in hyperdiverse groups such as Turrinae, where morphology is difficult to use and geographic distribution difficult to assess, a formalized, reproducible and testable methodology as the one proposed here is needed to perform an efficient primary partition of the diversity. However, similarly to the hypotheses provided by the morphologists, hypotheses obtained with the DBM or the GMYC methods are only a primary sorting that should be tested with other

characters and criteria. The combination of these two steps has the advantages of the DNA-taxonomy approach (a quick identification of entities that potentially correspond to species) and of the integrative taxonomy approach (a process of modification/validation of hypotheses of species delimitation), thus making DNA-taxonomy a real taxonomic science.

Acknowledgments

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the Tropical Deep Sea Benthos programme on board R/V Alis, that generated the deep-sea samples used in this study. The Total Foundation supported the Panglao 2004 and Santo 2006 biodiversity surveys and Ellen Strong and Yuri Kantor are thanked for their role in molecular sampling during these expeditions. We are also pleased to thank the staff of MNHN's "Service de Systématique Moléculaire" for technical facilities; the Consortium National de Recherche en Génomique, Génoscope for the sequencing facilities, and Alexander Sysoev, Baldomero Olivera and Yuri Kantor for the taxonomic identifications.

References

Agapow, P. M., O. R. P. Bininda-Edmonds, K. A. Crandall, J. L. Gittleman, G. M. Mace, J. C. Marshall, and A. Purvis. 2004. The impact of species concept on biodiversity studies. The quarterly review of biology 79:161-179.

Ahrens, D., M. T. Monaghan, and A. P. Vogler. 2007. DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae). Molecular Phylogenetics and Evolution 44:436–449.

Andjic, V., G. E. S. Hardy, M. N. Cortinas, M. J. Wingfield, and T. I. Burgess. 2007. Multiple gene genealogies reveal important relationships between species of *Phaeophleospora* infecting *Eucalyptus* leaves. FEMS Microbiology Letters 268:22-33.

Arzanlou, M., J. Z. Groenewald, R. A. Fullerton, E. C. A. Abeln, J. Carlier, M. F. Zapater, I. W. Buddenhagen, A. Viljoen, and P. W. Crous. 2008. Multiple gene genealogies and phenotypic characters differentiate several novel species of *Mycosphaerella* and related anamorphs on banana. Persoonia 20:19-37.

Barrett, R. D. H., and P. D. N. Hebert. 2005. Identifying spiders through DNA barcodes. Canadian Journal of Zoology 83:481-491.

Bouchet, P., and A. Waren. 1980. Revision of the north-east Atlantic bathyal and abyssal Turridae (Mollusca, Gastropoda). Journal of Molluscan Studies SuppI. 8:1-19.

Brower, A. V. Z. 2006. Problems with DNA barcodes for species delimitation: 'ten species' of *Astraptes fulgerator* reassessed (Lepidoptera: Hesperiidae). Systematics and Biodiversity 4:127-132.

De Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. Pages 57-75 *in* Endless forms. Species and speciation (D. J. Howard, and S. H. Berlocher, eds.). Oxford University Press, Oxford.

De Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879-886.

DeSalle, R. 2006. Species discovery versus Sspecies identification in DNA Barcoding efforts: response to Rubinoff. Conservation Biology 20:1545-1547.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299.

Fritz, U., D. Guicking, M. Auer, R. S. Sommer, M. Wink, and A. K. H. Hundsdörfer. 2008. Diversity of the Southeast Asian leaf turtle genus *Cyclemys*: how many leaves on its tree of life? Zoologica Scripta 37:367-390.

Fu, J., and X. Zeng. 2008. How many species are in the genus *Batrachuperus*? A phylogeographical analysis of the stream salamanders (family Hynobiidae) from southwestern China. Molecular Ecology 17:1469-1488.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology Evolution and Systematics 34:397-423.

Giraud, T., G. Refrégier, M. Le Gac, D. M. de Vienne, and M. E. Hood. 2008. Speciation in fungi. Fungal Genetics and Biology 45:791-802.

Gomez, A., P. J. Wright, D. H. Lunt, J. M. Cancino, G. R. Carvamho, and R. N. Hughes. 2007. Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon. Proceedings of the Royal Society B 274:199-207.

Hajibabaei, M., D. H. Janzen, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences 103:968-971.

Hajibabaei, M., G. A. C. Singer, P. D. N. Hebert, and D. A. Hickey. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics 23:167-172.

Hassouna, M., B. Michot, and J.-P. Bachellerie. 1984. The complete nucleotide sequence of mouse 28SrRNA gene. Implications for the process of size increase of the large subunit rRNA of higher eukaryotes. Nucleic Acid Research 12:3563-3583.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identifications through DNA Barcodes. Proceedings of the Royal Society B 270:313-321.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences 101:14812-14817.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis. 2004. Identification of birds through DNA barcodes. PLoS Biology 2:1657-1663.

Heralde, F. M., M. Watkins, J.-P. Ownby, P. K. Bandyopadhyay, A. D. Santos, G. P. Concepcion, and B. M. Olivera. 2007. Molecular phylogeny of some Indo-Pacific genera in the subfamily Turrinae H. Adams and A. Adams, 1853 (1838) (Gastropoda: Neogastropoda). Nautilus 121:131-138.

Hickerson, J. H., C. P. Meyer, and C. Moritz. 2006. DNA Barcoding will often fail to discover new animal species over broad parameter space. Systematic Biology 55:729-739.

Huelsenbeck, J. P., F. Ronquist, and B. Hall. 2001. MrBayes: bayesian inference of phylogeny. Bioinformatics 17:754-755.

Janzen, D. H., M. Hajibabaei, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2005. Wedding biodiversity inventory of a large and complex lepidopteria fauna with DNA Barcoding. Philosophical Transactions of the Royal Society B 360:1835-45.

Jovelin, R., and J.-L. Justine. 2001. Phylogenetic relationships within the Polyopisthocotylean monogeneans (Plathyhelminthes) inferred from partial 28S rDNA sequences. International Journal for Parasitology 31:393-401.

Knowles, L. L., and B. C. Carstens. 2007. Delimiting species without monophyletic gene trees. Systematic Biology 56:887-895.

Knowlton, N. 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420:73-90.

Maddison, W. P., and D. R. Maddison. 2007. Mesquite: a modular system for evolutionary analysis. Version 2.01.

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.

Mariaux, J., N. Lutzmann, and J. Stipala. 2008. The two-horned chamaeleons of East Africa. Zoological Journal of the Linnean Society 152:367-391.

Marshall, J. C. 2006. Delimiting species: comparing methods for mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. Evolution 60:1050-1065.

Meyer, P. C., and G. Paulay. 2005. DNA Bracoding: error rates based on comprehensive sampling. PLoS Biology 3:1-10.

Miller, S. E. 2007. DNA barcoding and the renaissance of taxonomy. Proceedings of the National Academy of Sciences 104:4775-4776.

Monaghan, M. T., M. Balke, J. Pons, and A. P. Vogler. 2006. Beyond barcodes: complex DNA taxonomy of a South Pacific Island radiation. Proceedings of the Royal Society B 273:887-893.

Moritz, C., and C. Cicero. 2004. DNA Barcodes : promise and pitfalls. PLoS Biology 2:1529-1531.

Nee, S., A. O. Mooers, and P. H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. The Proceedings of the National Academy of Sciences USA 89:8322-8326.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16:358-364.

Nielsen, R., and M. Matz. 2006. Statistical Approaches for DNA Barcoding. Systematic Biology 55:162-169.

Pfenninger, M., M. Cordellier, and B. Streit. 2006. Comparing the efficacy of morphologic and DNA-based taxonomy in the freshwater gastropod genus *Radix* (Basommatophora, Pulmonata). BMC Evolutionary Biology 6:100.

Pons, J., T. G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S. Kamoun, W. D. Sumlin, and A. P. Vogler. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55:595-609.

Posada, D., and K. A. Crandall. 2001. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). Molecular Biology and Evolution 18:897-906.

Powell, A. W. B. 1964. The family Turridae in the indo-pacific. Part I. The subfamily Turrinae. Indo-Pacific Mollusca 1:227-411.

Rezac, M., S. Pekar, and J. Johannesen. 2008. Taxonomic review and phylogenetic analysis of central European *Eresus* species (Araneae: Eresidae). Zoologica Scripta 37:263-287.

Rissler, L. J., and J. J. Apodaca. 2007. Ecological niche models and phylogeography help uncover cryptic biodiversity. Systematic Biology 56:924-942.

Rosenberg, N. A., and R. Tao. 2008. Discordance of species trees with their most likely gene trees: the case of five taxa. Systematic Biology 57:131-140.

Rubinoff, D. 2006. DNA Barcoding evolves into the familiar. Conservation Biology 20:1548-1549.

Sáez, A. G., and E. Lozano. 2005. Body doubles - Cryptic species: as we discover more examples of species that are morphologically indistinguishable, we need to ask why and how they exist. Nature 433:111.

Salter Kubatko, L., and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56:17-24.

Samadi, S., and A. Barberousse. 2006. The tree, the network, and the species. Biological Journal of the Linnean Society 89:509-521.

Savage, J. M., and J. I. Watling. 2008. Not so rare snakes: a revision of the *Geophis sieboldi* group (Colubridae: Dipsadinae) in lower Central America and Colombia. Zoological Journal of the Linnean Society 153:561-599.

Savolainen, V., R. S. Cowan, A. P. Vogler, G. K. Roderick, and R. Lane. 2005. Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society B 360:1805-1811.

Schander, C., and E. Willassen. 2005. What can biological barcoding do for marine biology? Marine Biology Research 1:79-83.

Sites, J. W., and J. C. Marshall. 2003. Delimiting species: a renaissance issue in systematic biology. Trends in Ecology and Evolution 19:462-470.

Sites, J. W., and J. C. Marshall. 2004. Operational criteria for delimiting species. Annual Review of Ecology Evolution and Systematics 35:199-227.

Smith, A. M., B. L. Fisher, and P. D. N. Hebert. 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society B 360:1825-1834.

Steinke, D., M. Vences, W. Salzburger, and A. Meyer. 2005. TaxI: a software tool for DNA barcoding using distance methods. Proceedings of the Royal Society B 360:1975-1980.

Steinley, D. 2003. Local optima in K-Means clustering: what you don't know may hurt you. Psychological Methods 8:294-304.

Steinley, D., and M. J. Brusco. 2007. Initializing K-means batch clustering: a critical evaluation of several techniques. Journal of Classification 24:99-121.

Swofford, D. L. 2002. PAUP*. *Phylogenetic Analysis Using Parsimony* and other methods*. Sinauer Associates.

Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Vogler. 2003. A plea for DNA taxonomy. Trends in Ecology and Evolution 18:70-74.

Tucker, J. K. 2004. Catalogue of recent and fossil turrids (Mollusca: Gastropoda). Zootaxa 682:1-1295.

Vogler, A. P., and M. T. Monaghan. 2007. Recent advances in DNA taxonomy. Journal of Zoological Systematics and Evolutionary Research 45:1-10.

Ward, R. D., T. S. Zemlak, B. H. Innes, P. R. Last, and P. D. N. Hebert. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B 360:1847-1857.

Wiens, J. J. 2007. Species delimitation: new approaches for discovering diversity. Systematic Biology 56:875-878.

Will, K. P., B. D. Mishler, and Q. D. Wheeler. 2005. The perils of DNA Barcoding and the need for integrative taxonomy. Systematic Biology 54:844-851.

C. 2nde étape : discussion des entités identifiées

Dans cette partie, je présente les résultats obtenus pour quatre groupes de Turridae. Je reviens tout d'abord sur la sous-famille des Turrinae, en détaillant les tests des hypothèses primaires de délimitation d'espèces qui ont été effectués. Je présente ensuite l'analyse des genres *Bathytoma*, *Benthomangelia* et *Xenuroturris*. Pour chacun d'entre eux, une analyse exploratoire basée sur le gène COI a été réalisée, et les hypothèses primaires de délimitation d'espèces ont été testées avec d'autres caractères et critères. J'insisterai dans chacun des trois cas sur l'originalité des caractères qui ont été utilisés.

1. La sous-famille des Turrinae

Plusieurs hypothèses de délimitation d'espèces proposées pour la sous-famille des Turrinae sont en accord avec les autres caractères et critères utilisés. Par exemple, un groupe de quatre espèces définies avec les méthodes DBM et GMYC sont monophylétiques avec le gène COI. De plus, elles correspondent également à cinq clades avec le gène 28S. Enfin, elles possèdent toutes une morphologie distincte.

Dans d'autres cas, même si la majorité des caractères et des critères ne rejettent pas les hypothèses primaires proposées, un caractère ou un critère ne permet pas de détecter des différences entre deux hypothèses d'espèces. Par exemple, la monophylie de certaines hypothèses d'espèces n'est pas soutenue avec le gène COI, ou elle n'est pas retrouvée avec le gène 28S. Dans plusieurs cas, les groupes définis avec le gène COI possèdent une morphologie très semblable.

Dans ces deux cas (tous les caractères ou critères sont congruents, ou un seul ne l'est pas), les hypothèses primaires n'ont pas été rejetées. Au total, 81 espèces ont donc été délimitées. Le statut de sept autres hypothèses primaires est incertain, d'autres caractères ou critères suggérant qu'elles incluent chacune deux ou trois espèces différentes. Par exemple, les deux hypothèses primaires *Lophiotoma cf. jickelli* 1 et *L. cf. jickelli* 2 sont séparées avec la méthode GMYC, et correspondent à deux groupes monophylétiques avec le gène COI, mais elles ne sont pas séparées avec la méthode DBM, ne sont pas distinguées par le gène 28S, et n'ont pas été discernées morphologiquement.

2. Approche morphologique traditionnelle : le genre Bathytoma

Les spécimens de *Bathytoma* ont été dans un premier temps séquencés pour le gène COI, et plusieurs hypothèses primaires de délimitation d'espèces ont été proposées. Ils ont ensuite été confiés à Alexander Sysoev, taxonomiste spécialiste des Turridae. Il les a regroupé en morphoespèces de façon traditionnelle, c'est-à-dire en se basant sur les caractères de la coquille et sur la répartition géographique des échantillons. Ces morphoespèces ont ensuite été comparées avec les résultats obtenus avec le gène COI (Article 6). Dans la plupart des cas, les morphoespèces proposées n'étaient pas congruentes avec les hypothèses primaires proposées avec le gène COI.

Un gène nucléaire variable (ITS2) a été également séquencé. Il a permis de tester l'absence de flux de gènes entre les hypothèses primaires proposées avec le gène COI. Les résultats indiquent que les généalogies de gènes sont identiques entre les hypothèses d'espèces (sauf dans deux cas, détaillés dans l'article) mais différentes au sein de chacune de ces hypothèses. Cela signifie que ces deux gènes recombinent au sein des hypothèses d'espèces mais pas entre ces mêmes hypothèses d'espèces. Les hypothèses primaires établies sur le COI sont donc confirmées par le gène ITS2, mais à l'inverse, celles basées sur la morphologie sont rejetées. Les caractères morphologiques qui entraient en contradiction avec

les hypothèses de délimitation d'espèces pourraient correspondre à des caractères déterminés par l'environnement (plasticité phénotypique) ou à des caractères soumis à une sélection liée à des conditions environnementales particulières (adaptation locale).

Le réexamen de la coquille, en tenant compte cette fois-ci des résultats moléculaires, a permis d'identifier des différences morphologiques entre les espèces délimitées par les gènes COI et ITS2. Ces différences portent sur des caractères qui avaient été déjà reconnus lors de la première analyse, mais considérés comme insignifiant au niveau spécifique, ou qui n'avaient pas été identifiés. Cette deuxième analyse montre qu'il est plus facile d'identifier des différences morphologiques lorsque des hypothèses de délimitation d'espèces sont disponibles, plutôt que de proposer des hypothèses d'espèces en n'utilisant que l'information morphologique.

Finalement, 14 espèces ont été définies. En se référant aux descriptions d'espèces disponibles dans la littérature pour le genre *Bathytoma*, il est apparu que seules deux d'entre elles pouvaient être associées à un nom d'espèce disponible. Les 12 autres correspondent donc à des espèces nouvelles. L'une d'entre elle n'était représentée que par un seul spécimen dans les collections du muséum, et n'a pour l'instant pas été décrite. Les descriptions des onze autres sont inclues dans l'article 6. Les résultats obtenus ont également permis de proposer des hypothèses quant au processus de spéciation à l'origine des patterns mis en évidence.

Loss of planktotrophy, fragmentation and speciation: the deep-water gastropod *Bathytoma* (Gastropoda: Conoidea) in the West Pacific

Nicolas Puillandre[1], Alexander Sysoev[2], Baldomero Olivera[3], A. Couloux[4] and Philippe Bouchet[1].

[1] Muséum National d'Histoire Naturelle, 55, Rue Buffon, 75005 Paris, FRANCE, puillandre@mnhn.fr
[2] Zoological Museum of Moscow State University, Bolshaya Nikitskaya str. 6, Moscow 103009, Russia
[3] University of Utah, Biology Department, Salt Lake City, Utah 84112, USA
[4] GENOSCOPE, Centre National de Séquençage, Evry, France

En préparation ; pour soumission à Zoologica Scripta

Abstract

Dispersal capabilities are crucial in how speciation patterns are determined in marine invertebrates. Species possessing a long-living planktonic larva apparently have a dispersal advantage over those with direct development, and their distant populations may exchange genetic material maintaining a broad geographical range for the species. Recent species of the genus *Bathytoma* (Conoidea) are all characterized by non-planktonic development, having most probably lost a free-swimming larva in the pre-Pliocene, as Miocene fossils have protoconconchs indicating planktotrophic larval development. All have a bathyal distribution (100-1500 m), which implies that their capability for direct expansion on the bottom is restricted by both deep-sea basins and shallow-water areas, especially in insular Central Indo-Pacific. Therefore, it can be hypothesized that *Bathytoma* populations should represent numerous, mostly allopatric taxa restricted to a single or contiguous island groups. We tested this hypothesis using molecular and morphological characters independently. At first, 138 specimens collected in the Philippines, Solomons, Vanuatu, and the Coral Sea were sequenced for one mitochondrial (COI) and one nuclear (ITS2) gene. Then these specimens were sorted into morphospecies, without taking into account the results obtained with molecular data. Finally, 14 groups were defined with DNA sequences, most of them corresponding to one or two morphospecies. A recurring analysis of shell characters allowed to clearly delimit the species on the basis of both conchological and molecular approaches. As a result, 13 species (11 of them unnamed) and 1 form of uncertain status are recognized: 7 occur in the Solomons and 6 in the Philippines. Large-scale (inter-archipelagic) sympatry is a rare event (3 species) and could be a remnant of the ancient (pre-Pliocene) radiation. On the whole, the phylogeographic pattern of the genus diversity is rather complex and reflects probably also processes of sympatric and fine-scale allopatric speciation, and local extinctions. Eleven new species are described and illustrated.

Introduction

Speciation patterns in marine environments have been found to be extremely diverse, ranging from multi-scaled allopatric speciation to sympatric speciation (Knowlton 2000; Briggs 2006). Several factors have been proposed to explain these patterns, focusing mainly on dispersal capacities (Collin 2003; Meyer *et al* 2005; Vallejo 2005) and ecological ranges (Bierne *et al* 2003). In this context, molluscs are of particular interest as their dispersal capacities can be easily inferred from protoconch morphology. Thus, a multispiral protoconch indicates the presence of a planktotrophic larva capable to stay in the water column for up to several weeks and to drift with currents for long distances. On the contrary, a paucispiral protoconch formed during non-planktotrophic or direct development indicates greatly reduced dispersal capacities (see *e.g*. Jablonski & Lutz 1980; Paulay & Meyer 2006).

Bathytoma is a genus of marine gastropods of the superfamily Conoidea, with currently about 25 nominal species known from deep waters of the tropical West Pacific and Indian Oceans, South Africa, and the Western Atlantic. Some species penetrate the more temperate waters of New Zealand and northern Japan. No species are yet known from the Eastern Atlantic and the Eastern Pacific. This genus has a well documented paleontological record. As a fossil, *Bathytoma* is known since the Eocene of Europe and North America and since the Miocene of New Zealand (Powell 1966). No species are known from the Eocene of what is today the Indo-Pacific region. It is interesting to note that in geological history the evolution of the genus was accompanied by a gradual loss of planktotrophic and, generally, planktonic larval development. European Eocene species had protoconchs of 3.5-4.5 whorls (data from Amitrov 1973), American Oligocene species – of 4.0-4.5 whorls (data from MacNeil & Dockery 1984), New Zealand Miocene species – of 3.0-3.5 whorls (data from Beu & Maxwell 1990), whereas in the Pliocene the number of the protoconch whorls in the Indo-Pacific decreased to 1.5, as evidenced by species from Okinawa (Noda 1980, 1988) and New Zealand (Beu & Maxwell 1990), though the European Pliocene *B. cataphracta*, the type species of the genus, still possessed 3.0 whorls (Bernasconi & Robba 1984). All known Recent species have protoconch of 1.5 whorls, *i.e*. lack a planktonic larva, and therefore possess a very limited dispersal capacity.

All known *Bathytoma* species are bathyal in their bathymetric distribution, occurring at 100 to about 1500 m, usually at 200-700 m. There is good evidence that it was so in the past as well (at least in the Mio-Pliocene, see Beu & Maxwell 1990 and Noda 1988 for the respective fossil formations). This kind of bathymetric preference suggests that the capability of molluscs for direct expansion is restricted by both deep-sea basins and shallow-water areas. The Central Indo-Pacific, where *Bathytoma* is known since at least the Miocene (India, New Zealand – Powell 1966), is characterized by a highly fragmented bathyal, both at macro- (groups of islands) and microgeographic (isolated seamounts and groups of them) scales, as opposite to the classical continuous "continental slope" (Hall 2002).

Given these circumstances, the genus *Bathytoma* possessing weak dispersal capacities at a highly fragmented habitat, could be expected to demonstrate evidence of allopatric speciation (Kirkendale & Meyer 2004; Meyer *et al* 2005). However, the lack of data regarding gene flow and thus the lack of robust species delimitation in this genus prevent a reliable confirmation or rejection of this hypothesis. We propose here to test it by (i) proposing species delimitation based on molecular and morphological analyses and (ii) inferring speciation patterns from their actual geographic distribution.

Available species descriptions in the genus *Bathytoma* are based on conchological and, sometimes, radular characters. Indeed, shell characters are commonly used to define species in marine molluscs, as they allow a proxy of the variability and can be observed in both dead and live-taken specimens. However, they are known to be highly plastic (Knowlton 2000),

and comparative studies have recently highlighted the influence of environment on intraspecific variability in some groups of molluscs (Samadi *et al* 2000; Yeap *et al* 2001; Bichain *et al* 2007; Brookes & Rochette 2007). Our lack of understanding the significance of variation of shell characters in *Bathytoma* has resulted in taxonomists reaching conclusions that are opinions rather than testable hypotheses. For instance, Sysoev & Bouchet (2001) synonymized many nominal species under *B. atractoides*, whereas Vera-Pelaez & Lozano-Francisco (2004) described three new species based solely on shell characters.
Molecular characters are now routinely used to hypothesize species delimitation and to evaluate the robustness of shell characters in alpha-taxonomy. Most studies (e.g., *Conus*, Duda & Kohn 2005; Duda & Rolan 2005; "*Littorina*", Reid 2006; Cypraeidae, Meyer *et al* 2005; *Xenuroturris*, Kantor *et al*, unpublished results) use molecular characters to test *a posteriori* the validity of species already recognized and described based on shell characters. We propose here to use a double-blind approach: (i) all material available was sequenced for two genes (a mitochondrial and a nuclear one), without any *a priori* morphological hypothesis; (ii) the same specimens were studied for shell characters in order to delimit species as is traditional in molluscan taxonomy, without knowledge of the molecular results. Our species delimitation is based on two criteria derived from the lineage species concept (De Queiroz 1998; Samadi & Barberousse 2006): genetic or morphological differences found between individuals, to delimit clusters of genetically or morphologically similar specimens (Hebert *et al* 2003; Sites & Marshall 2004), and monophyly of the genetic clusters (Mallet 2001; Sites & Marshall 2004). Following this approach, we are able to compare and discuss the two types of characters to propose robust species delimitation. Finally, based on the observed geographic distribution of the species involved, we discuss patterns of speciation that can be inferred from the analysis.

Material & Methods

Sampling

A total of 138 live-taken specimens of *Bathytoma*, collected between 2004 and 2007 in the Philippines, Solomon Islands, Vanuatu, and Coral Sea, were available (Appendix 1; Fig. 1) for both molecular and morphological analyses. Material for molecular studies had been preserved in 90 or 100% ethanol by clipping pieces of the head-foot from anesthetized specimens. Additional dead-collected shells from the same or different cruises were also available (see Appendix 2).

Sequencing

DNA was extracted from a piece of foot, using 6100 Nucleic Acid Prepstation system (Applied Biosystem) or DNeasy® 96 Tissue kit (Qiagen) for smaller specimens. Two gene fragments were analysed: (i) a fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene using universal primers LCO1490 and HCO2198 (Folmer *et al* 1994), (ii) a fragment of 500 bp of the ITS2 gene, using the primers ITS3 (Chombard *et al* 1998) and ITS4 (White *et al* 1990). All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq, QBiogene for all genes. Thermocycles used for COI gene are those described in Hebert *et al*. (2003); for ITS2 gene, they consisted in an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 50°C for ITS2 gene. The final extension was at 72°C for 10'. PCR products were

purified and sequenced by the Genoscope (genbank accession numbers EU592046-EU592124 and EU292126-EU592231). In all cases, both directions were sequenced to confirm accuracy of each haplotype sequence.
Two closely related specimens (Puillandre *et al* 2008) were used as outgroups: one specimen of the genus *Mitromorpha*, Conidae (17893, GenBank accession number: EU015716 for COI gene and EU592233 for ITS2 gene) and one specimen of the genus *Etrema*, Conidae (17869, GenBank accession number: EU015691 for COI gene and EU592232 for ITS2 gene).

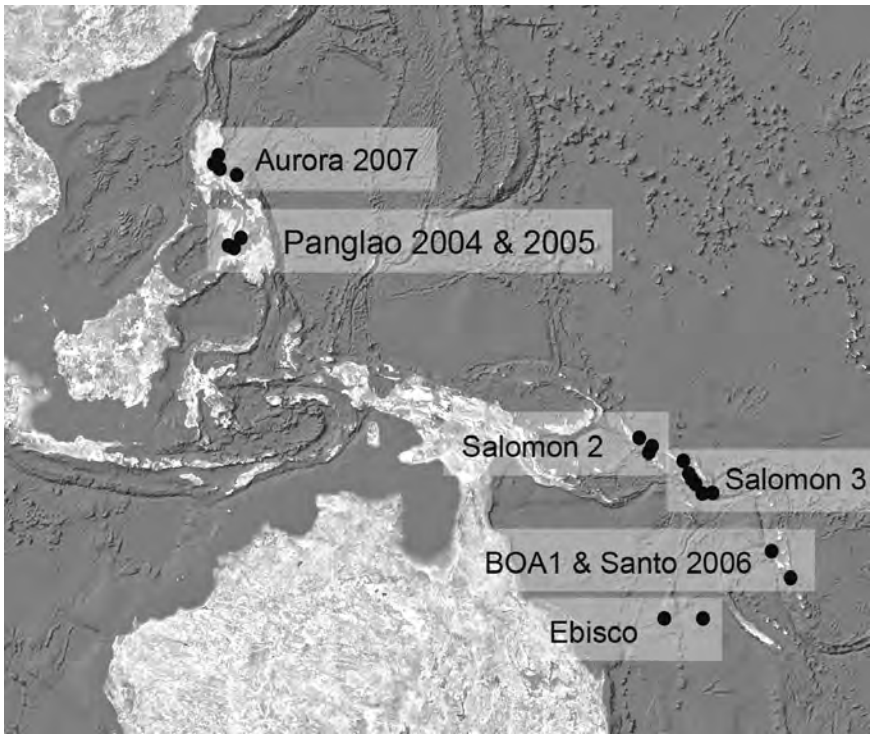


Fig. 1. Map of South-West Pacific showing sampling sites (black circles), grouped in cruise of collection (Panglao 2004 and Panglao 2005 took place in the same region; same thing for BOA1 and Santo 2006)

Genetic analysis

COI sequences were manually aligned, as no ambiguous indel were found. All sites were conserved for phylogenetic analysis. ITS2 sequences were aligned using ClustalW multiple alignment implemented in BioEdit version 7.0.5.3 (Hall 1999). Standard molecular diversity parameters were calculated using Arlequin 3.1 (Excoffier *et al* 2005).
Genetic pairwise distances for COI and ITS2 genes (excluding outgroups) were calculated with PAUP 4.0b10 (Swofford 2002), using the best fitting model of nucleotide substitution for each gene as defined by the program Modeltest (Posada & Crandall 2001), in conjunction with PAUP 4.0b10, following the Akaike information criterion (AIC). These distances were visualized on a NJ tree calculated using PAUP 4.0b10.
Phylogenetic reconstruction were conducted using Bayesian Analysis (BA), consisting of two Markov chains (2000000 generations each with a sampling frequency of one tree each hundred generations) run in four parallel analyses using Mr. Bayes (Huelsenbeck *et al* 2001). When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in. Only the number of nucleotide substitutions categories was fixed for BA. Phylogenetic analyses were performed on the cluster developed at the

MNHN (17 nodes, 2 Gb RAM per node, 30 AMD 64 bits CPU's for the slave nodes and 4 Xeon 32 bits CPU's for the two master nodes).
To test for potential relationships between geographic (calculated using Geographic Distance Matrix Generator V1.2; Ersts [Internet]) and genetic distances, a Mantel test was performed using zt software (Bonnet & Van de Peer 2002).
Mesquite software (Maddison & Maddison 2007) was used to trace geographic distribution history with parsimony ancestral states reconstruction method, in order to infer potential past events of migration between geographic regions. Specimens were grouped in four different geographic regions, corresponding to sampling sites distant from at most 850 km (Philippines, Solomon Islands, Lord Howe ridge and Vanuatu).
Potential within group structure was also tested between different sampling sites, using Arlequin 3.1 to perform AMOVA (with a 10000 permutations test) for each pair of populations.

Morphological analysis

All specimens used in the molecular analysis were morphologically analysed. Specimens were clustered in morphologically homogeneous groups, without taking into account the groups obtained with the molecular analysis. A traditional taxonomic approach was used, i.e. the shells were grouped into morphological species based on similarity of conchological characters, such as shell shape, proportions, and details of sculpture.

Bathymetric analyses

Depth of collection of each specimen has been used as a potential explanatory variable of genetic differentiation. All *Bathytoma* specimens were collected by dredging or trawling, and consequently the precise depth of collection is not known. The available information is limited to the start and the end of the dredging or trawling, as shown in Appendix 1. One sampling site can thus represent a line of several hundred meters, associated to a variation of the depth between the start and the end of the operation. The mean between these two depths was calculated for each specimen, and bathymetric differences between each genetically defined group were tested. Only genetic groups including more than six specimens were analysed, using Mann-Withney tests (or Kruskall-Wallis tests for more than two samples) implemented in Xlstat (XLSTAT2008, AddinSoft, http://www.xlstat.com).

Results

Genetic analysis

All specimens were sequenced for the COI gene, resulting in a 658 bp fragment without indel. The Kimura-3-parameter model (K81uf+I+G, with I = 0.579 and G = 0.971) was defined as the best fitting model. Ninety-four different haplotypes were found among the 138 analysed specimens, displaying 447 polymorphic loci, and a high haplotypic diversity (0.989). Distribution of genetic distances is shown in Fig. 2: the multi-modal distribution allows the recognition of several gaps between two successive modes of genetic distances. A first gap can be detected between 2.8 and 3.2% of divergence, and the use of this gap on the NJ tree delimits 14 groups of specimens (for the clarity of the discussion, we will use the term "group" to refer to group of specimens as well as separate specimens) characterized by genetic distances within groups inferior to 2.8% and genetic distances between groups

superior to 3.2% (Fig. 3). Similarly, the others gaps between 4.6 and 5.6%, 8.2 and 8.4% and 11.6 and 11.8% can be used to define respectively 13, 7 and 4 groups (results not shown).

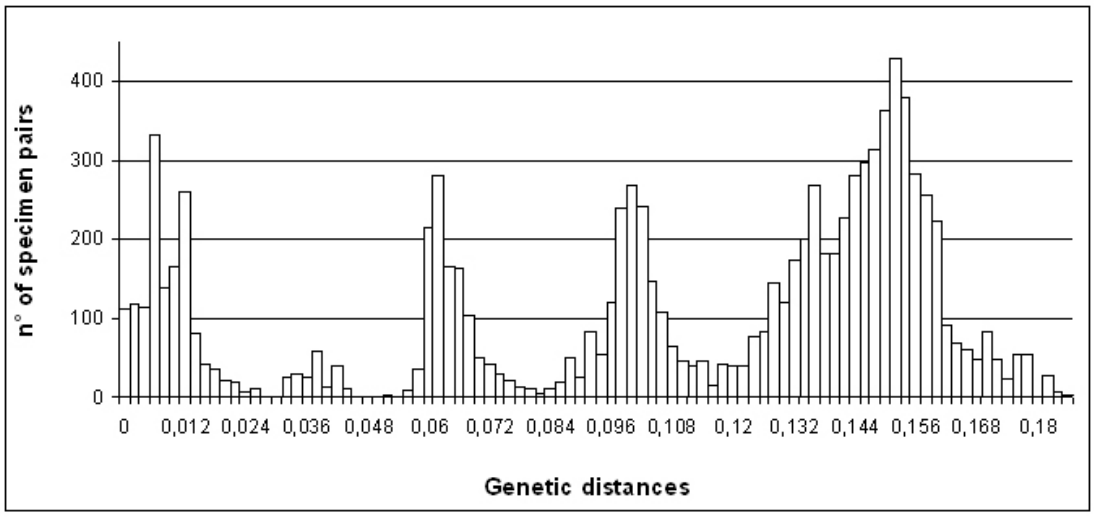


Fig. 2. Histogram of genetic distances for COI gene.

Fifty-one specimens, corresponding to the 14 groups defined with this first threshold of COI genetic distances were also sequenced for the ITS2 gene, resulting in a 557 bp fragment after alignment. All the cruises of collection within each group are represented. Thirty-seven different haplotypes were found, displaying 374 polymorphic loci, and a high haplotypic diversity (0.984). As with the COI gene, the phylogenetic tree obtained with the ITS2 gene is characterized by long branches at the end of which genetically similar specimens are clustered (Fig. 3). All groups defined with the first threshold of COI genetic distances (around 3%), except one, were found identical with the ITS2 gene (see details below). Conversely, most of the groups defined with the three other thresholds of COI genetic distances were not found with the ITS2 gene.

Out of the 14 groups sequenced for both genes, 13 were found identical, *i.e.* specimens clustered in the same group with the COI gene were also clustered together with the ITS2 gene. These groups defined by both COI and ITS2 genes are numbered from G1 to G12. Furthermore, phylogenetic analyses support the monophyly of these thirteen groups (posterior probabilities PP $\geq$ 0.98), for both genes independently (except for G7 with ITS2 gene: PP = 0.94). Each group consists of 1 to 44 specimens. Moreover, the specimens included within one group are generally issued from the same geographic region, except the group G5 that includes specimens collected in the Philippines and in the Solomon Islands, and the group G7 that includes specimens collected in Vanuatu and on the Lord Howe ridge. Eight groups include specimens from the same region but collected during different cruises (Salomon 2 and 3 for the G3, G5, G6, G11 and G13a, Panglao 04 and Panglao 05 for G9 and Panglao 05 and Aurora 07 for G10 and G12) (see Fig. 1).

In one case, we found a contradiction between the groups defined with the first threshold of COI genetic distances and the results obtained with the ITS2 genes. Two groups are defined with the COI gene (noted G13a and G13b, Fig. 3), one including specimens collected in the Solomon Islands and the other including specimens collected in the Philippines. All these specimens were not discriminated using the ITS2 gene.

Gene flows were estimated for the unique group (G12) including more than 5 specimens from different sampling sites. No structure was found between all pairs of sampling sites (p-value not significant; maximal F_{st} (0.07) found between stations CP2660 and CP2735). This

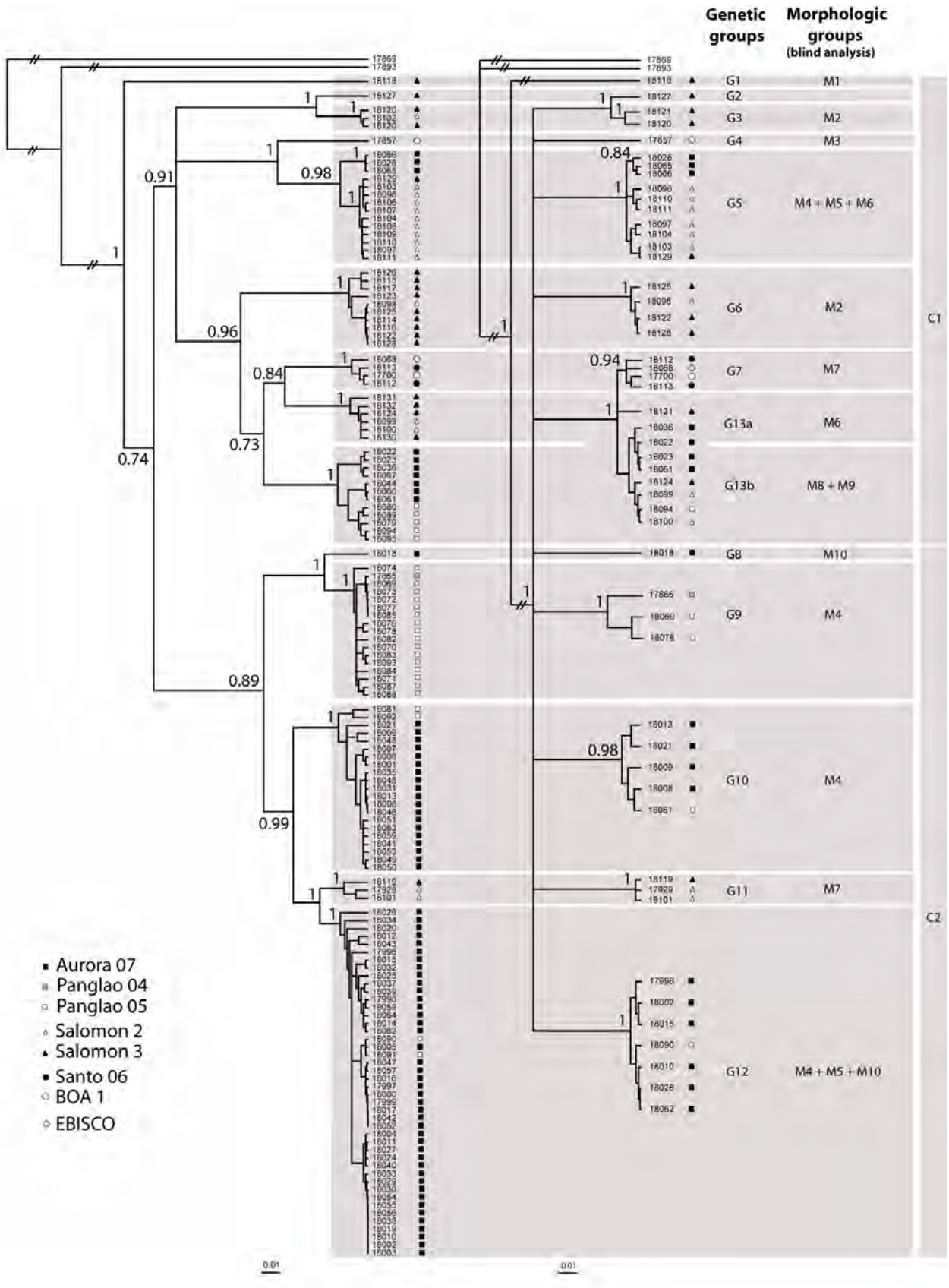


Fig. 3. Bayesian tree for COI (A) and ITS2 gene (B). Posterior Probabilities (superior to 0.5) are given for each node. Grey boxes represent genetic groups, numbered from G1 to G14b. Corresponding morphological groups are given for each genetic groups. Clades C1 and C2 are also figured.

surprising result for a non-planktotrophic species can reflect the small geographic distance between the compared sites (13 km).

Geographic distances between sites of collection are visualised as a function of COI genetic distances (Fig. 4). As previously noted, intragroup genetic distances are found in most cases between specimens from the same region. More interestingly, for specimens of different genetic groups, ranges of genetic distances seem to be similar at different geographic scale (within a geographic region but also between different geographic regions, whatever the geographic distance considered; Fig. 4). However, a Mantel's test shows up a strong correlation ($r = 0.509$, $p = 3.10^{-4}$) between genetic and geographic distances for specimens from different groups. Specimens from the same geographic region tend to be genetically more similar than specimens from the different geographic regions, even when different genetic groups are considered. However, this tendency is not a general rule, as genetic groups from the same geographic region do not cluster together in the phylogenetic tree.

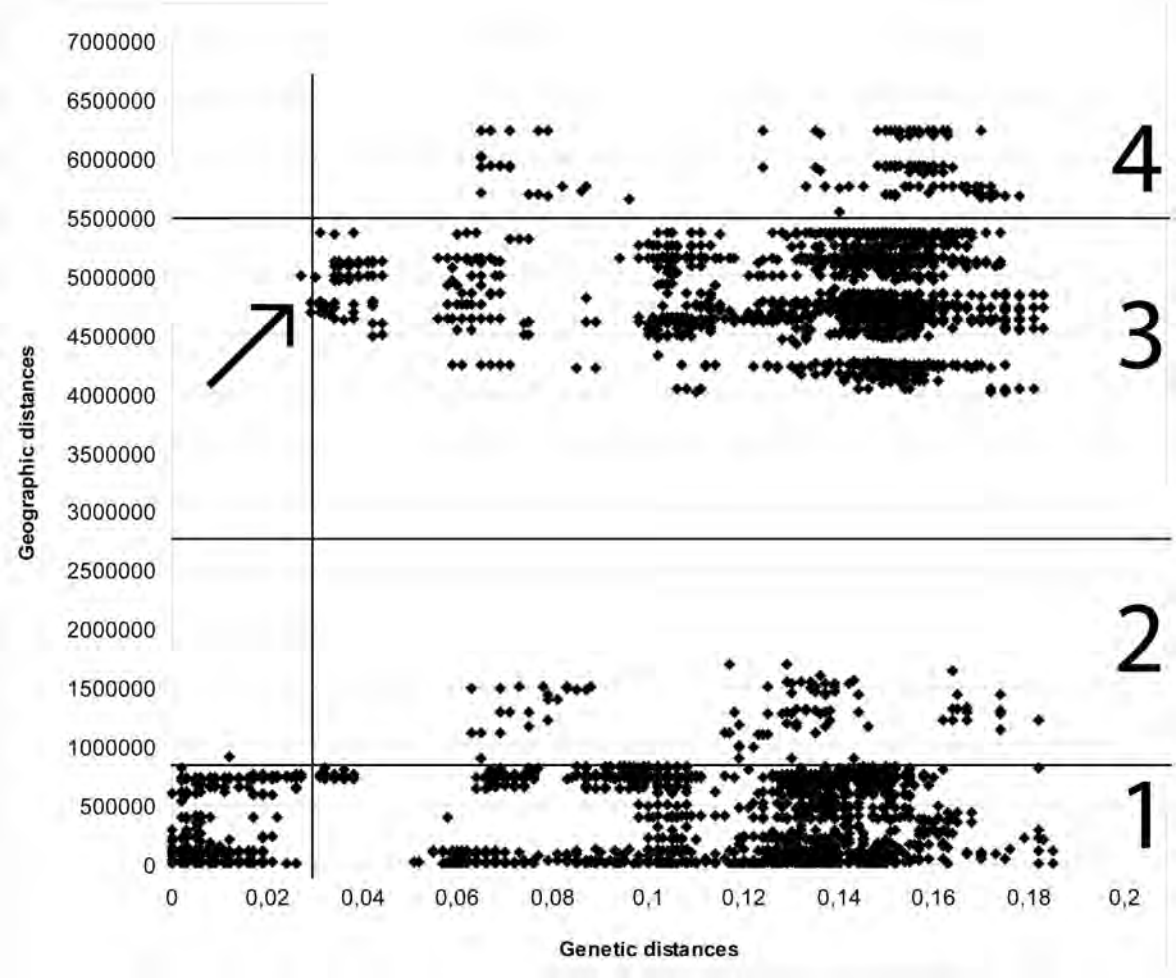


Fig. 4. Geographic distances as a function of COI genetic distances for all specimens. The vertical line represents the first threshold of COI genetic distances. The horizontal lines delimit four geographic regions (1: intra-regions; 2: inter Vanuatu-Solomon Islands-Lord Howe ridge; 3: inter Philippines-Solomon Islands; 4: Inter Philippines-Vanuatu and Philippines-Lord Howe ridge). The arrow indicates dots that correspond to Philippines-Solomon Islands comparisons within the genetic group G5.

Concordance with morphological analysis

Ten morphological groups were isolated among the same 138 specimens. Morphological identifications were uncertain for 37 specimens, most of them being juveniles (empty cells in the column "Morphological groups" of the Appendix 1). As a result, two morphological groups (M1 and M3), each represented by only one specimen, are completely congruent with molecular results, *i.e.* the morphological group is found in only one genetic group (Fig. 3). Two other morphological groups (M8 and M9) correspond to the same genetic group (G13b). Two other morphological groups are present in two genetic groups each: M2 in G3 and G6 and M7 in G7 and G11. In this case, each genetic group corresponds to a unique morphological group. Finally, the group M4 is present in the groups G5, G9, G10 and G12; M5 is present in G5 and G12; M6 is present in G5 and G13a and M10 in G8 and G12.
In the second analysis of the morphological characters, information provided by the molecular analysis was used to re-evaluate and eventually modify morphological groups defined in the first "blind" analysis. This resulted in recognition of eleven morphs that correspond to the genetic groups G1-G4, G6, G8-G12 and G13b. Two morphs were found within the genetic group G5. Conversely, genetic groups G7 and G13a were not differentiated morphologically: in the light of the genetic results, the morphological differences found in the first approach between these two groups (corresponding to M6 and M7 morphological groups) were reinterpreted as intraspecific variability

Bathymetric analysis

Only genetic groups G5, G6, G9, G10, G12 and G13b were represented by more than six specimens, and thus statistically tested for bathymetric differences. Bathymetric differences were found between most of the groups (Table 1), except between G5, G6 and G13b and between G10 and G12. Groups G5, G6 and G13b on the one hand and G9, G10, G12 on the other hand correspond to two clades (C1 and C2) in the phylogenetic tree (Fig. 3), although they are weakly supported. Kruskal-Wallis test between G5, G6 and G13b was statistically not significant ($p = 0.06$). Conversely, Kruskal-Wallis test between G9, G10 and G12 was statistically significant ($p < 10^{-4}$). Bathymetric differences between C1 and C2 (pooling all genetic groups) were statistically significant ($p < 10^{-4}$).

Table 1 : Bathymetric difference between genetic groups. Results of Mann-Whitney (or Kruskal-Wallis as indicated by the [1]) tests for each comparison

	U	p-value
G5 Vs G6	62.5	0.694
G5 Vs G9	30	0***
G5 Vs G10	52	0.001**
G5 Vs G12	481	0.001**
G5 Vs G13b	81	0.89
G6 Vs G9	9	$<10^{-4}$***
G6 Vs G10	54	0.027*
G6 Vs G12	313	0.036*
G6 Vs G13b	69	0.573
G9 Vs G10	337	$<10^{-4}$***
G9 Vs G12	71	$<10^{-4}$***
G9 Vs G13b	13	$<10^{-4}$***
G10 Vs G12	417	0.511
G10 Vs G13b	68.5	0.027*
G12 Vs G13b	372	0.027*
G5 Vs G6 Vs 13b[1]	15.497	0.06
G9 Vs G10 Vs 12[1]	32.456	$<10^{-4}$***
C1 Vs C2	3456	$<10^{-4}$***

Discussion

Reliability of the species delimitation

Ten groups were recognized with both COI and ITS2 genes and morphological characters (G1-G4, G6, and G8-G12). They correspond to lineages that are divergent and reciprocally monophyletic, satisfying two criteria of species delimitation derived from the lineage species concept. Furthermore, as such criteria are met both with a mitochondrial and a nuclear gene, they should correspond to non-interbreeding species (Avise 2004; Reid *et al* 2006). Among these groups, two correspond to already described species (G9 is *B. tippetti* and G12 is *B.*

atractoides) and eight correspond to new species that are described in the Appendix 2. As the group G8 represents a form of uncertain status (see Appendix 2), it is not described as a new species.
Within the group G13, two very distant and reciprocally monophyletic clades (G13a and G13b) are recognized with the COI gene. They also correspond to two distinct morphs. The lack of resolution for the ITS2 gene can be a consequence of incomplete lineage sorting (Funk & Omland 2003; Hickerson *et al* 2007), because of a smaller effective size for the mitochondrial genome or due to potential effect of selection on these genes (Bazin *et al* 2006). The group G13b has been secondarily linked to a distinct morph and is described as a new species (see Appendix 2). Furthermore, in order to not describe species solely on molecular characters, we choose temporarily to describe these two allopatric groups G7 and G13a, not distinguished in the second morphological analysis, as one species (see Appendix 2). Additional morphological analyses with additional material should be performed to find eventually the differences between these two groups and then describe them as two different species.
Finally, the group G5 includes two distinct morphs that correspond to two reciprocally monophyletic lineages detected within G5 with the COI gene (Fig. 3). Genetic distances between these two clades are just inferior to the threshold that we used (Fig. 4). These two clades within G5 are allopatric, one occurring only in the Philippines and the other in the Solomon Islands, and can thus represent recently diverged species. The lack of resolution for the ITS2 gene can be interpreted as for the G13 group. As data are congruent between COI gene and morphology, these two groups are described as two different species (see Appendix 2). Similarly, within the species G13b, two morphs are recognized, one from Panglao 05 and the other from Aurora 07 cruise. With both genes, these groups are not fully sorted. Pending more data, G13b can be described as one species, and morphs within it as intraspecific morphological variants.
As a result, 14 species-group taxa have been recognized in our material. This evidences both the underexploration of the fauna, because only 2 species were previously known for the studied region, and the importance of sometimes minor morphological characters (Meyer *et al* 2005), since some species well separated by molecular data can only hardly be characterized conchologically (see Appendix 2). Moreover, the genus diversity described in the paper is actually not exhaustive, with some 5-8 species remaining for description but yet lacking genetic data (data not shown).

Patterns of speciation of Bathytoma *in the Indo-Pacific*

Geographic distribution and potential long-distance dispersion

All species represented by several specimens, except groups G7/G13a (but see discussion before), exhibit much reduced distribution area. This result has been already found for other groups with non-planktonic larvae (Kohn & Perron 1994; Paulay & Meyer 2006). If we postulate that observed distribution ranges correspond to the real distribution area of each species, we can infer that common ancestor of all recognized *Bathytoma* species were in Solomon Islands, and that observed pattern of geographic distribution can be explained by eight events of migrations (results not shown). However, as suggested from the multimodal distribution of COI genetic distances (Fig. 2) and by the 5-8 morphological groups not represented in the genetic dataset (data not shown), our analysis is far from including the whole diversity of *Bathytoma*, and distribution area of several species might have been underestimated (Collin 2003). The opposite scenario can thus be also proposed, with an

ancestral species present in the whole sampled area that has experienced several events of speciation, potentially followed by local extinctions.
These two scenarios are both extreme, and alternative scenarios involving widely distributed species resulted from and/or associated to migration events can be proposed. Thus, in each case, the large-scale dispersion is likely to have occurred, which is quiet in contradiction with the dispersal capacities as inferred by the larval shell. Actually, if we use the assumption of a mutation rate of 2.8 per million year (as calculated for COI gene in teguline gastropods by Hellberg & Vacquier (1999)), we can estimate that the first speciation event within our dataset has occurred between 8 and 9 MY ago, a time when *Bathytoma* fossils from New Zealand were characterized by a planktotrophic larvae. It is also interesting to note that between 15 and 10 MY, the Philippines plate has getting closer to Indonesia and Australia (Hall 2002; Gaina & Muller 2007), which can be a helpful opportunity for ancestral populations of *Bathytoma* to expand their geographical distribution and to experience several events of speciation resulting in the actual pattern of diversity. Other historical processes could be mentioned, such as the periodic falls of sea level, by up to 120 m below the present stand, during the glacial intervals of the Plio-Pleistocene epoch (Reid *et al* 2006). However, contrary to littoral species, such a drop of the sea-level is weakly susceptible to have opened new deep pathways for *Bathytoma*, given the large range of bathymetry where the genus is present. Finally, to explain long-distance dispersal of Pacific topshells characterized by a short larval phase, Donald *et al* (2005) suggested that suitable platforms such as macroalgae could be used as rafts by adults to disperse on long distances.

Modes of speciation

Some well supported phylogenetic relationships seem to indicate that several types of speciation may have occurred within the genus *Bathytoma*. Firstly, of the 14 species-group taxa, most are allopatrically isolated. There are only 3 species for which the existence of allopatric populations is considered valid (see Appendix 2), even though in one case there are genetic data only for a Philippine population (not Indonesian one, see *B. netrion* in the Appendix 2), in one case there are morphological differences between the populations not significantly differentiated genetically (*B. episoma*, see Appendix 2), and in one case there are no reliable morphological differences to distinguish genetically different populations at the species level (*B. carnicolor*, see Appendix 2 and the discussion on G7 + G13a above). However, as shown in the phylogenetic tree, other effects already detailed (migration, local extinctions, etc.) have somewhat obscured the allopatric pattern, as we cannot observe a complete congruency between phylogeny and geographic distribution. Once again, such allopatric species can also be the artefactual result of a biased sampling.
Secondly, a species pair (G2 and G3) both include specimens collected during Salomon 3 cruise, and can thus constitute a case of sympatric speciation (Hellberg & Vacquier 1999). Even if the speciation pattern observed can be linked to past events of migration and local extinctions, sympatric speciation are also likely to explain several deeper relationships inferred from the phylogenetic analysis. For example, the relationships (G10 (G11 / G12)), with G10 and G12 sampled in the Philippines and G11 in the Solomon Islands, can be explained by a sympatric speciation in the Philippines followed by the migration of one species in the Solomon islands associated to an allopatric speciation. However, we have to be careful regarding our conclusions concerning sympatric speciation. Even if several species pairs (G5 and G12, G11 and G13b, to cite only two of the numerous examples) were collected in the same site, one sampling site can correspond to several hundreds of meters, and can hide micro-scaled allopatric speciation (Collin 2003, Bierne *et al* 2003, Hyde 2008). Apparent sympatric speciation can thus correspond to parapatric speciation, often linked to local adaptation (Briggs 2006). In this perspective, depth differences found between several species

can be a factor of species structure within a limited geographical area. Bathymetry is furthermore susceptible to explain some deep phylogenetic relationships, as species from the clade 1 significantly live deeper than species from the clade 2. Even though found for deeper organisms, such effect of bathymetry on species diversity has already been reported (Zardus *et al* 2006). Once again, we have to be careful as the collecting gear used does not provide the exact depth of collection for each specimen. Finally, co-occurrence of two species in the same site can also be the result of secondary contact between two species that have diverged allopatrically or by natural selection (Norman *et al* 2007). Bierne *et al*. (2003) give the example of two mussels that co-occurred in the same region but that are characterized by spawning asynchrony.

Conclusion

The integrative taxonomy approach quickly provided robust hypotheses of species delimitations, allowing the description of eleven new species. DNA clearly presents advantages, as it is genetically determined, easy to sequence given the actual technical facilities, and applicable to all stages of development. Juveniles specimens were clearly a problem for the morphological analysis, as characters used in alpha-taxonomy are mostly displayed by adults.
This analysis is a pre-requisite to subsequent ones that directly lies on robust species delimitation. In this way, we were able to propose scenario of speciation within the South-West Pacific range of the genus *Bathytoma*. Contrary to what have been hypothesized, not only cases of allopatric speciation were identified in *Bathytoma*, but also cases of sympatric speciation. Moreover, the whole diversification within the genus *Bathytoma* seems to be the product of multiple processes in space and time (Meyer *et al* 2005, Imron 2007, Kirkendale & Meyer 2004, Palumbi 2005). More investigations are needed, especially in the geographic and taxonomic coverage of the group, to clearly circumscribe species and their distribution and provide the basis of discussion on speciation patterns.

Acknowledgments
Key material for molecular studies originates from the following expeditions. (A) Philippines. The PANGLAO 2004 Marine Biodiversity Project was a joint project between University of San Carlos, Cebu City (USC; co-PI Danilo Largo) and MNHN (co-PI Philippe Bouchet), with funding from the Total Foundation and the French Ministry of Foreign Affairs. The follow-up PANGLAO 2005 cruise on board M/V DA-BFAR associated the USC (co-PI Danilo Largo), MNHN (co-PI Philippe Bouchet) and the Philippines Bureau of Fisheries and Aquatic Research (BFAR). The AURORA 2007 cruise also on board M/V DA-BFAR associated the National Museum of the Philippines (NMP, co-PI Marivene Manuel), MNHN (co-PI Philippe Bouchet) and BFAR, and was made possible through a grant from the Lounsbery Foundation. (B) Vanuatu. Material originates from the MNHN-IRD-PNI Santo 2006 expedition, which was made possible by grants from the Total Foundation and the Niarchos Foundation. (C) Coral Sea and Solomon Islands. The EBISCO and SALOMON cruises took place on board R/V Alis deployed from Nouméa by the Institut de Recherche pour le Développement (IRD).
Bertrand Richer de Forges was cruise leader for the Solomons, Coral Sea and Vanuatu expeditions. Yuri Kantor processed specimens during several of the Philippines and Vanuatu expeditions. Work was conducted in Paris by Alexander Sysoev during several visits under the MNHN visiting curatorships programme.

References

Amitrov, O. V. (1973). Turrid pozdnego Eotzena - rannego Miotzena juga SSSR. *Trudy Paleontologicheskogo Institutom, Akademij Nauk SSSR, 135*, 1-211.
Avise, J. C. (2004). *Molecular markers, natural history, and evolution. 2nd edn*. Sunderland, MA, USA.: Sinauer Associates.
Bazin, B., Glemin, S. & Galtier, N. (2006). Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals. *Science*, *312*, 570-572.
Bernasconi, M. P., & Robba, E.. (1984). The Pliocene Turridae from Western Liguria. I. Clavinae, Turrinae, Turriculinae, Crassispirinae, Borsoniinae, Clathurellinae. *Bollettino del Museo Regionale di Scienze Naturali (Torino), 2*, 257-358.
Beu, A. G., & Maxwell, P. A.. (1990). Cenozoic Mollusca of New Zealand. *New Zealand Geological Survey Paleontological Bulletin, 58*, 1-518.
Bichain, J. M., Boisselier, M. C., Bouchet, P. & Samadi, S. (2007). Delimiting species in the genus *Bythinella* (Mollusca: Caenogastropoda: Rissooidea): molecular and morphometric approachs. *Malacologia, 49*, 291-311.
Bierne, N., Bonhomme, F. & David, P. (2003). Habitat preference and the marine-speciation paradox. *Proceedings of the Royal Society of London Serie B, 270*, 1399-1406.
Bonnet, B. & Van de Peer, Y. (2002). zt: a software tool for simple and partial Mantel tests. *Journal of Statistical software, 7*, 1-12.
Briggs, J. C. (2006). Proximate sources of marine biodiversity. *Journal of Biogeography, 33*, 1-10.
Brookes, J. I. & Rochette, R. (2007). Mechanism of a plastic phenotypic response: predator-induced shell thickening in the intertidal gastropod *Littorina obtusata*. *Journal of Evolutionary Biology, 20*, 1015-1027.
Chombard, C., Boury-Esnault, N. & Tillier, S. (1998). Reassessment of Homology of Morphological Characters in Tetractinellid Sponges Based on Molecular Data. *Systematic Biology*, 351-366.
Collin, R. (2003). Phylogenetic Relationships Among Calyptraeid Gastropods and Their Implications for the Biogeography of Marine Speciation. *Systematic Biology, 52*, 618-640.
De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In D. J. Howard & S. H. Berlocher (Eds) *Endless forms. Species and speciation* (pp. 57-75). Oxford: Oxford University Press.
Donald, K. M., Kennedy, M. & Spencer, H. S. (2005). Cladogenesis as a result of long-distance rafting events in South Pacific topshells (Gastropoda, Trochidae). *Evolution, 59*, 1701-1711.
Duda, T. F. & Kohn, A. J. (2005). Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. *Molecular Phylogenetics and Evolution, 34*, 257-272.
Duda, T. F. & Rolan, E. (2005). Explosive radiation of Cape Verde *Conus*, a marine species flock. *Molecular Ecology, 14*, 267-272.
Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online, 1*, 47-50.
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.
Funk, D. J. & Omland, K. E. (2004). Species-level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics, 34*, 397-423.
Gaina, C. & Müller, D. (2007). Cenozoic tectonic and depth/age evolution of the Indonesian gateway and associated back-arc basins. *Earth-Science Reviews, 83*, 177-203.
Hall, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences, 20*, 353-431.
Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95-98.
Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003). Biological identifications through DNA Barcodes. *Proceedings of the Royal Society of London B, 270*, 313-21.
Hellberg, M. E. & Vacquier, V. D. (1999). Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Molecular Biology and Evolution, 16*, 839-848.

Hickerson, J. H., Meyer, C. P. & Moritz, C. (2006). DNA Barcoding Will Often Fail to Discover New Animal Species over Broad Parameter Space. *Systematic Biology*, *55*, 729-739.
Huelsenbeck, J. P., Ronquist, F. & Hall, B. (2001). MrBayes: bayesian inference of phylogeny. *Bioinformatics*, *17*, 754-755.
Hyde, J. R., Kimbrell, C. A., Budrick, J. E., Lynn, E. A. & R. D. Vetter, R. D. (2008). Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. *Molecular Ecology, 17,* 1122-1136.
Imron, B. J., Hale, P., Degnan, B. M. & Degnan, S. M. (2007). Pleistocene isolation and recent gene flow in *Haliotis asinina*, an Indo-Pacific vetigastropod with limited dispersal capacity. *Molecular Ecology, 16*, 289-304.
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: Plenum Press.
Kilburn, R.N. (1986). Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique.
Part 3. Subfamily Borsoniinae. *Annals of the Natal Museum*, 27, 633-720.
Kirkendale, L. A. & Meyer, C. P. (2004). Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal-driven marine system. *Molecular Ecology, 13*, 2749-2762.
Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia, 420*, 73-90.
Kohn, A. J. & Perron, F. E. (1994). *Life history and biogeography patterns in Conus*. Oxford: Clarendon Press.
Kosuge, S. (1986). Report on the family Turridae collected along the north-western coast of Australia (Gastropoda) (1). *Bulletin of the Institute of Malacology Tokyo*, 2, 80-90, pls. 30-34.
MacNeil, F. S., & Dockery, D. T. III. (1984). Lower Oligocene Gastropoda, Scaphopoda, and Cephalopoda of the Vicksburg Group in Mississippi. *Mississippi Geological, Economic and Topographical Survey, Bulletin, 124,* 1-415.
Maddison, W. P. & Maddison, D. R. (2007). Mesquite: *a modular system for evolutionary analysis. Version 2.01*. Available via http://mesquiteproject.org
Mallet, J. (2001). The speciation revolution. *Journal of Evolutionary Biology*, *14*, 887-888.
Meyer, C. P., Geller, J. B. & Paulay, G. (2005). Fine scale endemism on coral reefs: archipelagic differentiation in Turbinid gastropods. *Evolution, (1), , pp. , 59*, 113-125.
Noda, H. (1980). Molluscan fossils from the Ryukyu Islands, southwestern Japan. Part 1. Gastropoda and Pelecypoda from the Shinzato Formation in southeastern part of Okinawa-jima. *Science Report, Institute of Geoscience, University of Tsukuba, section B, 1,* 1-95.
Noda, H. (1988). Molluscan fossils from the Ryukyu Islands, southwest Japan. Part 2. Gastropoda and Pelecypoda from the Shinzato Formation in the middle part of Okinawa-jima. *Science Report, Institute of Geoscience, University of Tsukuba, section B, 9*, 29-85.
Norman, J. A., Rheindt, F. E., Rowe, D. L. & Christidis, L. (2007). Speciation dynamics in the Australo-Papuan Meliphaga honeyeaters. *Molecular Phylogenetics and Evolution*, *42*,
Palumbi, S. R. & Lessios, H. A. (2005). Evolutionary animation: How do molecular phylogenies compare to Mayr's reconstruction of speciation patterns in the sea? *Proceedings of the National Academy of Sciences*, *102*, 6566–6572.
Paulay, G. & Meyer, C. P. (2006). Dispersal and divergence across the greatest ocean region: Do larvae matter? *Integrative and Comparative Biology*, *46*, 269-281.
Posada, D. & Crandall, K. A. (2001). Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). *Molecular Biology and Evolution*, *18*, 897-906.
Powell, A. W. B. (1966). The molluscan families Speightiidae and Turridae. An evaluation of the valid taxa, both recent and fossil, with lists of characteristics species. *Bulletin of the Auckland Institute and Museum, 5*, 5-184.
Puillandre, N., Samadi, S., Boisselier, M. C., Sysoev, A. V., Kantor, Y. I., Cruaud, C., Couloux, A. & Bouchet, P. Starting to unravel the toxoglossan knot: Molecular phylogeny of the "turrids" (Neogastropoda: Conoidea). *Molecular Phylogenetics and Evolution*, *47*, 1122–1134.

Reid, D. G., Lal, K., Mackenzie-Dodds, J., Kaligis, F., Littlewood, D. T. J. & Williams, S. T. (2006). Comparative phylogeography and species boundaries in *Echinolittorina* snails in the central Indo-West Pacific. *Journal of Biogeography*, *33*, 990-1006.
Samadi, S. & Barberousse, A. (2006). The tree, the network, and the species. *Biological Journal of the Linnean Society*, *89*, 509-521.
Samadi, S., Roumegoux, A., Bargues, M. D., Mas-Coma, S., Yong, M. & Pointier, J. P. (2000). Morphological studies of Lymnaeid snails from the human Fasciolasis endemic zone of Bolivia. *Journal of Molluscan Studies*, *2000*, 31-44.
Sites, J. W. & Marshall, J. C. (2004). Operational criteria for delimiting species. *Annual Review of Ecology Evolution and Systematics*, *35*, 199-227.
Swofford, D. L. 2002. PAUP*. *Phylogenetic Analysis Using Parsimony* and other methods*. Sunderland: Sinauer Associates.
Sysoev, A.V. & Bouchet, P. (2001). New and uncommon turriform gastropods (Gastropoda: Conoidea) from the South-West Pacific. *Mémoires du Muséum National d'Histoire Naturelle, Paris*, série A, Zoologie, 185, 271-320.
Vallejo Jr., J. (2005). Inferring the mode of speciation in Indo-West Pacific *Conus* (Gastropoda: Conidae). *Journal of Biogeography*, *32*, 1429-1439.
Vera-Peláez, J. L. & Lozano-Francisco, M. C. (2004). Contribución al conocimiento del género *Bathytoma* Harris & Burrows, 1891 (Gastropoda, Turridae) en Japón, Taiwán y Filipinas con la descripción de tres especies nuevas. *Pliocénica*, *4*,
Watson, R.B. (1881). Mollusca of H.M.S. Challenger Expedition. Parts VIII-X. *Journal of the Linnean Society (London)*, 15, 388-475.
Watson, R.B. (1886). Report on the Scaphopoda and Gasteropoda collected by the H. M. S. Challenger during the years 1873-76. *Report on the Scientific Results of the Voyage of the 'Challenger' (Zoology)*, 15: 1-756, 50 pls.
White, T. J., Bruns, T., Lee, S. & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In D. H. G. M. A. Innis, J. J. Sninsky, and T. J. White (Eds) *PCR Protocols: A Guide to Methods and Applications* (pp. 315-322). San Diego: Academic Press, Inc.
Wilson, B. (1994). *Australian Marine Shells. Vol. 2 Prosobranch Gastropods. Part Two (Neogastropods)*. Kallaroo, Western Australia: Odyssey Publishing, 370 pp.
Yeap, K. L., Black, R. & Johnson, M. S. (2001). The complexity of phenotypic plasticity in the intertidal snail *Nodilittorina australis*. *Biological Journal of the Linnean Society*, *72*, 63-76.
Zardus, J. D., Etter, R. J., Chase, M. R., Rex, M. A. & Boyle, E. E. (2006). Bathymetric and geographic population structure in the pan-Atlantic deep-sea bivalve *Deminucula atacellana* (Schenck, 1939). *Molecular Ecology*, *15*, 639-651.

Appendix 1: Identification number, cruise, station, depth, coordinates, morphological and genetic groups (for COI and ITS2 genes) are given for each specimen.

N°	Cruise	Station	Depth	Lat	Long	Morpho	COI	Species
17700	BOA1	CP2462	618-641	16°37.5'S	167°57.4'E		G7	*Bathytoma carnicolor*
17857	EBISCO	CP2551	637-650	21°06'S	158°35'E	M3	G4	*Bathytoma neocaledonica*
17865	Panglao 04	P2	400	9°39'N	123°44'E	M4	G9	*Bathytoma tippetti*
17929	Salomon 2	CP2186	487-541	8°17,0'S	160°00,0'E	M7	G11	*Bathytoma paratractoides*
17995	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G10	*Bathytoma stenos*
17996	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M4	G12	*Bathytoma atractoides*
17997	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M10	G12	*Bathytoma atractoides*
17998	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M4	G12	*Bathytoma atractoides*
17999	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M4	G12	*Bathytoma atractoides*
18000	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18001	Aurora 07	CC2744	418-456	15°58.75N	121°50.36E	M4	G10	*Bathytoma stenos*
18002	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18003	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18004	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18005	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18007	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M4	G10	*Bathytoma stenos*
18008	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E	M4	G10	*Bathytoma stenos*
18009	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G10	*Bathytoma stenos*
18010	Aurora 07	CP2659	460-480	15°57.43N	121°50.26E	M10	G12	*Bathytoma atractoides*
18011	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E	M4	G12	*Bathytoma atractoides*
18012	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E		G12	*Bathytoma atractoides*
18013	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E	M4	G10	*Bathytoma stenos*
18014	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E		G12	*Bathytoma atractoides*
18015	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E		G12	*Bathytoma atractoides*
18016	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E		G12	*Bathytoma atractoides*
18017	Aurora 07	CP2660	506-542	15°52.06N	121°49.6E		G12	*Bathytoma atractoides*
18018	Aurora 07	CP2668	467-576	15°47.86N	121°45.65E	M10	G8	*Bathytoma sp.*
18019	Aurora 07	CP2658	422-431	15°59.45N	121°50.68E		G12	*Bathytoma atractoides*
18020	Aurora 07	CP2658	467-576	15°47.86N	121°45.65E		G12	*Bathytoma atractoides*
18021	Aurora 07	CP2673	431-493	14°59.83N	121°43.90E	M4	G10	*Bathytoma stenos*
18022	Aurora 07	CP2663	562	15°47.18N	121°46.09E	M8	G13b	*Bathytoma episoma*
18023	Aurora 07	CP2699	541-583	14°50.04N	123°34.50E		G13b	*Bathytoma episoma*
18024	Aurora 07	CP2707	368-442	15°04.79N	121°41.79E	M10	G12	*Bathytoma atractoides*
18025	Aurora 07	CP2707	368-442	15°04.79N	121°41.79E	M4	G12	*Bathytoma atractoides*
18026	Aurora 07	CP2707	368-442	15°04.79N	121°41.79E		G12	*Bathytoma atractoides*
18027	Aurora 07	CP2725	620-627	15°15.01N	121°36.18E		G12	*Bathytoma atractoides*
18028	Aurora 07	CP2725	620-627	15°15.01N	121°36.18E	M4	G5	*Bathytoma netrion*
18029	Aurora 07	CP2725	620-627	15°15.01N	121°36.18E	M4	G12	*Bathytoma atractoides*
18030	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18031	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G10	*Bathytoma stenos*
18032	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18033	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18034	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18035	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G10	*Bathytoma stenos*
18036	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M8	G13b	*Bathytoma episoma*
18037	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18038	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18039	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18040	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18041	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E		G10	*Bathytoma stenos*
18042	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18043	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M4	G12	*Bathytoma atractoides*
18044	Aurora 07	CP2734	453-460	15°56.41N	121°48.71E	M8	G13b	*Bathytoma episoma*
18045	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18046	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E		G10	*Bathytoma stenos*
18047	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18048	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18049	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18050	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18051	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18052	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18053	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18054	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E		G12	*Bathytoma atractoides*
18055	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18056	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18057	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18058	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G12	*Bathytoma atractoides*
18059	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M4	G10	*Bathytoma stenos*
18060	Aurora 07	CP2735	422-431	15°59.02N	121°50.25E	M8	G13b	*Bathytoma episoma*
18061	Aurora 07	CC2744	418-456	15°58.75N	121°50.36E	M8	G13b	*Bathytoma episoma*
18062	Aurora 07	CP2749	473	15°56.38N	121°49.46E	M4	G12	*Bathytoma atractoides*
18063	Aurora 07	CP2749	473	15°56.38N	121°49.46E	M4	G10	*Bathytoma stenos*
18064	Aurora 07	CP2749	473	15°56.38N	121°49.46E	M4	G12	*Bathytoma atractoides*
18065	Aurora 07	CP2750	518-538	15°53.81N	121°51.55E	M5	G5	*Bathytoma netrion*
18066	Aurora 07	CP2750	518-538	15°53.81N	121°51.55E	M5	G5	*Bathytoma netrion*
18067	Aurora 07	CP2750	518-538	15°53.81N	121°51.55E	M8	G13b	*Bathytoma episoma*

18068	EBISCO	DW2626	728-739	21°06'S	160°49'E		G7	*Bathytoma carnicolor*
18069	Panglao 05	CP2341	544-712	9°24.5'N	123°49.7'E	M4	G9	*Bathytoma tippetti*
18070	Panglao 05	CP2332	396-418	9°38.8'N	123°45.9'E	M4	G9	*Bathytoma tippetti*
18071	Panglao 05	CP2332	396-418	9°38.8'N	123°45.9'E	M4	G9	*Bathytoma tippetti*
18072	Panglao 05	CP2340	271-318	9°29.4'N	123°44.4'E	M4	G9	*Bathytoma tippetti*
18073	Panglao 05	CP2340	271-318	9°29.4'N	123°44.4'E	M4	G9	*Bathytoma tippetti*
18074	Panglao 05	CP2348	196-216	9°29.6'N	123°52.5'E	M4	G9	*Bathytoma tippetti*
18076	Panglao 05	CP2372	255-301	8°38.7'N	123°16.0'E	M4	G9	*Bathytoma tippetti*
18077	Panglao 05	CP2383	338-351	8°44.7'N	123°18.5'E	M4	G9	*Bathytoma tippetti*
18078	Panglao 05	CP2383	338-351	8°44.7'N	123°18.5'E	M4	G9	*Bathytoma tippetti*
18079	Panglao 05	CP2388	762-786	9°26.9'N	123°34.5'E	M9	G13b	*Bathytoma episoma*
18080	Panglao 05	CP2388	762-786	9°26.9'N	123°34.5'E	M9	G13b	*Bathytoma episoma*
18081	Panglao 05	CP2388	762-786	9°26.9'N	123°34.5'E		G10	*Bathytoma stenos*
18082	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E	M4	G9	*Bathytoma tippetti*
18083	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E	M4	G9	*Bathytoma tippetti*
18084	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E	M4	G9	*Bathytoma tippetti*
18085	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E	M4	G9	*Bathytoma tippetti*
18087	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E		G9	*Bathytoma tippetti*
18088	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E		G9	*Bathytoma tippetti*
18089	Panglao 05	CP2395	382-434	9°36.2'N	123°43.8'E	M9	G13b	*Bathytoma episoma*
18090	Panglao 05	CP2334	606-631	9°37.5'N	123°40.2'E	M5	G12	*Bathytoma atractoides*
18091	Panglao 05	CP2334	606-631	9°37.5'N	123°40.2'E	M5	G12	*Bathytoma atractoides*
18092	Panglao 05	CP2334	606-631	9°37.5'N	123°40.2'E		G10	*Bathytoma stenos*
18093	Panglao 05	CP2332	396-418	9°38.8'N	123°45.9'E	M4	G9	*Bathytoma tippetti*
18094	Panglao 05	CP2333	584-596	9°38.2'N	123°43.5'E	M9	G13b	*Bathytoma episoma*
18095	Panglao 05	CP2333	584-596	9°38.2'N	123°43.5'E	M9	G13b	*Bathytoma episoma*
18096	Salomon 2	CP2184	464-523	8°16,9'S	159°59,7'E	M6	G5	*Bathytoma solomonensis*
18097	Salomon 2	CP2245	582-609	7°43,1'S	156°26,0'E	M6	G5	*Bathytoma solomonensis*
18098	Salomon 2	CP2227	508-522	6°37,2'S	156°12,7'E		G6	*Bathytoma consors*
18099	Salomon 2	CP2219	650-836	7°58,3'S	157°34,4'E	M6	G13a	*Bathytoma carnicolor*
18100	Salomon 2	CP2219	650-836	7°58,3'S	157°34,4'E	M6	G13a	*Bathytoma carnicolor*
18101	Salomon 2	CP2186	487-541	8°17,0'S	160°00,0'E	M7	G11	*Bathytoma paratractoides*
18102	Salomon 2	CP2212	400-475	7°37,8'S	157°41,7'E	M2	G3	*Bathytoma badifasciata*
18103	Salomon 2	CP2288	509-520	8°36,3'S	157°26,5'E	M6	G5	*Bathytoma solomonensis*
18104	Salomon 2	CP2214	550-682	7°41,6'S	157°43,8'E	M6	G5	*Bathytoma solomonensis*
18106	Salomon 2	CP2284	195-197	8°38,4'S	157°21,5'E	M6	G5	*Bathytoma solomonensis*
18107	Salomon 2	CP2213	495-650	7°38,7'S	157°42,9'E	M6	G5	*Bathytoma solomonensis*
18108	Salomon 2	CP2263	485-520	7°54,8'S	156°51,3'E		G5	*Bathytoma solomonensis*
18109	Salomon 2	CP2263	485-520	7°54,8'S	156°51,3'E		G5	*Bathytoma solomonensis*
18110	Salomon 2	CP2263	485-520	7°54,8'S	156°51,3'E	M6	G5	*Bathytoma solomonensis*
18111	Salomon 2	CP2263	485-520	7°54,8'S	156°51,3'E	M6	G5	*Bathytoma solomonensis*
18112	Santo 06	AT122		16°36,0S	167°59,4E	M7	G7	*Bathytoma carnicolor*
18113	Santo 06	AT122		16°36,0S	167°59,4E	M7	G7	*Bathytoma carnicolor*
18114	Salomon 3	CP2834	545-753	10°45,70	162°19,41	M2	G6	*Bathytoma consors*
18115	Salomon 3	CP2839	506-567	10°24,86	161°18,79	M2	G6	*Bathytoma consors*
18116	Salomon 3	CP2838	581-510	10°26,92	161°21,65	M2	G6	*Bathytoma consors*
18117	Salomon 3	CP2838	581-510	10°26,92	161°21,65		G6	*Bathytoma consors*
18118	Salomon 3	DW2829	220-292	10°46,6	162°18,9	M1	G1	*Bathytoma cranaos*
18119	Salomon 3	CP2768	447-503	09°25,1	160°29,4		G11	*Bathytoma paratractoides*
18120	Salomon 3	CP2857	473-505	9°43,12	160°47,50	M2	G3	*Bathytoma badifasciata*
18121	Salomon 3	CP2857	473-505	9°43,12	160°47,50		G3	*Bathytoma badifasciata*
18122	Salomon 3	CP2767	416-425	09°19,222	160°07,245	M2	G6	*Bathytoma consors*
18123	Salomon 3	CP2857	473-505	9°43,12	160°47,50	M2	G6	*Bathytoma consors*
18124	Salomon 3	CP2835	735-862	10°43,22	162°20,17		G13a	*Bathytoma carnicolor*
18125	Salomon 3	CP2848	414-456	9°32,9	160°45,7	M2	G6	*Bathytoma consors*
18126	Salomon 3	CP2848	414-456	9°32,9	160°45,7	M2	G6	*Bathytoma consors*
18127	Salomon 3	CP2828	414-456	9°32,9	160°45,7		G2	*Bathytoma punicea*
18128	Salomon 3	CP2833	585-565	10°44,45	162°20,10	M2	G6	*Bathytoma consors*
18129	Salomon 3	CP2833	585-565	10°44,45	162°20,10	M6	G5	*Bathytoma solomonensis*
18130	Salomon 3	CP2850	502-621	9°35,31	160°44,93		G13a	*Bathytoma carnicolor*
18131	Salomon 3	CP2850	502-621	9°35,31	160°44,93		G13a	*Bathytoma carnicolor*
18132	Salomon 3	CP2850	502-621	9°35,31	160°44,93		G13a	*Bathytoma carnicolor*
17893	Panglao 04	T41						*Mitromorpha sp.*
17869	Panglao 04	S21						*Etrema sp.*

Appendix 2: Taxonomic description of the species

Genus *Bathytoma* Harris & Burrows, 1891
Type species (by monotypy): *Murex cataphractus* Brocchi, 1814

Powell (1966) recognized four subgenera in *Bathytoma*, an opinion later followed by Kilburn (1986). However, presently no clear differences can be drawn between these subgenera, and their validity needs to be re-evaluated. The molecular data do not show distinct clades that can be conchologically associated with existing nominal subgenera. Therefore, we prefer not to use any subgeneric division within *Bathytoma*.

Beyond the 14 species reprented by live-taken material, the material we have examined from the West Pacific also contains specimens of what appears to be additional species represented by only empty shells, or worn or juvenile specimens. These putative new taxa are not considered here.

All material is in MNHN (Muséum National d'Histoire Naturelle, Paris) unless stated otherwise.

Abbreviations:
AH – aperture height
dd – dead-collected specimen
LW – last whorl height
lv – live-collected specimen
SD – shell diameter
SH – shell height

Bathytoma atractoides (Watson, 1881) (Fig. 1 A-F)

Pleurotoma (Genota) atractoides Watson, 1881: 407; Watson 1886: 301, pl. 20, fig. 8.
Bathytoma atractoides (part.) – Sysoev & Bouchet 2001: 292, figs. 84-85 (holotype), not figs. 86-96.
Not *Bathytoma atractoides* – Kosuge 1986: 85, pl. 30, figs. 5, 6; Wilson 1994: 182, pl. 39, figs. 4a-c; Vera-Peláez 2004: 109-112, pl. 3, figs. 6-9.

Type locality. Philippines, 9°26' N, 123°45' E, 375 fm (686 m).

Material examined. Holotype (The Natural History Museum, London, BMNH 1887.2.9.985); Philippines: Bohol Sea. PANGLAO 2005, sta. CP2333, 9º38.2' N, 123º43.5' E, 584-596 m, 1 dd; sta. CP2334, 9º37.5' N, 123º40.2' E, 606-631 m, 2 lv; sta. CP2335, 9º34.3' N, 123º37.8' E, 729-733 m, 2 dd; sta. CP2351, 9º30.7' N, 124º03.0' E, 810-812 m, 3 dd; sta. CP2352, 9º27.3' N, 124º03.1' E, 923-1260 m, 4 dd; sta. CP2385, 8º51.0' N, 123º10.0' E, 982-989 m, 1 dd; sta. CP2388, 9º26.9' N, 123º34.5' E, 762-786 m, 28 dd; sta. CP2389, 9º27.9' N, 123º38.4' E, 784-786 m, 7 dd; sta. CP2390, 9º27.4' N, 123º43.1' E, 627-645 m, 2 dd; Philippines: Philippine Sea. AURORA 2007, sta. CP2658, 15º59.45' N, 121º50.68' E, 422-431 m, 10 lv; sta. CP2659, 15º57.43' N, 121º50.26' E, 460-480 m, 1 lv; sta. CP2660, 15º52.06' N, 121º49.6' E, 506-542 m, 6 lv; sta. CP2668, 15º47.86' N, 121º45.65' E, 467-576 m, 1 lv; sta. CP2707, 15º04.79' N, 121º41.79' E, 368-442 m, 3 lv; sta. CP2725, 15º15.01' N, 121º36.18' E, 620-627 m, 2 lv; sta. CP2734, 15º56.41' N, 121º48.71' E, 453-460 m, 10 lv; sta. CP2735, 15º59.02' N, 121º50.25' E, 422-431 m, 7 lv; sta. CP2749, 15º56.38' N, 121º49.46' E, 473 m, 2 lv; Philippines: S and SE of Mindoro. MUSORSTOM 2, sta. CP36, 13º31' N, 121º24' E, 569-595 m, 1 dd; sta. CP44, 13º23' N, 122º20' E, 760-820 m, 1 dd; MUSORSTOM 3, sta. CP122, 12º20' N, 121º42' E, 673-675 m, 17 dd; sta. CP123, 12º10' N, 121º45' E, 673-675 m, 4 dd; sta. CP128, 11º50' N, 121º42' E, 673-675 m, 1 dd.

The understanding of this species has been most unfortunate, as it was repeatedly cited and illustrated in the literature but every time misidentified. However, based on the holotype (Fig. 1 D) and the material collected close to the type locality (e.g. Fig. 1 C), the species can be characterized and distinguished from other similar members of the genus. *B. atractoides* is a medium-sized species (the height of shell of the holotype is 35.5 mm, the largest shell is 53.1 mm high) with a rather slender species and high spire. Whorls are distinctly angulated well below periphery, with numerous (ca. 40-50 on the last whorl, depending on the shell size) rather weak peripheral gemmules. Subsutural slope is broad, steep and flattened, weakly or moderately concave. Spiral cords on it are strong and one (rarely two) of them on the subsutural fold is remarkably stronger, which is characteristic of the species. Columellar pleat is absent or very weak, at most as a light oblique swelling. Color off-white to yellowish-white, with rather narrow light orange-brown band on peripheral row of gemmules and lowermost subsutural ramp.

The northeastern population (Philippines Sea) differs in having shells with more concave subsutural ramp, sometimes comparatively higher spire and better developed plicae on subsutural fold (Fig. 1 E-F).

Distribution. Northeastern and central Philippines, living specimens found at 368-631 m, empty shells to 1260 m.

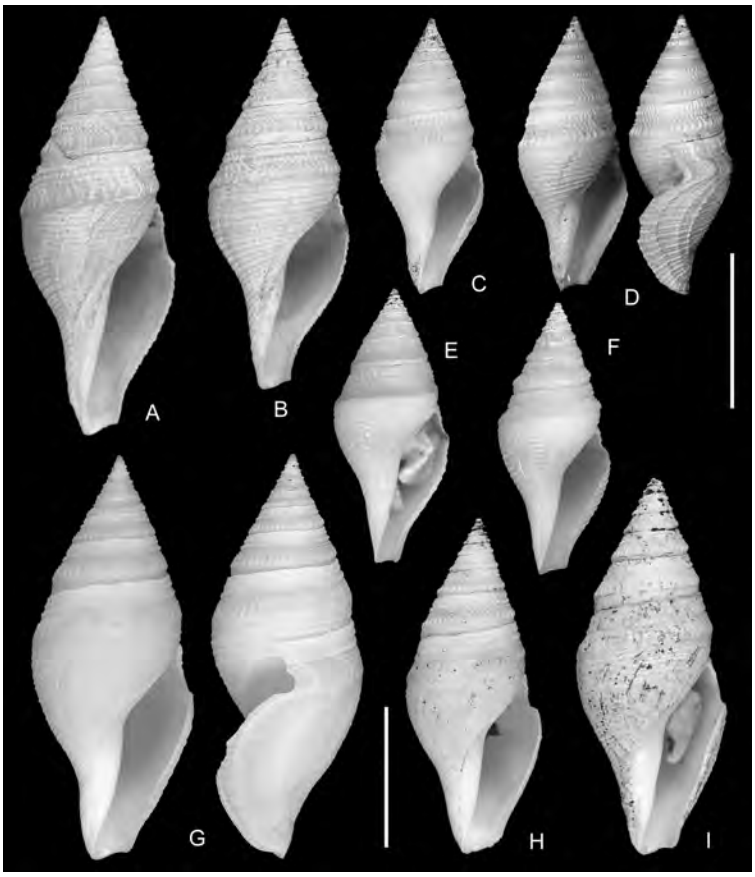


Fig. 1. A-F. *Bathytoma atractoides*. Scale bar 2 cm. – A. MUSORSTOM 3, sta. CP122, 53.0 mm. – B. MUSORSTOM 3, sta. CP122, 47.2 mm. – C. MUSORSTOM 3, sta. CP122, 35.6 mm. – D. Holotype, 35.5 mm. – E. AURORA 07, sta. CP2658, 35.3 mm. – F. AURORA 2007, sta. CP2660, 35.1 mm. – G-I. *Bathytoma paratractoides* sp. nov. Scale bar 2 cm. – G. Holotype. – H. Paratype, SALOMON 1, sta. CP1859, 47.4 mm. – I. Paratype, SALOMON BOA, sta. CP2768, 52.9 mm.

Bathytoma paratractoides sp. nov. (Fig. 1 G-I)

Type data. SALOMON 2, sta. CP2186, 8º17.0' S, 160º00.0' E, 487-541 m, 2 lv (holotype MNHN XXX and paratype MNHN XXX); SALOMON 1, sta. CP1859, 9º32.6' S, 160º37.3' E, 283-305 m, 1 lv (paratype MNHN XXX); SALOMON BOA, sta. CP2768, 9º21.1' S, 160º30.9' E, 447-503 m, 1 lv (paratype MNHN XXX).
Other material examined. Solomons Islands, 4 dd, 505-609 m.

Etymology. Para (Greek), meaning "near", and *atractoides*, the name of the closest species, referring to the similarity between them.

Description (holotype). Shell medium-sized for genus (less than 60 mm), solid, biconic, with rather low spire comprising 29% of SH. Last whorl large, inflated. Protoconch 1.1 mm in diameter, of about 1.5 light-brown smooth whorls developing thin arcuate folds in gradual transition to teleoconch. Teleoconch whorls 9.5, moderately convex, spire with flattened profile. Suture weakly and narrowly channelled due to elevated subsutural fold. Subsutural ramp broad, steep, nearly flat, with weak subsutural fold. Spiral cords rather narrow, variable in width and prominence but without clear separation into primary and secondary, interspaces narrower than cords themselves. Cords smoothed just abapically of whorl periphery, on shell base of about same width till canal. Two strong and broadly spaced cords on subsutural fold, followed abapically by four finely tuberculate ones on lower subsutural ramp, ca. 40 cords on last whorl abapically of periphery. Growth lines not very prominent but making spiral cords tuberculate on spire whorls. Peripheral gemmules weak, densely set, numerous, looking as thin, arcuate folds, becoming less distinct and irregular on last whorl, especially in its last half, and situated much

abapically of periphery, nearly on suture in earlier spire whorls. Shell base convex, well rounded, distinctly curved in passing to moderately long, twisted, broadly open and obliquely truncated canal. Aperture narrowly pyriform, anal sinus rather shallow, broad, U-shaped, with outer lip strongly projecting abapically of it. Columella bearing a narrow swelling in its middle part, better seen deep inside aperture. Color off-white. Dimensions: SH 56.3 mm, LW 39.8 mm, AH 32.0 mm, SD 22.5 mm.

Other shells differ slightly from the holotype in prominence of some sculptural elements, particularly the cords on subsutural ramp and fold and the peripheral gemmules. The subsutural ramp may be weakly concave. The columellar swelling may be stronger than in the holotype, or may be lacking.

Remarks. Although the molecular data support the distinction between them at species level, *B. paratractoides* is conchologically very similar to and even not easily distinguishable from *B. atractoides*. The main difference between them is the relative diameter/height ratio: in *B. paratractoides* this ratio is 0.395-0.509 (mean 0.434, n=8) at SH 22.8-56.3 mm; in *B. atractoides*, this ratio is in the range 0.354-0.453 (mean 0.388, n=20) at SH 28.1-53.1 mm. *B. atractoides* is so far known only from the Philippines, whereas *B. paratractoides* occurs in the Solomons.

Bathytoma stenos sp. nov. (Fig. 2 A-C)

Type data. Philippines, Bohol Sea, PANGLAO 2005, sta. CP2389, 9º27.9' N, 123º38.4' E, 784-786 m (holotype NMP XXX and 2 paratypes MNHN XXX).

Other material examined. Philippines, Bohol Sea, 3 lv in 583-762 m, 623 dd in 410-784 m; Philippines Sea, 19 lv in 431-506 m.

Etymology. Stenos (Greek) – narrow, referring to slender shell of the species.

Description (holotype). Shell small for genus (ca. 40 mm), strong, slender, with high spire comprising 35% of SH. Protoconch 1.05 mm in diameter, of about 1.5 smooth whorls developing few thin arcuate folds in transition to teleoconch. Teleoconch whorls 8, weakly convex, angulated abapically of periphery on early spire whorls but only subangulate on last whorls. Suture narrowly channelled by subsutural fold. Subsutural ramp broad, steep, concave in middle part, with moderately developed subsutural fold. Spiral cords strong, separated by interspaces not exceeding cord width, on shell base with one thin cord in most interspaces. Two thin cords on subsutural fold, 4 stronger and serrated cords on lower subsutural ramp, 3 thin cords running on peripheral row of gemmules, and ca. 35 primary cords on last whorl abapically of periphery. Growth lines strong, thickened as prosocline arcuate plicae on subsutural slope, cutting spiral cords and making them tuberculate. Peripheral gemmules weak, not numerous (ca. 30 on last whorl), looking as short, thin, rather broadly spaced folds. Shell base flatened, smoothly passing to moderately long, broadly open, obliquely truncated canal. Aperture subrectangular, anal sinus moderately deep, broadly U-shaped, with outer lip strongly projecting abapically of it. Columella bearing a weak oblique swelling in its middle part. Color yellowish-white.

Dimensions: SH 40.6 mm, LW 27.5 mm, AH 22.4 mm, SD 14.0 mm.

The species is rather variable in both shell proportions and details of sculpture. Some shells are slenderer and some broader than the holotype. The degree of development of the peripheral gemmules also slightly varies. Most shells have a comparatively very strong columellar pleat, some possessing even two pleats plus sometimes a third, weak pleat. Vast majority of the shells, including the live-collected specimens, do not show any shell coloration; two juvenile shells possess a light red-brown band on the periphery. The largest shell is 48.0 mm high.

Distribution. Philippines: Bohol and Philippine seas, at 382-786 m, with living molluscs found at 418-786 m.

Remarks. This small species is rather similar to *B. atractoides*, differing in narrower shell with more slowly expanding whorls, more concave subsutural slope without a strong spiral cord on subsutural fold. Both species occur sympatrically and have been collected together at 11 stations taken in the same cruises (out of 23 for *B. stenos* and out of 18 for *B. atractoides*).

Bathytoma tippetti Vera-Peláez, 2004 (Fig. 2 D-H)

Vera-Peláez, 2004: 120-122, pl. 3, fig. 5, pl. 4, figs. 1-20, pl. 5, figs. 1-4.

Type material. Holotype in Museo Nacional de Ciencias Naturales, Madrid, 15.05/46693.

Type locality. Philippines, Aliguay, 140 m.

Material examined.

Philippines: Bohol Sea. PANGLAO 2004, sta. P2, 9º39' N, 123º44' E, 400 m, 1 lv; PANGLAO 2005, sta. CP2332, 9º38.8' N, 123º45.9' E, 396-418 m, 3 lv, 1 dd; sta. CP2333, 9º38.2' N, 123º43.5' E, 584-596 m, 2 dd; sta. CP2340, 9º29.4' N, 123º44.4' E, 271-318 m, 2 lv, 1 dd; sta. CP2341, 9º24.5' N, 123º49.7' E, 544-712 m, 1 lv; sta. CP2348, 9º29.6' N, 123º52.5' E, 196-216 m, 1 lv; sta. CP2350, 9º31.4' N, 124º00.6' E, 602-738 m, 1 lv; sta.

CP2360, 8º48.9' N, 123º37.6' E, 357-372 m, 1 dd; sta. CP2372, 8º38.7' N, 123º16.0' E, 255-301 m, 1 lv, 1 dd; sta. CP2381, 8º43.3' N, 123º19.0' E, 259-280 m, 3 lv, 1 dd; sta. CP2383, 8º44.7' N, 123º18.5' E, 338-351 m, 2 lv; sta. CP2392, 9º29.0' N, 123º41.1' E, 242-400 m, 2 lv; sta. CP2395, 9º36.2' N, 123º43.8' E, 382-434 m, 8 lv, 1 dd; sta. CP2396, 9º36.3' N, 123º42.0' E, 609-673 m, 1 dd; sta. CP2404, 9º39.4' N, 123º43.3' E, 481-505 m, 3 dd; sta. CP2406, 9º40.6' N, 123º46.8' E, 334-387 m, 9 lv, 1 dd. Philippines: Sibuyan Sea. MUSORSTOM 2, sta. CP40, 13º08' N, 122º40' E, 280-440 m, 1 lv, 6 dd.

This medium-size species (SH up to 52.2 mm) is characterized by a rather narrow slender shell with high spire and distinctly angulated whorls. Subsutural ramp weakly concave, covered by strong uniform cords, subsutural fold usually weakly developed. Peripheral gemmules numerous, poorly pronounced, sometimes even subobsolete. Living shells either uniformly off-white, or with narrow, rather pale orange-brown band along peripheral row of gemmules. The species is similar to *B. stenos* but differs in broader shell with more gradate whorls.

Distribution. Central Philippines, at 196-738 m, with living molluscs found throughout this depth range.

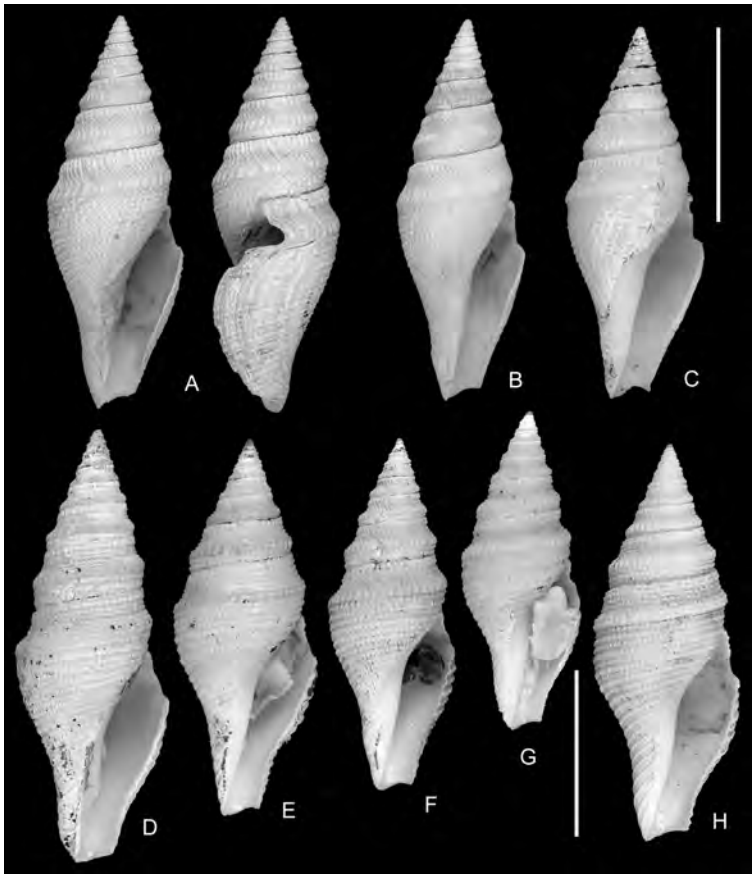


Fig. 2. A-C. *Bathytoma stenos* sp. nov. Scale bar 2 cm. – A. Holotype. – B. Paratype, PANGLAO 2005, sta. CP2389, 39.6 mm. – C. PANGLAO 2005, sta. CP2388, 38.5 mm. – D-H. *Bathytoma tippetti*. Scale bar 2 cm. – D. PANGLAO 2005, sta. CP2348, 50.4 mm. – E. PANGLAO 2005, sta. CP2383, 43.9 mm. –F. PANGLAO 2005, sta. CP2406, 41.6 mm. – G. PANGLAO 2005, sta. CP2395, 37.9 mm. – H. PANGLAO 2005, sta. CP2381, 46.6 mm.

Bathytoma episoma sp. nov. (Fig. 3 A-D, G-H)

Bathytoma atractoides – Vera-Peláez 2004: 109-112, pl. 3, figs. 6-9.
Type material. Holotype NMP XXX and 2 paratypes in MNHN.
Material examined.
Typical form. Philippines: Bohol Sea. PANGLAO 2005, sta. CP2333, 9º38.2' N, 123º43.5' E, 584-596 m, 2 lv; sta. CP2350, 9º31.4' N, 124º00.6' E, 602-738 m, 1 dd; sta. CP2388, 9º26.9' N, 123º34.5' E, 762-786 m, 2 lv (holotype and paratype), 6 dd; sta. CP2389, 9º27.9' N, 123º38.4' E, 784-786 m, 1 dd; sta. CP2390, 9º27.4' N, 123º43.1' E, 627-645 m, 1 dd; sta. CP2395, 9º36.2' N, 123º43.8' E, 382-434 m, 1 lv; sta. CP2396, 9º36.3' N, 123º42.0' E, 609-673 m, 25 dd; sta. CP2397, 9º34.9' N, 123º41.7' E, 642-669 m, 6 dd; sta. CP2399, 9º31.7' N, 123º41.9' E, 309-342 m, 1 dd sta. DW2401, 9º31.0' N, 123º40.4' E, 397-410 m, 1 dd; sta. CP2404, 9º39.4' N, 123º43.3' E, 481-505 m, 1 dd.

Philippines: Bohol Province, Panglao Island: Momo Beach, tangle net, "deep water", from fishermen, 1 dd; off Momo Beach, Looc, tangle nets, 200 fathoms (378 m), 1 lv (paratype).

"Broad form". Philippines: Philippine Sea. AURORA 2007, sta. CP2663, 15º57.18' N, 121º46.09' E, 562 m, 1 lv; sta. CP2699, 14º50.04' N, 123º34.50' E, 541-583 m, 1 lv; sta. CP2734, 15º56.41' N, 121º48.71' E, 453-460 m, 2 lv; sta. CP2735, 15º59.02' N, 121º50.25' E, 422-431 m, 1 lv; sta. CC2744, 15º58.75' N, 121º50.36' E, 418-456 m, 1 lv; sta. CP2750, 15º53.81' N, 121º51.55' E, 518-538 m, 1 lv.
Indonesia: Tanimbar Islands. KARUBAR, sta. CP 54, 8º21' S, 131º43' E, 836-869 m, 1 lv; sta. CP 72, 8º36' S, 131º33' E, 676-699 m, 1 lv; sta. CP 91, 8º44' S, 131º05' E, 884-891 m, 1 lv.
Indonesia: Makassar Stait. CORINDON, sta. CH209, 00º07' S, 117º53' E, 490 mm, 1 dd.

Etymology. Episomos (Greek) – bulky, fat, referring to large shell with well rounded whorls.

Description (typical form, holotype). Shell large for genus (nearly 80 mm), strong, rather slender, with moderately high spire comprising 1/3 of SH. Protoconch missing, initial spire whorls eroded. Remaining teleoconch whorls 8.5, moderately convex, with rounded profile. Suture narrowly channelled due to elevated subsutural fold. Subsutural ramp narrow, deeply concave in middle part, with prominent collar-like subsutural fold. Spiral cords strong and broad, separated by interspaces narrower than cords themselves, with one thin cord in each interspace. On shell base cords becoming progressively narrower and weaker towards canal. Two cords on subsutural fold, four increasing in strength on lower subsutural ramp abapically of excavation, ca. 30 cords on last whorl abapically of periphery. Growth lines strong, thickened, cutting spiral cords, especially on spire whorls. Peripheral gemmules weak, numerous (more than 60 on last whorl), looking as thin, arcuate, densely set folds, situated at about mid-whorl on later spire whorls. Shell base weakly convex, slightly curved in passing to moderately long and broadly open canal. Aperture narrowly pyriform, anal sinus deep and broad, U-shaped, with outer lip strongly projecting abapically of it. Columella bearing a swelling in its adapical part, better seen deep inside aperture. Color off-white, with aperture cream-orange inside.

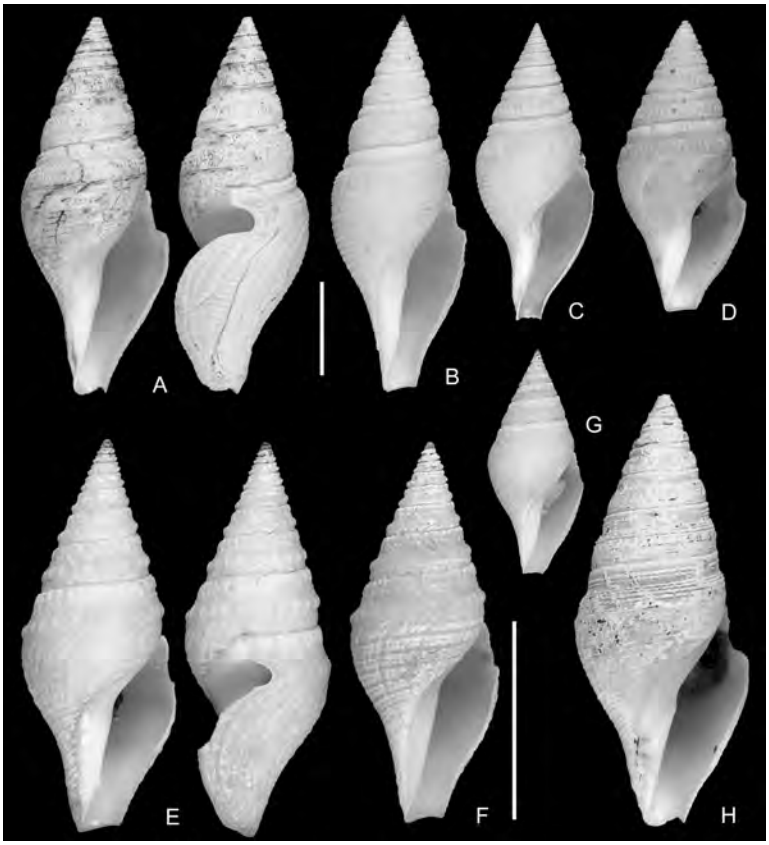


Fig. 3. A-D, G-H. *Bathytoma episoma* sp. nov. Scale bar 2 cm. – A. Holotype. – B. Paratype, Panglao Island, off Momo Beach, 79.0 mm. – C. PANGLAO 2005, sta. CP2388, 62.5 mm. – D. KARUBAR, sta. CP91, 60.9 mm. – G. AURORA 2007, sta. CP2699, 47.0 mm. -H. KARUBAR, sta. CP 72, 88,3 mm. E-F. *Bathytoma solomonensis* sp. nov. Scale bar 2 cm. – E. Holotype. – F. Paratype, SALOMON 2, sta. CP2288, 34.7 mm.

Dimensions: SH 78.1 mm, LW 52.9 mm, AH 44.7 mm, SD 29.1 mm.

Other shells show a rather low variability, concerning mostly the spiral sculpture, which can be almost completely cut into separate tubercles by growth lines, and the relative prominence of peripheral gemmules. The columellar swelling is always present; its prominence slightly varies, and sometimimes there is a weak pleat on external surface of the columella. The largest shell is 79.1 mm high.

Protoconch 1.0-1.1 mm in diameter, of about 1.5 smooth whorls developing thin arcuate folds in gradual transition to teleoconch.

Several shells collected off northeastern Luzon in the Philippine Sea, with none of them being fully adult, differ considerably in having far less concave subsutural slope and a weaker subsutural fold (Fig. 3 G). In fact, the shell shape and sculpture approach rather those of *B. carnicolor*, but that species possesses slightly weaker sculpture, less concave subsutural ramp and more slowly expanding whorls. At the same time, the molecular analysis shows that these shells belong to the very same clade as *B. episoma*. A conchological comparison showed that they are conspecific with several shells collected in Indonesia (Fig. 1 D, H) and represent a "broad form" of this species. Therefore, there are two conchologically distinct and constant allopatric forms within the species that differ in their non-overlapping distribution, and they do not show genetic distinction at the species level. It is possible that the typical form is an emerging species having not yet acquired genetic differences of a species level. The largest shell of the "broad form" is 88.3 mm high.

Distribution. The typical form of the species is so far known only from the Bohol Sea. The depth of occurrence is 309-786 m, with living molluscs found since 378 m and deeper. The shells of the typical form illustrated by Vera-Peláez (2004) also came from off Panglao in the Bohol Sea. The "broad form" was recorded in the Philippine Sea, off Tanimbar Islands an in the Makassar Stait in Indonesia. It occurs at 418-891 m, with living molluscs collected throughout this bathymetric range. Therefore, the range of the species is rather broad, from central and southeastern Indonesia to northern Philippines, with the typical form occurring in the inner sea of the Philippines at least partially separated from other known parts of the range by a land mass of Luzon Island or deep-sea basins of the Celebes and Banda seas.

Remarks. This is one of the largest species of the genus, with the shell height attaining 79.1 mm for the typical form (our data and Vera-Peláez, 2004) and 88.3 mm for the "broad form". The species is easily distinguishable by its rounded, not angulated whorls, deeply excavated subsutural ramp, distinct subsutural fold, and very numerous and weak peripheral gemmules.

Bathytoma netrion sp. nov. (Fig. 4 E-H)

Type material. Holotype and 3 paratypes in MNHN.

Material examined. Indonesia: Tanimbar Islands. KARUBAR, sta. CC56, 8º16' S, 131º59' E, 552-549 m, 2 dd; sta. CC57, 8º19' S, 131º53' E, 603-620 m, 1 dd; sta. CP59, 8º20' S, 132º11' E, 399-405 m, 3 lv, 5 dd; sta. CP69, 8º42' S, 131º53' E, 356-368 m, 1 lv, 9 dd; sta. CP70, 8º41' S, 131º47' E, 410-413 m, 2 dd; sta. CP72, 8º36' S, 131º33' E, 676-699 m, 2 lv (holotype and paratype), 2 dd; sta. CP77, 8º57' S, 131º27' E, 352-346, 1 dd. Philippines: Philippine Sea. AURORA 2007, sta. CP2725, 15º15.01' N, 121º36.18' E, 620-627 m, 1 lv; sta. CP2750, 15º53.81' N, 121º51.55' E, 518-538 m, 2 lv (paratypes).

Etymology. Netrion (Greek, diminutive of spindle), referring to small broadly fusiform shell.

Description (holotype). Shell small for genus (about 35 mm), comparatively thin-walled, broadly biconic, with rather low spire comprising 27% of SH. Protoconch missing, initial teleoconch eroded. Remaining whorls 8. Whorl angulation on spire whorls strongly shifted abapically, nearly to suture, so exposed whorl surface in fact represented only by very broad and weakly concave subsutural ramp, thus giving spire outline nearly straight sides. Suture narrowly channelled, subsutural fold rather weak but distinct and wide. Spiral cords moderately strong, rather narrow, on shell base thin and closely set just abapacally of periphery, then becoming wider and separated by interspaces exceeding cord width, with a thin secondary cord in some interspaces. One thin cord on middle of subsutural fold followed at some distance by 4 thicker, closely set, tuberculate cords. Nearly 30 cords on last whorl abapically of periphery, becoming much weaker on canal. Growth lines not particularly strong but making spiral cords on shell base slightly tuberculate and forming regular thin prosocline plicae on subsutural fold. Peripheral gemmules small but disinct, not numerous (38 on last whorl), slightly elongate and arcuate. Shell base moderately convex, nearly smoothly passing to moderately long, broadly open, obliquely truncated canal. Aperture narrowly pyriform, anal sinus rather deep, broadly U-shaped, with outer lip strongly projecting abapically of it. Columella with weak, narrow pleat better seen inside aperture. Color yellowish-white, without any colored zones.

Dimensions: SH 35.6 mm, LW 25.8 mm, AH 21.6 mm, SD 15.7 mm.

The species variability is rather high and expresses itself mainly in slightly varying relative shell width, the degree development of peripheral gemmules which can be even subobsolete (Fig. 4 F), and various development

of oblique plicae on subsutural fold, which can be much thinner and more numerous than in the holotype (Fig. 4 H). Shells from the Philippinean population (Fig. 4 G-H) are quite similar to the holotype from Indonesia. Protoconch is missing or eroded in all available specimens, in most intact shell it consists of about 1.5 smooth whorls. The largest shell is 40.6 mm high.

Distribution. Southeastern Indonesia (Tanimbar Islands) and central Philippines (Bohol Sea), at 352-699 m, living molluscs were found nearly throughout the depth range (356-699 m).

Remarks. This small species can be easily distinguised by its broad biconic shell with low-set whorl angulation and nearly straight profile of the spire.

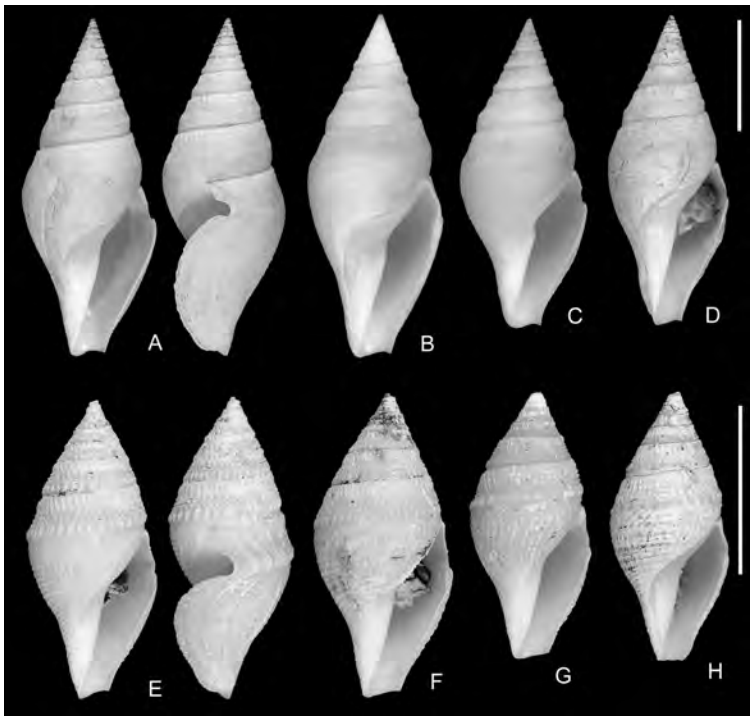


Fig. 4. A-D. *Bathytoma carnicolor* sp. nov. Scale bar 2 cm. – A. Holotype. – B. MUSORSTOM 7, sta. DW571, 61.3 mm. – C. Paratype, SALOMON 2, sta. CP2243, 56.0 mm. – D. SANTO 2006, sta. AT122, 56.0 mm. – E-H. *Bathytoma netrion* sp. nov. Scale bar 2 cm. – E. Holotype. – F. Paratype, KARUBAR, sta. CP72, 36.0 mm. – G. Paratype, AURORA 2007, sta. CP2750, 31.8 mm. – H. Paratype, AURORA 2007, sta. CP2750, 32.0 mm.

Bathytoma carnicolor sp. nov. (Fig. 4 A-D)

Type material. Holotype and 3 paratypes in MNHN.

Material examined. Solomon Islands. SALOMON 1, sta. CP1771, 8º17.1' S, 160º38.4' E, 411-498 m, 1 dd; sta. CP1793, 9º13.4' S, 160º07.8' E, 505-510 m, 2 dd; sta. CP1794, 9º16.1' S, 160º07.7' E, 494-504 m, 13 dd; sta. CP1805, 9º35.0' S, 160º42.7' E, 367-500 m, 1 dd; sta. CP1808, 9º45.5' S, 160º52.5' E, 611-636 m, 3 dd; SALOMON 2, sta. CP2194, 8º24.8' S, 159º26.7' E, 440-521 m, 3 lv (including holotype); sta. CP2219, 7º58.3' S, 157º34.4' E, 650-836 m, 2 lv (paratypes); sta. CP2227, 6º37.2' S, 156º12.7' E, 508-522 m, 5 dd; sta. CP2228, 6º34.7' S, 156º10.5' E, 609-625 m, 1 dd; sta. CP2243, 7º42.9' S, 156º27.3' E, 518-527 m, 1 lv (paratype); sta. CP2297, 9º08.8' S, 158º16.0' E, 728-777 m, 1 dd; SALOMON BOA, sta. DW2835, 10º40.9' S, 162º19.8' E, 735-862 m, 1 lv; sta. CP2850, 9º37.4' S, 160º47.2' E, 502-621 m, 2 lv.

Vanuatu. SANTO 2006, sta. AT122, 16º37.0' S, 167º59.5' E, 567-580 m, 2 lv.

Coral Sea. EBISCO, sta. DW2626, 21º06' S, 160º49' E, 728-739 m, 1 lv.

Wallis & Futuna: Waterwitch Bank. MUSORSTOM 7, sta. DW571, 12º31' S, 176º52' W, 502-508 m, 1 lv.

Etymology. Carnis (Latin, meaning flesh) + color, referring to the shell color.

Description (holotype). Shell rather large for genus (about 60 mm), biconic, strong, broad, with high spire comprising 34% of SH. Protoconch 1.15 mm in diameter, of about 1.5 smooth brown whorls developing few thin arcuate folds in rather rapid transition to teleoconch. Teleoconch whorls 10, spire whorls weakly convex, last whorl rather inflated. Early whorls angulate in abapical part, angulation shifting adapically and weakening towards nearly evenly rounded last whorl.

Suture weakly and narrowly channelled. Subsutural ramp broad, very steep, only weakly concave in middle, with poorly pronounced subsutural fold. Spiral sculpture rather weak: two broadly spaced thin cords on subsutural fold, five thin closely set tuberculate cords in middle of subsutural ramp, 2-3 smoothed cords abapically of periphery on spire whorls. Last whorl abapically of periphery covered with low, flattened cords, alternating in width and separated by narrow interspaces. Peripheral gemmules rather prominent only on early spire whorls, becoming obsolete on last whorls, looking as arcuate, closely set folds just abapically of whorl periphery on later spire whorls. Growth lines moderately strong, forming numerous thin, prosocline, oblique plicae on subsutural fold. Shell base convex, slightly concave in passing to moderately long, straight, broadly open, obliquely truncated canal. Aperture narrowly oval, anal sinus relatively shallow, narrowly U-shaped, with outer lip projecting abapically of it. Columella it its middle part bearing a weak oblique pleat better seen deep inside aperture. Color light-corneous.

Dimensions: SH 60.6 mm, LW 39.8 mm, AH 32.6 mm, SD 23.4 mm.

The species shows some variability in conchological characters. The whorl angulation and the gemmules may be more pronounced, even on last whorl (Fig. 4 C). The shells from Vanuatu have more rough plicae on subsutural ramp, whereas the sculpture of the shell from Waterwitch Bank is more smoothed (Fig. 4 B). One shell has a rather strong columellar pleat, whereas in two adult specimens from Vanuatu the pleat is either obsolete or absent. Juvenile shells lack any pleat or a swelling on columella. Juvenile, dead-collected but fresh shells from SALOMON 2 sta. CP2227 show a weak but distinct orange-brown color band running along periphery. The largest shell is 64.3 mm high.

Molecular data indicate that there are certain genetic differences between populations of the species from the Solomons and Vanuatu. However, conchological characters do not allow to distinguish two species in this case, and we prefer to consider these populations as a single species.

Distribution. The species is known with confidence from the Solomon Islands, Vanuatu, and Samoan Archipelago. The only available specimen from the Coral Sea is an early juvenile and conchologically cannot be positively assigned to this species, even though the molecular data indicate its belonging to *B. carnicolor*. The depth of occurrence is 367-862 m, with living molluscs found at 440-836 m.

Remarks. This small species is characterized by its strong sculpture with well-defined peripheral gemmules and rough growth lines forming prominent oblique plicae on subsutural fold.

Bathytoma badifasciata sp. nov. (Fig. 5 A-C)

Type material. Holotype and 2 paratypes in MNHN.

Material examined. Solomon Islands. SALOMON 1, sta. CP1786, 9º21.3' S, 160º24.6' E, 387 m, 1 dd; sta. CP1800, 9º21.4' S, 160º23.9' E, 357-359 m, 1 dd; SALOMON 2, sta. CP2194, 8º24.8' S, 159º26.7' E, 440-521 m, 1 lv; sta. CP2212, 7º37.8' S, 157º41.7' E, 400-475 m, 1 lv (paratype); sta. CP2290, 8º40.0' S, 157º31.7' E, 384-418 m, 1 dd; SALOMON BOA, sta. CP2857, 9º43.6' S, 160º48.5' E, 473-505 m, 2 lv (holotype and paratype).

Etymology. *Badius* (Latin, reddish-brown) + *fasciatus* (Latin, with bands), referring to the reddish-brown color band on the shell periphery.

Description (holotype). Shell of medium size for genus (slightly exceeding 60 mm), narrowly biconic, strong, comparatively slender, with high spire comprising 33% of SH. Protoconch 1.1 mm in diameter, of about 1.5 smooth light-brown whorls developing several thin arcuate folds in transition to teleoconch. Teleoconch whorls 9.5, spire whorls weakly convex, last whorl moderately so. Early whorls angulate in abapical part just near suture, angulation shifting adapically almost to periphery and weakening towards nearly evenly rounded last whorl.

Suture weakly and narrowly channelled. Subsutural ramp broad, very steep, only weakly and evenly concave, with very poorly developed subsutural fold. Spiral sculpture: 6-7 thin but distinct, broadly spaced serrated cords on subsutural ramp, 2 cords running along peripheral row of gemmules, 1-4 cords abapically of periphery on spire whorls. Last whorl abapically of periphery covered with low, flattened cords, alternating in width but not clearly separated into primary and secondary ones. Interspaces much narrower than cords themselves. Peripheral gemmules on all later whorls subobsolete, forming a narrow row of arcuate, closely set folds. Growth lines comparatively weak, not prominent but rendering spiral cords serrated, not forming any plicae on subsutural ramp. Shell base flattened, smoothly passing to moderately long, straight, broadly open canal. Aperture narrowly oval, anal sinus, judging from growth lines, relatively shallow, broadly U-shaped, with outer lip projecting abapically of it. Columella it its middle part bearing a distinct narrow oblique pleat better seen inside aperture. Color light-chestnut, with a broad reddish-brown band encircling whorl periphery and most intense on row of peripheral gemmules; abapical two spiral cords on subsutural slope also stained with same color.

Dimensions: SH 61.0 mm, LW 40.9 mm, AH 32.5 mm, SD 21.9 mm.

In studied material, the species is rather constant in conchological characters. The columellar pleat is not developed or weak in juvenile or subadult specimens. The largest shell is 58.6 mm high.

Distribution. Solomon Islands, at 377-521 m, with living molluscs found at 400-521 m.
Remarks. The species is characterized by very narrow row of weak peripheral gemmules, uniform spiral cords on subsutural ramp that lacks oblique plicae on the obsolete subsutural fold.

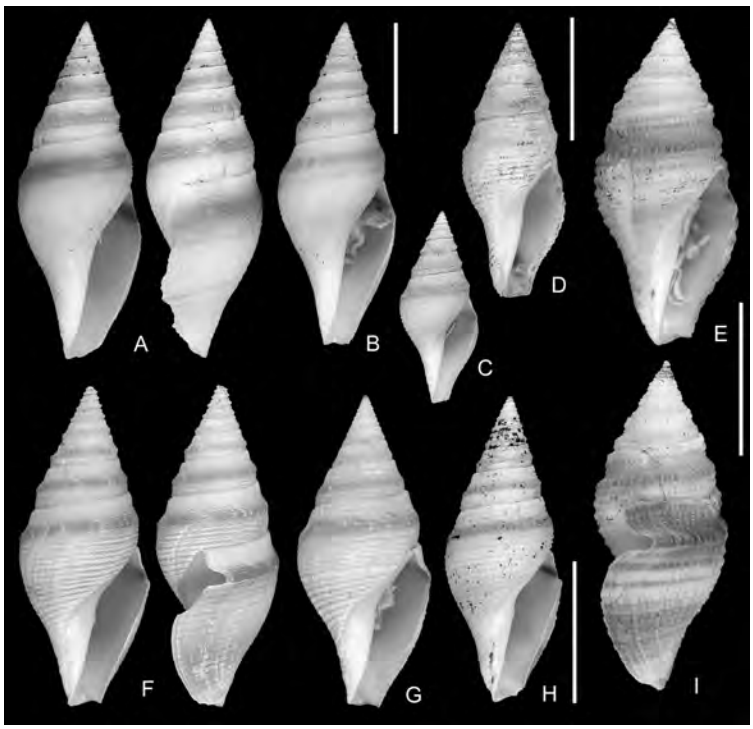


Fig. 5. A-C. *Bathytoma badifasciata* sp. nov. Scale bar 2 cm. – A. Holotype. – B. Paratype, SALOMON 2, sta. CP2212, 58.6 mm. – C. SALOMON 1, sta. CP1786, 35.6 mm. – D. *Bathytoma* sp., AURORA 2007, sta. CP2668, 44.7 mm. – E, I. *Bathytoma punicea* sp. nov., holotype. Scale bar 2 cm. – F-H. *Bathytoma consors* sp. nov. Scale bar 2 cm. – F. Holotype. – G. Paratype, SALOMON BOA, sta. CP2839, 44.2 mm. – H. SALOMON BOA, sta. CP2848, 43.8 mm.

Bathytoma consors sp. nov. (Fig. 5 F-H)

Type material. Holotype and 5 paratypes in MNHN.
Material examined. Solomon Islands. SALOMON 1, sta. CP1836, 10º10.3' S, 161º21.7' E, 439-486 m, 1 dd (paratype); SALOMON 2, sta. CP2217, 6º37.2' S, 156º12.7' E, 508-522 m, 1 lv; SALOMON BOA, sta. CP2767, 9º19.0' S, 160º05.9' E, 416-425 m, 1 lv (paratype); sta. CP2833, 10º42.4' S, 162º18.8' E, 565-585 m, 1 lv; sta. CP2834, 10º44.0' S, 162º20.7' E, 545-753 m, 1 lv; sta. CP2838, 10º25.0' S, 161º19.6' E, 510-581 m, 2 lv (holotype and paratype); sta. CP2839, 10º25.6' S, 161º20.4' E, 506-567 m, 1 lv (paratype); sta. CP2848, 9º34.9' S, 160º47.1' E, 414-456 m, 2 lv; sta. CP2857, 9º43.6' S, 160º48.5' E, 473-505 m, 1 lv (paratype).
Etymology. *Consors* (Latin, meaning brother, partner), referring to similarity to the preceding species.
Description (holotype). Shell of medium size for genus (less than 50 mm), biconic, strong, comparatively stout, with rather high spire comprising 31% of SH. Protoconch missing, remaining teleoconch whorls 9 in number. Spire whorls moderately convex, angulate at about periphery or just abapically of it. Suture weakly and narrowly channelled. Subsutural ramp moderately broad, rather steep, distinctly concave in its middle, with prominent but not very strong subsutural fold. Spiral sculpture well pronounced: 4 rather weak, low, smoothened, closely set cords on subsutural fold, 4 stronger and slightly serrated cords on subsutural ramp, 2 indistinct cords running along peripheral row of gemmules, 3 cords abapically of periphery on last spire whorls. Last whorl abapically of periphery covered with rather strong, rounded, nearly uniform cords, rarely bearing secondary cord in interspaces between some of them. Interspaces increasing in width abapically, from narrower than cords to slightly wider. Peripheral gemmules distinct, widely spaced, strongly curved, becoming irregular on last whorl. Growth lines comparatively weak but rendering spiral cords serrated and forming dense oblique plicae on subsutural ramp. Shell base convex, smoothly passing to moderately long, straight, broadly open canal. Aperture subrectangular-oval, anal sinus not particularly deep, narrowly U-shaped, with outer lip projecting abapically of it. Columella it its middle part bearing a distinct narrow oblique pleat better seen inside

aperture. Color light-chestnut, with a rather narrow reddish-brown band encircling whorl periphery along row of peripheral gemmules.

Dimensions: SH 45.9 mm, LW 31.6 mm, AH 24.7 mm, SD 18.5 mm.

Shells of the species may somewhat vary in the shell proportions and details of sculpture. The whorl angulation varies in prominence and its position may be more abapical. The plicae on subsutural folds also slightly vary in strength and may be more distant from each other. Protoconch 1.10-1.15 mm in diameter, of about 1.5 smooth light-brown whorls developing several thin arcuate folds in transition to teleoconch. The largest shell is 48.6 mm high.

Distribution. Solomon Islands, at 414-753 m, with living molluscs found throughout this depth range.

Remarks. The species is rather similar to *B. badifasciata* but differs in having a somewhat smaller and broader shell, with distinct subsutural fold covered by oblique axial plicae which form a kind of characteristic lattice when crossing the spiral cords, and a narrower color band. The ranges of these two species formally completely overlap, but they were collected together at only one station out of 14 (see material examined).

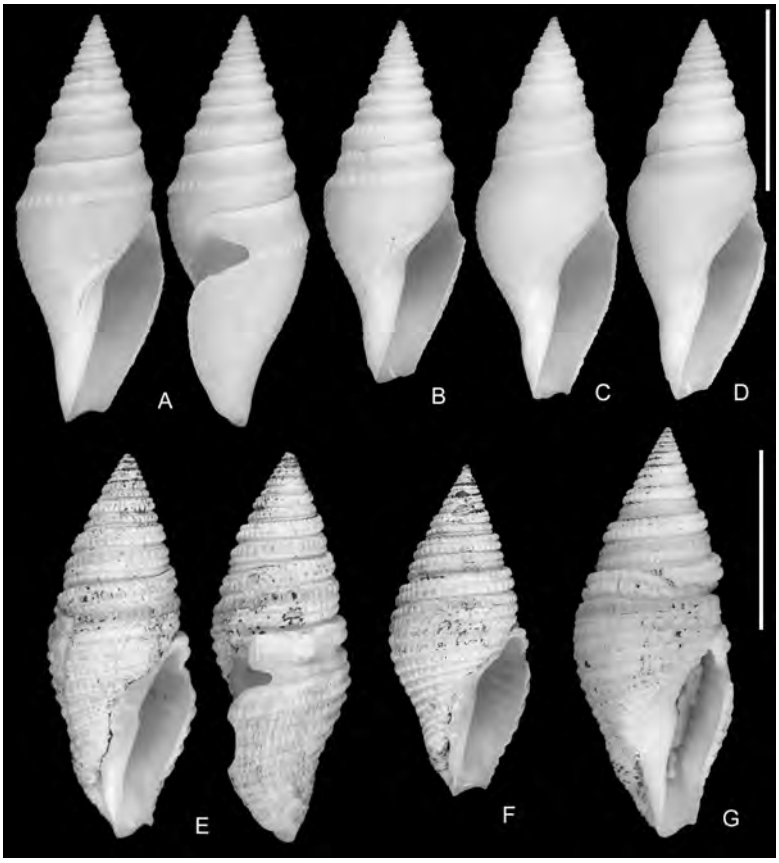


Fig. 6. A-D. *Bathytoma neocaledonica* sp. nov. Scale bar 2 cm. – A. Holotype. – B. Paratype, HALIPRO 1, sta. CP850, 40.3 mm. – C. Paratype, EBISCO, sta. CP2551, 44.4 mm. – D. SMIB 2, sta. DW23, 44.1 mm. – E-G. *Bathytoma cranaos* sp. nov. Scale bar 2 cm. – E. Holotype. – F. Paratype, SALOMON 1, sta. CP1801, 37.2 mm. – G. Paratype, SALOMON BOA, sta. DW2829, 45.0 mm.

Bathytoma cranaos sp. nov. (Fig. 6 E-G)

Type material. Holotype and 2 paratypes in MNHN.

Material examined. Solomon Islands. SALOMON 1, sta. CP1801, 9º25.0' S, 160º25.9' E, 264-273 m, 2 lv (holotype and paratype); sta. CP1837, 10º12.8' S, 161º28.6' E, 381-383 m, 1 dd; SALOMON BOA, sta. DW2829, 10º46.4' S, 162º19.6' E, 220-292 m, 1 lv (paratype).

Etymology. Kranaos (Greek) – rugged, rocky, referring to the rough sculpture of the species.

Description (holotype). Shell rather small for genus (less than 45 mm), thick and strong, slender, with moderately high spire comprising 30% of SH. Protoconch missing. Remaining teleoconch whorls ca. 6.5, with flattened profile. Suture broadly channelled due to massive subsutural fold. Subsutural ramp narrow, steep, deeply concave in lower part, with very strong collar-like subsutural fold. Spiral

sculpture of very strong and broadly spaced cords. Interspaces wider than cords themselves, on last whorl with one thinner cord in each interspace. Two strong cords on subsutural fold, subsutural ramp with only one weak cord just adapically of peripheral gemmules; 12 cords on last whorl abapically of periphery. On shell base cords becoming progressively narrower and weaker towards canal. Growth lines rough, thickened, cutting spiral cords into tubercles and giving them a beaded appearance. Peripheral gemmules numerous (ca. 50 on last whorl), weak, being not stronger than tubercles on subsutural fold, looking as arcuate, densely set folds, situated abapically of whorl periphery on later spire whorls. Shell base weakly convex, slightly concave in passing to fasciole obliquely encircling moderately long, curved and broadly open canal. Aperture narrowly oval, anal sinus deep, narrowly U-shaped, with outer lip projecting abapically of it. Columella bearing an oblique pleat in its middle, thick callus on its abapical part somewhat detached, forming a false umbilicus. Color off-white, with aperture pale cream-orange inside, and row of peripheral tubercled slightly stained with red-brown.

Dimensions: SH 42.9 mm, LW 29.8 mm, AH 24.2 mm, SD 16.0 mm.

One paratype (SALOMON BOA sta. DW2829, 45.0 mm high) is quite similar to the holotype in all essential characters, whereas the other (type locality) differs in having a proportionally stouter shell with nearly straight canal (Fig. 6 F). The elongate shape of the two shells may be due to strong damage to the shell that occurred on the last whorl and possibly resulted in a distorted growth, while the latter shell was severily broken only on a spire whorl and the last whorl growth proceeded in a normal way.

Protoconch 1.0 mm in diameter, of about 1.5 smooth whorls developing few thin arcuate folds in rather rapid transition to teleoconch.

Distribution. Solomon Islands, at 220-383 m, with living molluscs found at 220-292 m.

Remarks. The species is readily distinguishable by its very strong, rough spiral sculpture. For differences from *B. punicea*, see that species.

All available shells of the species show heavy scars of the shell breaks and repairing, indicating a high predatory pressure in that particular habitat.

Bathytoma punicea sp. nov. (Fig. 5 E, I)

Type material. Holotype in MNHN.

Material examined. Solomon Islands. SALOMON 1, sta. CP1800, 9º21.4' S, 160º23.9' E, 357-359 m, 4 dd; sta. CP1801, 9º25.0' S, 160º25.9' E, 264-273 m, 4 dd; sta. CP1804, 9º32.0' S, 160º37.4' E, 309-328 m, 4 dd; sta. CP1837, 10º12.8' S, 161º28.6' E, 381-383 m, 5 dd; sta. CP 1860, 9º22' S, 160º31'E, 620 m, 1 dd; SALOMON BOA, sta. CP2828, 10º25.6' S, 161º57.9' E, 173-379 m, 1 lv (holotype).

Etymology. Puniceus (Latin) – purplish-red, referring to coloration of the shell peripheral zone.

Description (holotype). Shell rather small for genus (less than 45 mm), thick and strong, slender, with moderately high spire comprising 27% of SH. Protoconch and initial teleoconch whorls eroded, so protoconch/teleoconch border cannot be traced. Total number of whorls 9.5, angulated just near suture on earlier whorls and abapically of periphery on penultimate whorl. Suture narrowly channelled due to massive subsutural fold. Subsutural ramp wide, steep, deeply concave in middle, with strong collar-like subsutural fold. Spiral sculpture of strong, broadly spaced cords on last whorl, with interspaces wider than cords themselves, often with one thinner but varying in strength cord in interspaces. Two strong cords on subsutural fold followed by 3-4 thinner cords on rest of subsutural ramp; 19 cords on last whorl abapically of periphery. On shell base cords becoming progressively narrower and lower towards canal. Growth lines well expressed, numerous, thin, making spiral cords tuberculate. Peripheral gemmules rather numerous (less than 50 on last whorl), not prominent, looking as arcuate, densely set folds. Shell base convex, slightly concave in passing to straight, broadly open and obliquely truncated canal. Aperture narrowly oval, anal sinus deep, narrowly U-shaped, with outer lip projecting abapically of it. Columella bearing an oblique pleat in its middle. Color better expressed on last whorl: brownish-red, with lighter band on shell base and darker peripheral row of gemmules, spiral cords stained stronger; color faded on earlier whorls of spire.

Dimensions: SH 42.9 mm, LW 31.5 mm, AH 25.7 mm, SD 18.6 mm.

All other available specimens are juvenile shells, similar in all essential characters to the holotype but slightly varying in the degree of subsutural ramp concavity and often having more heavily tuberculated cords on subsutural fold. Protoconch 1.0-1.05 mm in diameter, of about 1.5 smooth whorls developing 2-3 thin arcuate folds in rather rapid transition to teleoconch.

Distribution. Solomon Islands, at 173-620 m, the only living mollusc was found at 173-379 m.

Remarks. This species is very similar to *B. cranaos*, differing in broader shell with lower spire, much weaker development of subsutural fold with weaker cords on it, broader subsutural ramp, more abapical position of peripheral gemmules, and better expressed shell coloration.

Bathytoma solomonensis sp. nov. (Fig. 3 E-F)

Type material. Holotype and 5 paratypes in MNHN.
Material examined. Solomon Islands. SALOMON 1, sta. CP1794, 9º16.1' S, 160º07.7' E, 494-504 m, 1 dd; sta. CP1808, 9º45.5' S, 160º52.5' E, 611-636 m, 2 dd; sta. CP1859, 9º32.6' S, 160º37.3' E, 283-305 m, 2 dd; SALOMON 2, sta. CP2184, 8º16.9' S, 159º59.7' E, 464-523 m, 1 lv; sta. CP2194, 8º24.8' S, 159º26.7' E, 440-521 m, 1 lv (holotype); sta. CP2206, 7º43.4' S, 158º29.0' E, 391-623 m, 1 dd; sta. CP2213, 7º38.7' S, 157º42.9' E, 495-650 m, 1 lv (paratype), 5 dd; sta. CP2214, 7º41.6' S, 157º43.8' E, 550-682 m, 2 lv (one paratype); sta. CP2228, 6º34.7' S, 156º10.5' E, 609-625 m, 1 dd; sta. CP2245, 7º43.1' S, 156º26.0' E, 582-609 m, 1 lv (paratype); sta. CP2263, 7º54.8' S, 156º51.3' E, 485-520 m, 2 lv; sta. CP2268, 7º48.7' S, 156º53.3' E, 632-640 m, 1 lv (paratype); sta. CP2284, 8º38.4' S, 157º21.5' E, 195-197 m, 1 lv (paratype); CP2288, 8º36.3' S, 157º26.5' E, 509-520 m, 1 lv (paratype), 2 dd; SALOMON BOA, sta. DW2833, 10º42.4' S, 162º18.8' E, 565-585 m, 1 lv.

Etymology. The species is named after the region of collection, the Solomon Islands.

Description (holotype). Shell rather small for genus (less than 40 mm), biconic, strong, slender, with high spire comprising 36% of SH. Protoconch 1.15 mm in diameter, of about 1.5 smooth whorls developing few thin arcuate folds in rather rapid transition to teleoconch. Teleoconch whorls 8.5, weakly convex, angulate abapically of periphery. Suture weakly and narrowly channelled. Subsutural ramp broad, steep, moderately concave in middle, with prominent subsutural fold. Spiral sculpture rather smoothed on spire whorls and last whorl periphery: two rather indistinct cords on subsutural fold, two narrow cords in middle of subsutural ramp, 2-3 weak cords abapically of periphery on spire whorls. Shell base covered with more distinct broad cords, with interspaces about equal to cords in width, sometimes with one thinner cordlet in interspace. Peripheral gemmules prominent, widely spaced and not numerous (26 on last whorl and 23 on penultimate), rather strong, being not stronger than tubercles on subsutural fold, looking as arcuate, densely set folds, situated abapically of whorl periphery on later spire whorls. Growth lines rough, thickened, forming characteristic strong, prosocline, oblique plicae on subsutural fold. Number of plicae per whorl slightly exceeding that of peripheral gemmules. Shell base rather convex, slightly concave in passing to moderately long, straight and broadly open canal. Aperture narrowly oval, anal sinus relatively shallow, broadly U-shaped, with outer lip projecting abapically of it. Columella it its adapical part bearing an oblique pleat looking as low and broad swelling better seen deep inside aperture. Color uniformly off-white, without any color markings.

Dimensions: SH 36.1 mm, LW 23.2 mm, AH 19.5 mm, SD 14.1 mm.

Other specimens examined only slightly vary in minor details of sculpture and in the number of plicae on subsutural fold and peripheral gemmules (up to 32 on last whorl). The columella may lack any pleat or a distinct swelling. One shell has peripheral row of gemmules slightly darker than background shell color. The holotype is the largest specimen.

Distribution. Solomon Islands, at 195-650 m, with living molluscs found throughout this depth range.

Remarks. This small species is characterized by its strong sculpture with well-defined peripheral gemmules and rough growth lines forming prominent oblique plicae on subsutural fold.

Bathytoma neocaledonica sp. nov. (Fig. 6 A-D)

Type data. New Caledonia. BATHUS 1, sta. CP708, 21º43' S, 166º39' E, 550-580 m, 3 lv (paratypes MNHN XXX), 1 dd (holotype MNHN XXX); HALIPRO 1, sta. CP850, 21º43' S, 166º39' E, 541-580 m, 3 dd (paratypes MNHN XXX); EBISCO, sta. CP2551, 21º06' S, 158º35' E, 637-650 m, 1 lv (paratype MNHN 17857).
Other material examined. New Caledonia, 6 lv in 580-650 m, 65 dd in 300-720 m.

Etymology. The species is named after New Caledonia.

Description (holotype). Shell of medium size for genus (less than 40 mm), strong, slender, with high spire comprising 33% of SH. Protoconch 1.0 mm in diameter, of about 1.5 light-brown smooth whorls developing few thin arcuate folds in transition to teleoconch. Teleoconch whorls 8.8, weakly convex, angulate abapically of periphery. Suture weakly and narrowly channelled. Subsutural ramp broad, steep, rather concave, with moderately developed subsutural fold. Spiral cords rather smoothed, low, broadly spaced: 2-3 thin and weak cords on subsutural fold, 5 thin cords on rest of subsutural ramp, 2 indistict cords covering peripheral row of gemmules, 2-4 weak cords abapically of periphery on spire whorls, and more than 30 cords on shell base and canal. Cords on shell base alternating in width and becoming narower and more uniform toward canal. Peripheral gemmules small but prominent, widely spaced, not numerous (33 on last whorl and 36 on penultimate), rather clearly cut, as short arcuate folds. Growth lines not very prominent but making spiral cords tuberculate and forming numerous, weak, prosocline, oblique plicae on subsutural fold. Shell base rather convex, slightly concave in smooth passing to moderately long, broadly open, obliquely truncated canal. Aperture narrowly oval, anal sinus rather

deep, broadly U-shaped, with outer lip projecting abapically of it. Columella nearly smooth, with only faint swelling in its middle. Color uniformly off-white, without any color markings.

Dimensions: Holotype SH 45.4 mm, LW 30.4 mm, AH 25.5 mm, SD 16.5 mm. The largest shell is 54.9 mm high.

Distribution. Around New Caledonia and in adjacent part of the Coral Sea, alive in 541-800 m, empty shells in 250-950 m.

Remarks. *B. neocaledonica* is represented by two intergrading forms: "smooth" and typical "gemmate". The former is characterized by more rounded whorl periphery and subobsolete peripheral gemmules (Fig. 6 C-D). These forms differ in their bathymetric, but not in their geographic, distribution: the "smooth" form was found in 250-570 m and the "gemmate" form in 515-950 m. *B. neocaledonica* resembles *B. tippetti* but differs in having a uniformly colored and more stout shell with weak spiral sculpture.

Bathytoma sp. (Fig. 5 D)

Material examined. Philippines: Philippine Sea. AURORA 2007, sta. CP2668, 15º47.86' N, 121º45.65' E, 467-576 m, 1 lv.

This single specimen is somewhat similar to *B. paratractoides* from the Solomons but differs in its more slender shell with very weak subsutural fold, and uniform, thin, widely spaced cords on subsutural ramp. This species is well differentiated genetically as well. However, we refrain from formal description because we have only one shell in not very good condition.

3. Morphométrie géométrique : le genre Benthomangelia

La même approche moléculaire a été adoptée pour le genre *Benthomangelia* (Article 7). Les gènes COI et 28S ont été séquencés et permettent tout deux de définir cinq entités. Une analyse similaire à celle qui a été faite pour le genre *Bathytoma* a également été réalisée, mais n'a pas permis de mettre en évidence des différences entre les hypothèses d'espèces proposées avec les gènes COI et 28S. La variabilité de la coquille a alors été analysée par une approche de morphométrie géométrique, en collaboration avec Michel Baylac. Le contour du dernier tour de la spire de chacun des spécimens analysés moléculairement a été modélisé par la méthode des transformés de Fourier (EFA – Elliptic Fourier Analysis – Rohlf 1996). Une analyse discriminante a permis de mettre en évidence des différences morphologiques entre les cinq entités génétiques.

Parmi les noms d'espèces disponibles dans la littérature, *Benthomangelia trophonoidea* était le seul dont la description d'espèce qui y était associée correspondait à l'une des cinq espèces délimitées. Ce nom *B. trophonoidea* est donc maintenant associé à cette espèce. Les quatre autres correspondent à des espèces nouvelles.

Cependant, au sein de l'une des espèces, trois populations correspondant aux trois régions géographiques représentées dans cette espèce (Philippes, Iles Salomon et Vanuatu) présentent une structuration génétique importante détectée avec le gène COI (le gène 28S n'est pas variable au sein de ce groupe). Des différences morphologiques ont également été détectées entre ces trois populations. Le statut de ces trois populations (trois espèces différentes ou trois populations au sein d'une seule espèce) reste à évaluer, en utilisant par exemple un gène nucléaire plus variable que le gène 28S. Si ces trois populations appartiennent à trois espèces différentes, le gène nucléaire permettra comme le gène COI de mettre en évidence une structuration génétique entre elles.

L'approche de morphométrie géométrique a ainsi permis de mettre en évidence des différences morphologiques entre des espèces qui n'avaient pas été différenciées par les méthodes traditionnelles (espèces pseudo-cryptiques – Knowlton 2000). Contrairement à l'approche morphologique traditionnelle qui a été appliquée pour le genre *Bathytoma*, cette approche morphométrique a également l'avantage d'être reproductible : les caractères analysés ont été clairement définis et peuvent être réutilisés à l'identique par un autre taxonomiste. Elle permettra ainsi facilement d'ajouter d'autres spécimens, pour notamment améliorer la représentativité géographique et bathymétrique de l'échantillonnage. A ce titre, les spécimens qui sont conservés séchés dans les collections du MNHN et qui n'ont pas pu être analysés moléculairement pourront également être inclus dans l'analyse. Cela permettrait notamment de tester si les différences morphologiques qui ont été identifiées entre les cinq espèces sont maintenues si une plus grande diversité est inclus dans l'analyse, ou encore de tester l'influence de facteurs environnementaux, comme la profondeur, sur la variabilité morphologique.

An integrative approach to species delimitation in *Benthomangelia* (Mollusca: Conoidea)

N. Puillandre[1,*], M. Baylac[2], M.-C. Boisselier[1], C. Cruaud[3], S. Samadi[1]

[1] Université Pierre et Marie Curie (UPMC), UMR 7138, Systématique, adaptation, évolution (UPMC/IRD/MNHN/CNRS), Paris, France.
[2] Muséum national d'Histoire naturelle (MNHN), USM 0601 Origine, structure et évolution de la biodiversité, Paris, France
[3] GENOSCOPE, Centre National de Séquençage, Evry, France

* Corresponding author: UMR 7138, Systématique, adaptation, évolution, UPMC; IRD; MNHN; CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.; puillandre@mnhn.fr; +33 1 40 79 37 43

Biological Journal of the Linnean Society, in press

Running title: Integrative taxonomy in *Benthomangelia*

Abstract

DNA sequences are currently used to propose primary hypotheses of species delimitation, especially when morphological variability is difficult to assess. In an integrative taxonomy framework, these hypotheses are then compared with other characters, such as morphology or geography, to produce robust species delimitations. For this purpose, the COI gene has been sequenced for almost 50 specimens of the genus *Benthomangelia*, a deep-sea marine gastropod genus, collected in the South-West Pacific. Five genetic groups, displaying low and high genetic distances respectively within and between groups, were defined. COI hypotheses were compared to both the results obtained with the independent nuclear 28S gene and to an Elliptic Fourier Analysis of the shape of the last whorl of the shell. 28S gene analysis confirmed the same well-supported groups as COI, and Elliptic Fourier Analysis identified several morphological characters that vary similarly to genetic variability.

Keywords
28S rRNA - COI gene – *Benthomangelia* – DNA taxonomy – Elliptic Fourier Analysis - Integrative taxonomy – Molluscs.

Introduction

Given the sizeable number of species that remains to be described, especially in the context of a rapidly increasing rate of species extinctions, we urgently need to rethink the work of alpha-taxonomy to improve the rate of species description. In marine environments, the amount of unknown diversity is important, especially for molluscs (Bouchet *et al*., 2002). Among marine molluscs, the Conoidea are one of the most diverse taxa (Bouchet, 1990; Taylor, Kantor & Sysoev, 1993), and it includes a significant portion of undescribed species. Morevover, most of the described taxa from shallow water and deeper ecosystems remain largely underexplored. The difficulty of sampling in such environment and the lack of specialists for most of these groups explain this taxonomic impediment (Boyle *et al*., 2004;).

Until recently most species descriptions of shelled molluscs were based exclusively on shell characters, leading to brief diagnoses that in some cases apply to more than one taxon. Furthermore, shell variability is difficult to characterize with discrete characters (Pfenninger, Cordellier & Streit, 2006), and consequently the analysis of such characters, as done traditionally, is difficult to reproduce. Finally, the broad plasticity of shell characters for some well known species was demonstrated (Hollander *et al*., 2006; Brookes & Rochette, 2007), but this problem is rarely integrated in mollusc taxonomy. For all these reasons, taxonomists may differ in how they interpret variability of shell characters: what one specialist interprets as geographical or bathymetrical variation can be interpreted by another as specific differences. As a consequence, quite contradictory opinions about species delimitations in literature are not rare (e.g. Röckel**,** Rolan & Monteiro (1980) vs Monteiro, Tenorio & Poppe (2004) [Cape Verde *Conus*]).

The difficulty of confronting different opinions is also increased by the lack of a solid theoretical and methodological framework, *de facto* rendering many taxonomical opinions untestable hypotheses. Recently, concepts and methods underlying the delimitation of terminal taxa were clarified, advocating for the use of an integrative framework. De Queiroz (1998), followed by Samadi & Barberousse (2006), first suggested that a unified view of "what is a species" is possible if species are considered as definitively diverging lineages and most of the other so-called "species concepts" as criteria for delimiting species (Sites & Marshall, 2003). In this context, molecular characters are classically used to propose primary hypotheses of species delimitation based on genetic distances (Floyd *et al*., 2002; Vogler & Monaghan, 2007). Contrary to the morphological characters used in most mollusc species description, molecular characters are strictly heritable and reproducible. Within an integrative framework (De Queiroz, 2007) these primary molecular hypotheses can then be tested against several criteria, namely: (i) monophyly of primary species hypotheses (Wheeler & Meier, 2000; Meyer & Paulay, 2005), (ii) independent genetic markers, such as a nuclear gene, as gene trees do not necessarily reflect the species tree (Nichols, 2001; Funk & Omland, 2003), and (iii) morphological analysis to discuss species delimitations based on genetic data (Bichain *et al*., 2007; Pfenninger *et al*., 2006).

In this integrative context, we aim to combine several criteria to propose new hypotheses of species delimitations in the genus *Benthomangelia* (Conoidea, Conidae). This is a widely distributed genus of bathyal and abyssal marine molluscs (Bouchet & Warén, 1980). It was described by Thiele (1925), grouping together species previously placed in other genera. Since 1925, new species have been described, based on morphological characters such as ornamentations and shape of the shell (Bouchet & Warén, 1980; Sysoev & Ivanov, 1985; Sysoev, 1988). Ten species are now considered as valid (Tucker 2004), of which six are present in the Pacific. Several potential new forms have been recently collected during cruises organized by the Museum National d'Histoire Naturelle (MNHN) and the Institut de Recherche pour le Développement (IRD), but the great and sometimes continuous

morphological variability complicate the delimitation of species based solely on these characters (Bouchet & Sysoev, 2001).
We demonstrate the delimitation of genetic groups within the genus *Benthomangelia*, using molecular characters and a tree-based method. Two independent genes are used, one mitochondrial (COI) and one nuclear (28S) gene. An Elliptic Fourier Analysis (EFA; Rohlf, 1996) of the shape of the last whorl of the shell is also performed, allowing detailed analysis of complex structures as a whole (Monti, Baylac & Lalanne-Cassou, 2001). Contrary to shell characters used in most molluscan species descriptions, EFA has the advantage to be formalized and thus reproducible. EFA analysis complements the DNA-based species delimitation, as morphological characters are indispensable for the delimitation of species in an integrative framework. Indeed, the genetically defined groups can be described morphologically, on the basis of shell characters linked to interspecific genetic variation, as opposed to the morphological variability linked to geographical or ecological factors.

Material & Methods

Sampling
Specimens of *Benthomangelia* were collected between 2004 and 2007 during several deep-sea cruises conducted by the MNHN and the IRD in Philippines, Vanuatu and Solomon Islands (Table 1, Figure 1). Living specimens were anesthetized in $MgCl_2$ and fixed in 95 % ethanol. Shells were kept intact, so the same individuals were used for both molecular and morphological analyses.

Sequencing
DNA was extracted from a piece of foot, using 6100 Nucleic Acid Prepstation system (Applied Biosystem). Two gene fragments were amplified: (i) a fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene using universal primers LCO1490 and HCO2198 (Folmer *et al*., 1994) and (ii) a fragment of 900 bp of the rDNA 28S gene, using the primers C1 and D3 (Jovelin & Justine, 2001). All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq (MPBiomedicals) for all genes. COI gene amplifications were performed according to Hebert *et al*. (2003); for 28S gene, they consisted in an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 52°C and extension at 72°C for 1'. The final extension was at 72°C for 10'. PCR products were purified and sequenced by the Genoscope (genbank accession numbers EU015528, EU015628, EU015644, EU015743 and EU428956-EU429039). In all cases, both directions were sequenced to confirm accuracy of each haplotype sequence.
A specimen of the sister group of *Benthomangelia* (Puillandre *et al*., in press), the genus *Toxicochlespira* (17925, genbank accession number: EU015738 and EU015623 for COI and 28S gene respectively) and two other Conoidea (specimens 17866, *Mangelia*, EU015688 and EU015573 and 17754, *Turris*, EU015677 and EU015562) are used as outgroup. Outgroups were chosen to form a non-monophyletic group, as recommended by Darlu and Tassy (1993).

Table 1: Identification number, cruise, station, depth, coordinates and species number as determined by the molecular analysis are given for each specimen. In the column COI and 28S, a cross indicates that the specimen has been successfully sequenced for the gene.

ID	Cruise	Station	Depth (min-max)	Latitude	Longitude	DNA species	COI	28S
17950	Aurora 07	CP2678	507-540m	14°46.54N	123°09.69E	2		x
17952	Aurora 07	CP2729	593-600m	15°19.04N	121°37.02E	2	x	x
17953	Aurora 07	CP2729	593-600m	15°19.04N	121°37.02E	2	x	x
17954	Aurora 07	CP2734	453-460m	15°56.41N	121°48.71E	2	x	x
17955	Aurora 07	CP2734	453-460m	15°56.41N	121°48.71E	2	x	x
17956	Aurora 07	CP2734	453-460m	15°56.41N	121°48.71E	2	x	x
17957	Aurora 07	CP2749	473m	15°56.38N	121°49.46E	2	x	x
17958	Aurora 07	CP2749	473m	15°56.38N	121°49.46E	2	x	x
17959	Aurora 07	CP2749	473m	15°56.38N	121°49.46E	2	x	x
17960	BOA 1	CP2427	745-827m	14°56.2'S	166°54.8'E	1	x	x
17961	BOA 1	CP2422	667-750m	14°55.1'S	166°55.4'E	1	x	x
17835	BOA 1	CP2462	618-641m	16°37.5'S	167°57.4'E	1	x	x
17962	BOA 1	CP2462	618-641m	16°37.5'S	167°57.4'E	1	x	x
17963	BOA 1	CP2462	618-641m	16°37.5'S	167°57.4'E	2	x	x
17964	BOA 1	CP2461	582-614m	16°35.5'S	167°57.9'E	1		x
17965	BOA 1	CP2461	582-614m	16°35.5'S	167°57.9'E	1	x	x
17966	Panglao 05	CP2333	584-596m	9°38.2'N	123°43.5'E	1	x	x
17967	Panglao 05	CP2334	606-631m	9°37.5'N	123°40.2'E	1	x	x
17968	Panglao 05	CP2358	569-583m	8°52.1'N	123°37.1'E	1	x	x
17969	Panglao 05	CP2388	762-786m	9°26.9'N	123°34.5'E	1	x	x
17970	Panglao 05	CP2388	762-786m	9°26.9'N	123°34.5'E	1	x	x
17971	Panglao 05	CP2388	762-786m	9°26.9'N	123°34.5'E	1	x	x
17972	Panglao 05	CP2381	259-280m	8°43.3'N	123°19.0'E	4	x	x
17973	Panglao 05	CP2384	624-647m	8°46.2'N	123°16.1'E	1	x	x
17974	Panglao 05	CP2386	2120-2149m	8°49.3'N	123°01.9'E	1	x	x
17975	Panglao 05	CP2389	784-786m	9°27.9'N	123°38.4'E	1	x	x
17976	Panglao 05	CP2388	762-786m	9°26.9'N	123°34.5'E	1	x	x
17977	Panglao 05	CP2389	784-786m	9°27.9'N	123°38.4'E	1	x	x
17978	Panglao 05	CP2389	784-786m	9°27.9'N	123°38.4'E	1	x	x
17979	Panglao 05	CP2392	242-400m	9°29.0'N	123°41.1'E	4	x	x
17980	Panglao 05	CP2396	609-673m	9°36.3'N	123°42.0'E	1	x	x
17981	Salomon 2	CP2268	632-640m	7°48.7'S	156°53.3'E	5	x	x
17982	Salomon 2	CP2175	579-585m	9°05.8'S	158°59.9'E	1	x	x
17983	Salomon 2	CP2219	650-836m	7°58.3'S	157°34.4'E	3	x	x
17984	Salomon 2	CP2219	650-836m	7°58.3'S	157°34.4'E	3	x	x
17985	Salomon 2	CP2182	762-1060m	8°47'0S	159°37.9'E	1	x	
17930	Salomon 2	CP2269	768-890m	7°45.1'S	156°56.3'E	3	x	x
17986	Salomon 2	CP2196	724-765m	8°25.6'S	159°25.9'E	1	x	x
17987	Salomon 2	CP2196	724-765m	8°25.6'S	159°25.9'E	1	x	x
17988	Salomon 2	CP2196	724-765m	8°25.6'S	159°25.9'E	1	x	x
17989	Salomon 2	CP2264	515-520m	7°52.4'S	156°51.0'E	1	x	x
17990	Salomon 2	CP2195	543-593m	8°25.5'S	159°26.4'E	1	x	x
17991	Salomon 2	CP2194	440-521m	8°24.8'S	159°26.7'E	1	x	
17992	Salomon 2	CP2213	495-650m	7°38.7'S	157°42.9'E	1	x	x
17993	Salomon 2	CP2213	495-650m	7°38.7'S	157°42.9'E	1	x	x
17994	Santo 06	AT105	408-444m	15°03.0'S	166°34.5'E	4	x	x
17925	Salomon 2	CP2227	508-522m	6°37.2'S	156°12.7'E	*Toxicochlespira*	x	x
17866	Panglao 04	S19	3-4m	9°42.1'N	123°51.4'E	*Mangelia*	x	x
17754	Panglao 04	R42	8-22m	9°37.1'N	123°52.6'E	*Turris*	x	x

Genetic analysis

COI and 28S sequences were manually aligned, as no ambiguous indels were found. Standard molecular diversity indices were calculated using Arlequin 3.1 (Excoffier, Laval & Schneider 2005).

Genotypic clustering

Genetic pairwise distances (excluding outgroups) for each gene separately were calculated with PAUP 4.0b10 (Swofford, 2002), using the best fitting model of nucleotide substitution for each gene as defined by the program Modeltest (Posada & Crandall, 2001), in conjunction with PAUP 4.0b10, following the Akaike information criterion (AIC). These distances are visualized on a NJ tree calculated using PAUP 4.0b10 to define groups of specimens with low genetic distances within groups and high distances between groups.

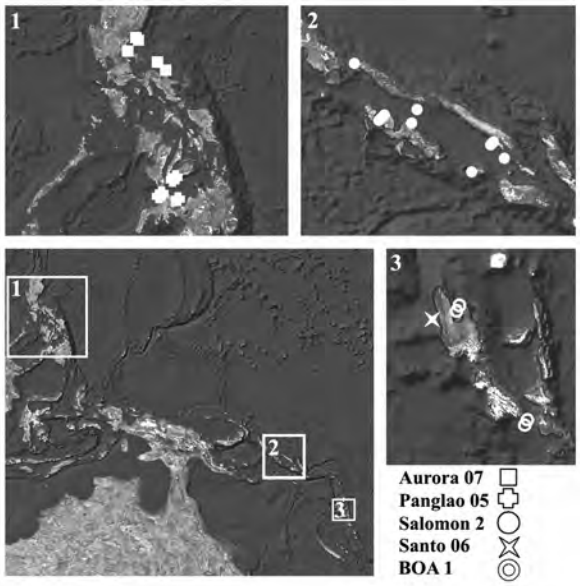


Figure 1: South-west Indo-Pacific map, with emphazis on Philippines (1), Solomon Islands (2) and Vanuatu (3). Sampling sites for each cruise are represented, with the same symbols used for CVA analyses (Figures 5, 6 and S2).

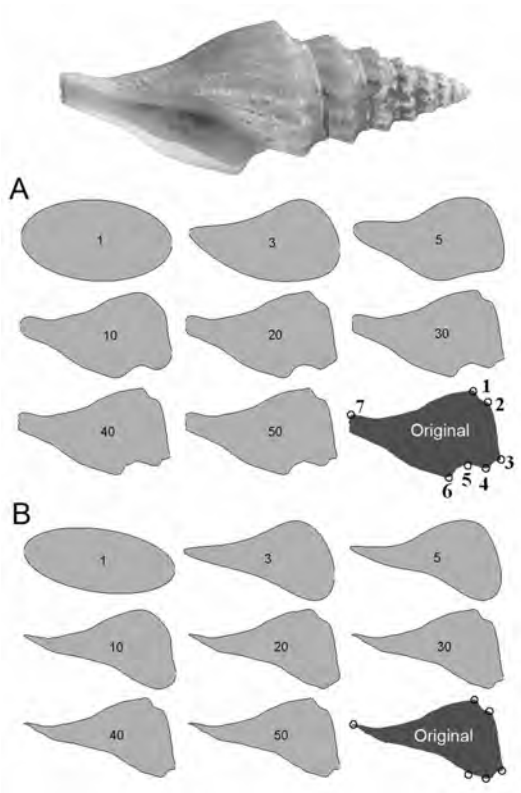


Figure 2: Outlines reconstructions with increasing number of harmonics indicated within outlines, for outline 1 (A) and outline 2 (B). The seven landmarks are represented on the original outline 1; only six landmarks were used for the second outline 2.

Phylogenetic analyses

Phylogenetic reconstruction were conducted using Bayesian Analysis (BA), consisting of two Markov chains (2000000 generations each with a sampling frequency of one tree each hundred generations) run in four parallel analyses using Mr. Bayes (Huelsenbeck, Ronquist & Hall, 2001). When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in. Only the number of nucleotide substitutions categories was fixed for BA. Phylogenetic analyses were performed on the cluster housed at the MNHN (17 nodes, 2 Go Ram per node, 30 AMDs 64 bits CPU's for the slave nodes and 4 Xeon 32 bits CPU's for the two master nodes)

Gene flow estimation

Within each group of specimens defined by genotypic clustering and phylogenetic analyses, population structure was estimated between groups of specimens collected in different geographic region. Arlequin 3.1 was used to perform AMOVA (with a 10000 permutations test) for each pair of populations.

Fourier analysis

Morphometric analyses were performed on the same 46 shells used for the molecular analysis. The shape of the last whorl has been previously considered informative in the alpha-taxonomy of *Benthomangelia* (Bouchet & Warén, 1980). In order to test the consistency of

this hypothesis with the results obtained with the genetic data, the morphometric analysis considered only this part of the shell. Shells were placed horizontally, aperture up, and digitized using a macro stand. To evaluate the potential positioning error, each shell was digitized twice. The same error was also estimated by digitizing three shells (chosen as three very similar ones) five times each (Position Test Dataset, PTD). As the labrum of some shells was broken, the lateral orientation of the shell could variate from one shell to another. To estimate potential error due to this problem of parallax, three shells were digitized five times, turning the shell around the columellar axis between each picture but keeping the columellar axis parallel to the support (Orientation Test Dataset, OTD).

Six homologous landmarks (Bookstein 1991) were defined in the last whorl of the shell, corresponding to the junction with the previous whorl (landmarks 2 and 3) and the aperture (landmark 5) and to the apex of the peripheral chord (landmarks 1, 4 and 6; Figure 2). As these six points covered only the upper part of the whorl, we also defined a seventh point located at the apex of the siphonal canal. The position of this point could be problematic for some shells, as the extremity of the canal can be broken, but the aim was to cover the whole outline. The seven landmarks and the outlines were digitized using TpsDig (Rohlf, 1996). We always used the first landmark as starting point. As the labrum of some shells was broken, two outlines were defined: the first corresponds to the whole last whorl (outline 1) and the second corresponds to the whole last whorl except the labrum by joining the extremity of the siphonal canal to the upper part of the aperture following the columellar lip (outline 2, Figure 2). These two outlines were analysed for the 46 shells photographed twice (complete dataset = 92 pictures) and for the PTD and OTD. All pictures and outlines were taken by the same person (NP).

Outlines were used as input for an EFA (Baylac & Friess, 2005). The seven landmarks were used as control points to rotate the outlines into the same orientation. For the analysis of the outline 2, the sixth landmark, placed on the labrum of the shell, was removed (Figure 2B). The images were then centred and normalised for size (using square roots of the surface). A visualization of Fourier reconstructions using different numbers of harmonics, compared to the original outline, shows that 40 harmonics are sufficient to reconstruct the outlines with high accuracy (Figure 2).

The variates resulting from the EFA and used in the subsequent analyses correspond to the different Fourier coefficients, as described in Rohlf 1996. PTD and OTD were studied using Principal Components Analysis (PCA) for both outlines 1 and 2. In order to eliminate potential error due to the positioning of the shell, the mean of each pair of Fourier outlines corresponding to the two replicates of each specimen was used for the complete dataset, resulting in the study of 46 outlines. As few specimens were used, an exploratory analysis was first performed for the complete dataset using PCA in order to evaluate the level of variability within and among the groups then tested with Canonical Variate Analysis (CVA). For CVA analysis, variability was maximised following two different grouping variables: the genetic groups as defined by the molecular analysis and the cruise of collection (Table 1). Visualizations of the outline deformations along the canonical axes were made using the procedure described in Monti *et al*. (2001). Multivariate regression parameters of Fourier coefficients were calculated using the depth of collection of each specimen as the independent grouping variable (the depth of collection is calculated as the mean between the depth at the beginning of the trawl and the depth at the end of the trawl, see table 1). Effect of the size of the last whorl on shape was also estimated, plotting the shape of the shell (measured as the projection on each axis of the CVA of the Fourier coordinates) as a function of the size of the specimens. All analyses were performed using specially devised MATLABv5.2 functions implemented by MB.

Results

Molecular analysis

Forty-four specimens were sequenced for COI, resulting in a 658 bp fragment. No indels were found. Forty-four specimens were sequenced for 28S (42 in common with COI, table 1), resulting in a 908 bp fragments after the alignment. The Kimura-3-parameter model (K81uf+I+G, with I = 0.4654 and G = 0.5381) model for the COI gene and the Tamura-Nei model (TrN+I+G, with I = 0.6746 and G = 0.6616) model for the 28S gene were defined as the best fitting models.

Genetic clustering

For the COI gene, 25 different haplotypes were found among the 44 specimens, displaying 186 polymorphic loci, and a high haplotypic diversity (0.948). The distribution of pairwise genetic distances for the COI gene is clearly bimodal (Figure 3). Indeed, the genetic distance between two specimens is either lower than 2.5% or higher than 7.5%. These two categories of distances are visualized in the NJ tree (not shown) as five long branches at the end of which from 1 to 28 genetically similar specimens are clustered. For the clarity of the discussion, we will refer to these branches as "groups" numbered from 1 to 5 (Table 1), although one includes only one specimen (17981). Specimens are from Panglao 05, BOA 1 and Salomon 2 for group 1, Aurora 07 and BOA 1 for group 2, Salomon 2 for group 3, Panglao 05 and Santo 06 for group 4 and Salomon 2 for group 5.

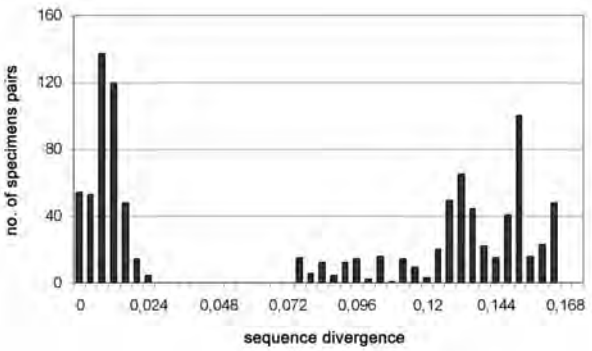


Figure 3: Histogram of genetic distances for COI gene.

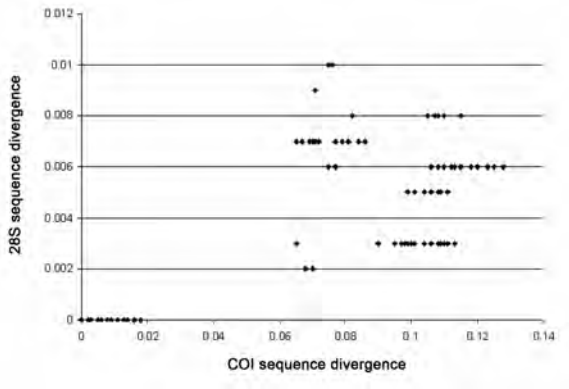


Figure S1: Comparison of 28S and COI genetic distances for each specimen.

Results obtained with the 28S are congruent with those obtained with COI gene. Overall variability for the 28S gene is less than for the COI gene, with only 5 different haplotypes among the 44 specimens, 36 polymorphic loci and a haplotypic diversity of 0.584. This low level of variability does not allow the recognition of a gap between short and large genetic distances, but specimens displaying short distances with the COI gene possess exactly the same 28S sequence. This complete congruency between both genes is also visualized in the figure S1 (supplementary material). Indeed, the same five groups are recognized in the 28S and COI trees.

Phylogenetic analysis

The four genetic groups containing several specimens are found monophyletic with the BA for the COI gene, although the monophyly of one group is not supported (posterior probabilities >0.99 for groups 1, 4 and 5, posterior probabilities = 0.91 for group 1; results not

shown). Since no incongruence is found between the two genes, a dataset combining the two genes for 42 specimens is created. Here, the data were separated into two different unlinked partitions corresponding to the two analyzed genes, each following the best fitting model of substitution estimated for each gene. The four groups are in this case found highly supported, but the relationships among the five groups are not confidently resolved (posterior probabilities < 0.95; Figure 4).

Gene flow estimation

It was possible to estimate gene flows only for group 1 because it was the only one that sampled several specimens from each of the different geographic regions. Population structure was evaluated only with the COI gene since the 28S gene did not vary within groups. Three populations were defined: the first including five specimens from BOA 1 (Vanuatu), the second 13 specimens from Panglao 05 (Philippines) and the last 10 specimens from Salomon 2 (Solomon Islands). The pairwise populations comparisons (F_{st}) among the three populations are 0.066 (p-value = 0.106) between BOA 1 and Salomon 2, 0.321 (p-value = 0.014) between BOA 1 and Panglao 05 and 0.239 (p-value = 0.003) between Salomon 2 and Panglao 05.

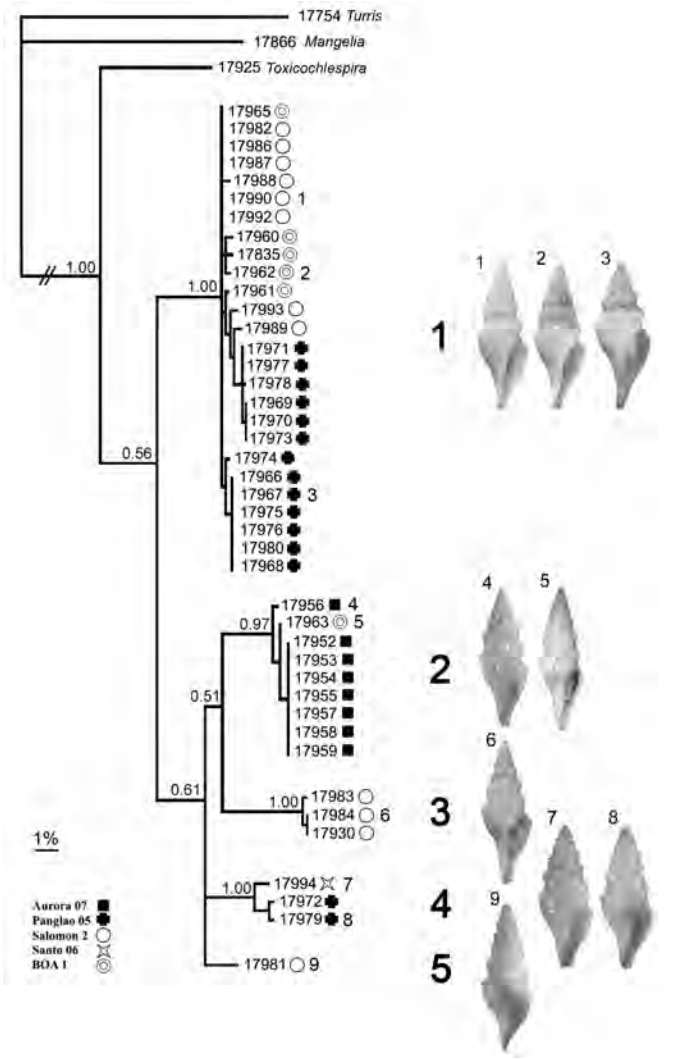


Figure 4: Bayesian tree for combined dataset corresponding to both genes. Posterior Probabilities (superior to 0.5) are given for each node. Groups are numbered top downwards from 1 to 5. For each group and for each cruise within group, one shell is illustrated (numbered from 1 to 9).

Fourier analysis

Analyses of potential error linked to positioning (PTD) or orientation (OTD) both show that variability between the replicates of the same shell is always less than the variability between different shells, as revealed by the PCA, except for the PTD with outline 1 (results not shown).

For both outlines 1 and 2, the PCA revealed high variability within and between genetic groups, allowing separation of two groups (genetic groups 1 and 3 vs genetic groups 2, 4 and 5; supplementary material, Figure S2A). The groups of specimens that correspond to different cruises of collection are not separated (Figure S2B).

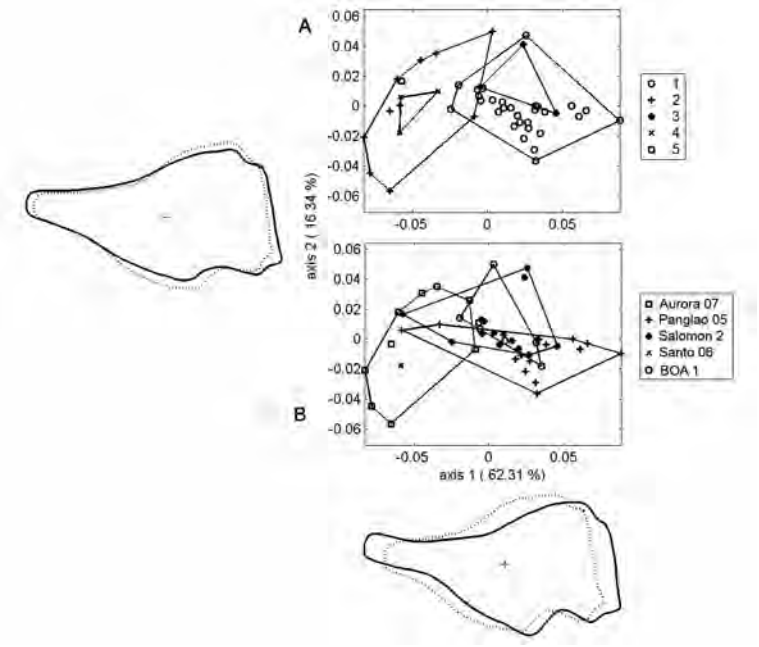


Figure S2: PCA for the complete dataset, with a visualisation of the genetic groups (A) and the cruise of collection (B). Only results for the first outline are shown. Superimposed outlines for minimum (dotted line) and maximum (full line) projections onto the two principal axes are represented.

The variable "genetic groups" is significantly discriminant using CVA (*F* test associated to the Wilk's lambda = 2.65, d*f* = 80, $p<10^{-4}$ and *F* = 2.90, d*f* = 80, $p<10^{-4}$ for outline 1 and 2 respectively). Axes 1 and 2 represent together 91.13% of the variance for outline 1 (Figure 5A) and 86.19% for outline 2 (Figure 5C). They allow the separation of the five genetic groups, except the groups 4 and 5 for outline 2. The variable "cruise of collection" is also significantly discriminant (*F* = 3.27, df = 80, $p<10^{-4}$ and *F* = 2.27, d*f* = 80, $p<10^{-4}$ for outline 1 and 2 respectively). Axes 1 and 2 represent together 83.2% of the variance for outline 1 (Figure 5B) and 80.84% for outline 2 (Figure 5D). However, contrary to the results obtained with the genetic groups, specimens from Panglao 05, Salomon 2 and BOA 1 are not completely discriminated.

The first axis shows up an opposition between short and large shells on one hand and more elongated and thin shells on the other hand for both CVA analyses (using "genetic groups" or "cruise of collection" as variables). This can be due to the strong relationships between these two variables (Chi^2 test: $p<10^{-4}$). On the second axis, genetic groups are separated by the length of the siphonal canal (Figure 5A and C), and specimens from different cruises of

collection are separated by the shape of the curvature at the beginning of the siphonal canal (Figure 5B and D), more marked for some specimens collected during Panglao 05.

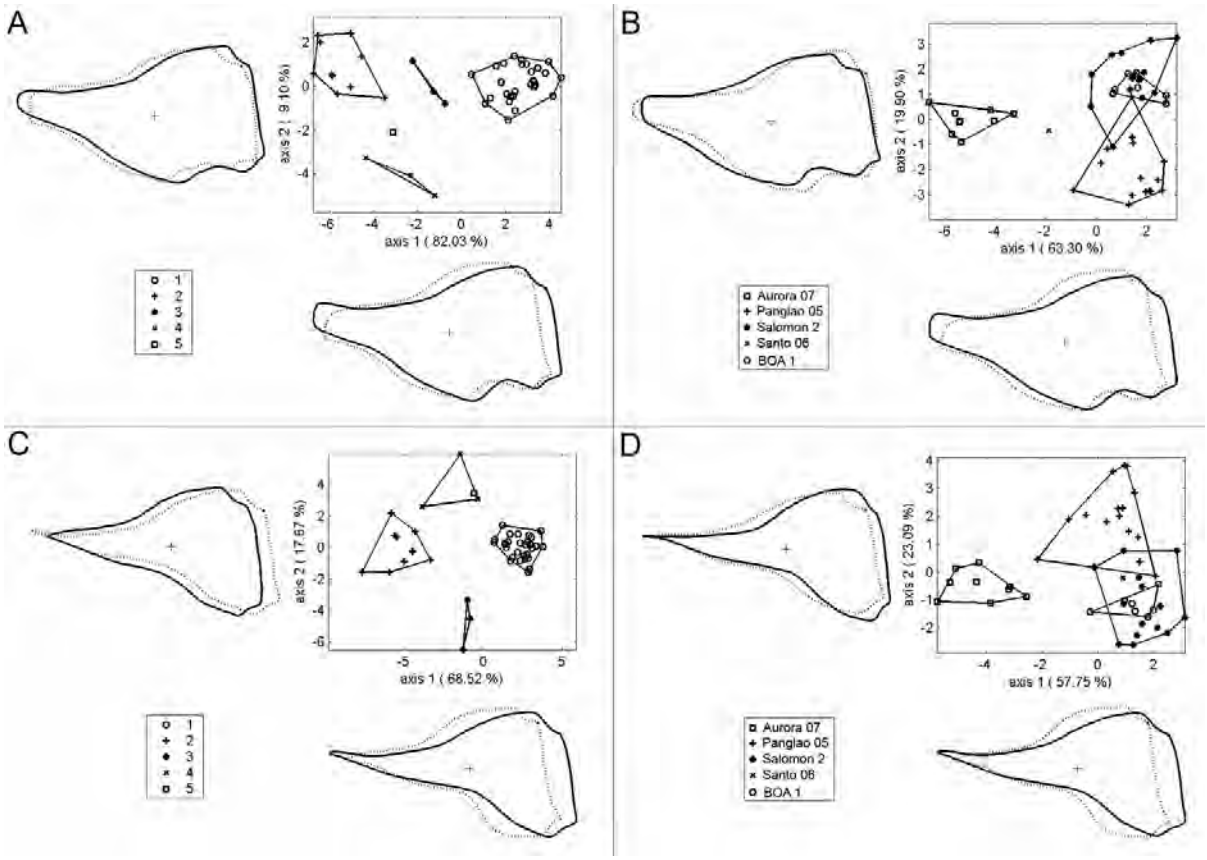


Figure 5: CVA for the complete dataset. CVA for the first outline using genetic groups as grouping variable (A), CVA for the first outline using cruise of collection as grouping variable (B), CVA for the second outline using genetic groups as grouping variable (C), CVA for the second outline using cruise of collection as grouping variable (D). Superimposed outlines for minimum (dotted line) and maximum (full line) projections onto the two principal axes are represented.

The following axes do not allow a better separation of the specimens from different cruises. The same patterns are conserved if the unique specimen of the group 5 is removed (CVA using "genetic groups" as variables), if the unique specimen of the cruise Santo 06 is removed (CVA using "cruise of collection" as variables) or if only the groups with the most specimens are considered (genetic groups 1 and 2; CVA using "genetic groups" as variables).

Multivariate regressions using the depth of collection as the independent grouping variable are not significant, whatever the number of axes used: $\alpha = 0.86$, 0.65, 0.50 and 0.79 for outline 1 and $\alpha = 0.77$, 0.85, 0.63 and 0.62 for outline 2, for 3, 5, 10 and 20 components respectively.

To avoid problems due to the correlation of the different variables, and to estimate the level of morphological variability between different cruises of collection within a genetic group, CVA was performed within the one genetic cluster with multiple samples from multiple localities (group 1: 6 specimens from BOA 1, 13 from Panglao 05 and 10 from Salomon 2) to test for geographic signal within group. The three different cruises are clearly separated on the two first axes, for both outline 1 and 2 (results shown only for outline 1, Figure 6). The first axis represents more than 87% of the variability and distinguishes Panglao 05 specimens from

Salomon 2 and BOA 1 specimens, based again on the shape of the curvature at the beginning of the siphonal canal, as found with CVA including the complete dataset.

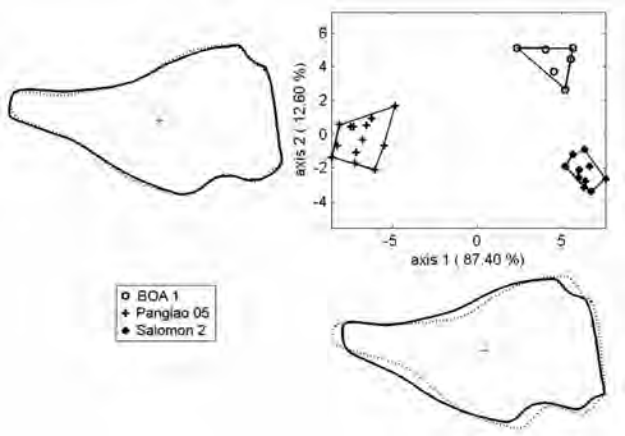


Figure 6: CVA on the first outline for the genetic group 1. Superimposed outlines for minimum (dotted line) and maximum (full line) projections onto the two principal axes are represented.

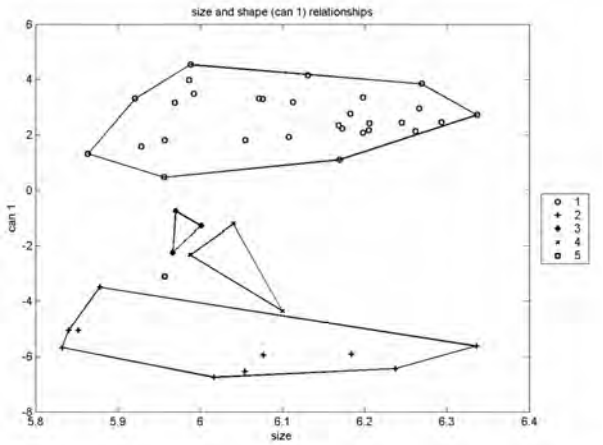


Figure S3: Plots of the shape of the last whorl of the shell as a function of the size. The five genetic groups are represented. Size of the shells was measured as the square root of the area.

Multivariate regression using the depth of collection was also performed for the genetic groups 1 and 2: no significant correlation is found (group 1, outline 2: $F = 0.505$, $p = 0.770$; group 2, outline 1: $F = 1.281$, $p = 0.396$; group 2, outline 2: $F = 0.898$, $p = 0.0.546$), except for the group 1 with the outline 1 ($F = 6.53$, $p<10^{-4}$). However, when removing one of the specimen (17987), whose labrum is broken, test is not significant anymore ($F = 1.018$, $p = 0.1879$).

Finally, a potential effect of size on the shape of the last whorl between genetic groups or cruises of collection was tested: no significant correlation were found, as the entire range of size is found in each numerous groups (genetic groups 1 and 2 and cruise of collection Panglao 05, BOA 1, Aurora 07 and Salomon 2). Results are shown only for the genetic groups, using the first axis of the CVA (supplementary material, Figure S3).

Discussion

Species delimitation based on genetic data

The number of specimens included in the molecular analysis performed with COI and 28S genes allows for the detection of five genotypic groups, four of which are represented by several specimens. Since there is a strict congruency between the two unlinked genes, our groups are not only clusters of haplotypes but real genotypic clusters (Mallet, 2001; Sites & Marshall, 2004). Phylogenetic analyses also suggest that the groups defined with the genotypic clustering criterion are monophyletic, thus each displaying an independent event of coalescence for two independent genes (COI and 28S), suggesting the absence of genetic exchange between the groups (Ferguson, 2002). Moreover, as three of the five groups include specimens from different geographic regions (e.g. Philippines and Vanuatu), the genetic groups do not reflect a geographic structure among distant populations. The expected dispersal capacity of the larvae of *Benthomangelia* supports the observed genotypic clustering as members of this genus have a planktotrophic protoconch (between 2.5 and 3 whorls),

indicating that the larvae can stay in the water column for a long period and can potentially disperse long distances (Shuto, 1974).
Thus, on the basis of our results, five lineages can be recognized, and we propose these lineages as primary hypotheses of species delimitation in *Benthomangelia*. The first species (group 1) is identified as *B. trophonoidea*, and the four other species does not correspond to any of the nine other described species of *Benthomangelia*, and could thus constitute new species (A. Sysoev, pers. com.).
However, recent species are not necessarily characterized by high genetic distances between species or by reciprocal monophyly (Hickerson, Meyer & Moritz, 2006; Knowles & Carstens 2007). Consequently we cannot exclude that the moderate genetic structure within group 1, especially between the Philippines (Panglao 05 cruise) and the two others archipelagos (Solomon Islands and Vanuatu), is due to the presence of several species. This structure can also be interpreted as intraspecific genetic structure linked to geographic differentiation, as Philippines and Solomon islands are separated by more than 4000 km, and Philippines and Vanuatu by more than 5000 km. Similarly, weak but non-negligible genetic structure is found between Solomon Islands and Vanuatu, separated by only 1200 km.
The only group present in three different geographic regions is also the group with the most specimens. Consequently, the apparently limited geographic distribution of some species, illustrated by the strong correlation between the two variables "genetic groups" and "cruise of collection", can only be an artifact due to undersampling of some species. Furthermore, it is important to note that our data set does not cover the whole genus since the whole Pacific is not sampled and putative distinct species of *Benthomangelia* from the Atlantic are not represented (Bouchet & Warén, 1980).

Comparison with morphological results

While the heritability of morphologic characters is difficult to document, the use of objective and repeatable morphometric analysis allows retrieving the five genetic groups. A large part of the morphological variability of the last whorl of the shell gives valuable characters to delimit species. Specimens from genetic group 1 are characterized by a more elongated shell and a long canal, those from group 2 also by a long canal, but a relatively short last whorl, and those from group 4 by a short shell and a short canal. Outline 2 does not distinguish genetic group 4 and 5, suggesting that the labrum could also be a valuable character to delimit species (Kohn, pers. com.).
However, the morphometric analysis also revealed that a part of the variability of the shape of the shell imparts the separation of specimens collected during different cruises within group 1. Thus, specimens of group 1 collected during Panglao 05 are characterized by a more vaulted siphonal canal, in opposition to the specimens collected during Salomon 2 and BOA 1. The morphologic differentiations between geographic groups within group 1 can be either associated to the genetic structure found between the same groups or to a morphologic plasticity of the shell in response to different environment conditions encountered in the different archipelagos. The potential combination of genetic and environmental effects on the shape of the shells constitutes one of the reasons why it is difficult to use solely the morphological characters in such a problematic groups.

On the use of shell characters to delimit species

Shell characters are known to be highly plastic, and several studies clearly identify characters that vary according to environmental factors: shells can be thicker (Brookes & Rochette, 2007) and shape can be modified in response to the presence of predators (Palmer, 1990), to flow velocity (Baker *et al.*, 2004) or biochemical conditions (Chiu *et al.*, 2002), and ornamentation of the shell can vary along with the environment (Yeap, Black & Johnson,

2001). The analysis of several character sets, such as DNA and morphology, is needed to avoid the use of characters that reflects environmental differences rather than taxonomic differences (Kantor *et al*., unpublished data; Samadi *et al*., 2000; Bichain *et al*., 2007). For example, in the pulmonate genus *Radix*, shell characters are correlated with environmental factors, and do not differ constantly among the species defined with DNA (Pfenninger *et al*., 2006). These authors suggest that erroneous species delimitations based on unsuitable morphological characters are likely to be widespread in molluscs with species description based solely on shell characters. Without a test of morphological characters using an integrative approach, taxonomists risk defining species based on non-genetically determined characters (Godfray, 2007). Mating trials are useful to test for species boundaries (Ribi & Porter, 1995; Pickles & Grahame 1999), but such tests are difficult with deep-sea groups such as *Benthomangelia*. In these cases, combining genetic and morphological analyses appears to be a robust method to propose hypotheses of species delimitations.

Conclusion

In *Benthomangelia*, we demonstrated that the morphological characters analysed are congruent with the groups recognized by genetic analyses. They can thus be used in systematic to delimitate species and identify specimens. Furthermore, the use of reproducible methods, for both genetic and morphological analyses, will allow future tests of these hypotheses of species delimitation, including more replicates but also other species that are not represented here. The identification of taxonomically valuable morphological characters (i.e. genetically determined) is of particular interest for marine gastropods, where the majority of collected specimens are empty shells. In the perspective of a more complete assessment of species diversity of *Benthomangelia*, CVA analyses performed with genetically and morphologically characterized specimens can be used as a guideline to include in the same analysis empty shells for which DNA characters are not available.

Acknowledgements

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the *Tropical Deep Sea Benthos* programme on board R/V *Alis*, that generated the deep-sea samples used in this study. We are also pleased to thank the staff of MNHN's "Service de Systématique Moléculaire" for technical facilities; the Consortium National de Recherche en Génomique, Génoscope for the sequencing facilities, Allowen Evin, Raphaël Cornette and Vincent Debat ("Plateforme Morphométrie") for their help in the morphological analysis and Philippe Jarne for constructive comments on the manuscript. The authors thank Line Le Gall, Mandë Holford, Alan J. Kohn and two anonymous referees that greatly improved the quality of the manuscript.

References

Baker AM, Sheldon F, Somerville J, Walker KF, and Hughes JM. 2004. Mitochondrial DNA phylogenetic structuring suggests similarity between two morphologically plastic genera of Australian freshwater mussels (Unionoida: Hyriidae). *Molecular Phylogenetics and Evolution* **32:** 902-912.

Baylac M, and Friess M. 2005. Fourier descriptors, Procrustes superimpositions and data dimensionality: An example of cranial shape analysis in modern human populations. In: D.E. S, ed. *Modern Morphometrics in Physical Anthropology*. Chicago: Kluwer. 145-165.

Bichain JM, Boisselier MC, Bouchet P, and Samadi S. 2007. Delimiting species in the genus *Bythinella* (Mollusca: Caenogastropoda: Rissooidea): molecular and morphometric approachs. *Malacologia* **49:** 291-311.

Bookstein FL. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.

Bouchet P. 1990. Turrid genera and mode of development: the use and abuse of protoconch morphology. *Malacologia* **32:** 69-77.

Bouchet P, Lozouet P, Maestrati P, and Héros V. 2002. Assessing the magnitude of species richness in tropical marine environments : exceptionnally high numbers of molluscs at a new Caledonian site. *Biological Journal of the Linnean Society* **75:** 421-36.

Bouchet P, and Sysoev A. 2001. *Typhlosyrinx*-like tropical deep-water turriform gastropods (Mollusca, Gastropoda, Conoidea). *Journal of Natural History* **35:** 1693-1715.

Bouchet P, and Warén A. 1980. Revision of the north-east Atlantic bathyal and abyssal Turridae (Mollusca, Gastropoda). *Journal of Molluscan Studies* **SuppI. 8:** 1-19.

Boyle EE, Zardus JD, Chasea MR, Ettera RJ, and Rex MA. 2004. Strategies for molecular genetic studies of preserved deep-sea macrofauna. *Deep-Sea Research I* **51:** 1319–1336.

Brookes JI, and Rochette R. 2007. Mechanism of a plastic phenotypic response: predator-induced shell thickening in the intertidal gastropod *Littorina obtusata*. *Journal of Evolutionary Biology* **20:** 1015-1027.

Chiu YW, Chen HC, Lee SC, and Chen CA. 2002. Morphometric Analysis of Shell and Operculum Variations in the Viviparid Snail, *Cipangopaludina chinensis* (Mollusca: Gastropoda), in Taiwan. *Zoological Studies* **41:** 321-331.

Darlu P, and Tassy P. 1993. *La reconstruction phylogénétique. Concepts et méthodes*. Masson, Paris.

De Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJaB, S.H, ed. *Endless forms. Species and speciation*. Oxford: Oxford University Press. 57-75

De Queiroz K. 2007. Species Concepts and Species Delimitation. *Syst. Biol.* **56:** 879-886.

Excoffier L, Laval G, and Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47-50.

Ferguson JWH. 2002. On the use of genetic divergence for identifying species. *Biological Journal of the Linnean Society* **75:** 509-516.

Floyd R, Abebe E, Papert A, and Blaxter M. 2002. Molecular barcodes for
soil nematode identification. *Molecular Ecology* **11:** 839-850.

Folmer O, Black M, Hoeh W, Lutz R, and 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.

Funk DJ, and Omland KE. 2003. Species-level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* **34:** 397-423.

Godfray HCJ. 2007. Linnaeus in the information age. *Nature* **446:** 259-260.

Hebert PDN, Cywinska A, Ball SL, and deWaard JR. 2003. Biological identifications through DNA Barcodes. *Proceedings of the Royal Society of London B* **270:** 313-21.

Hickerson JH, Meyer CP, and Moritz C. 2006. DNA Barcoding Will Often Fail to Discover New Animal Species over Broad Parameter Space. *Systematic Biology* **55:** 729-739.

Hollander J, Collyer ML, Adams DC, and Johannesson K. 2006. Phenotypic plasticity in two marine snails: constraints superseding life history. *Journal of Evolutionary Biology* **19:** 1861–1872.

Huelsenbeck JP, Ronquist F, and Hall B. 2001. MrBayes: bayesian inference of phylogeny. *Bioinformatics* **17:** 754-755.

Jovelin R, and Justine JL. 2001. Phylogenetic relationships within the Polyopisthocotylean monogeneans (Plathyhelminthes) inferred from partial 28S rDNA sequences. *International Journal for Parasitology* **31:** 393-401.

Knowles LL, and Carstens BC. 2007. Delimiting Species without Monophyletic Gene Trees. *Systematic Biology* **56:** 887-895.

Mallet J. 2001. The speciation revolution. *Journal of Evolutionary Biology* **14:** 887-888.

Meyer PC, and Paulay G. 2005. DNA Bracoding: Error rates based on comprehensive sampling. *PLoS Biology* **3:** 1-10.

Monteiro A, Tenorio M, and Poppe G. 2004. *The West African and Mediterranean species of Conus*. ConchBooks, Hackenheim.

Monti L, Baylac M, and Lalanne-Cassou B. 2001. Elliptic Fourier analysis of the form of genitalia in two *Spodoptera* species and their hybrids (Lepidoptera: Noctuidae). *Biological Journal of the Linnean Society* **72:** 391-400.

Nichols R. 2001. Gene trees and species trees are not the same. *Trends in Ecology & Evolution* **16:** 358-364.

Palmer AR. 1990. Effect of crab effluent and scent of damaged conspecifics on feeding, growth, and shell morphology of the Atlantic dogwhelk *Nucella lapillus* (L.). *Hydrobiologia* **193::** 155-182.

Pfenninger M, Cordellier M, and Streit B. 2006. Comparing the efficacy of morphologic and DNA-based taxonomy in the freshwater gastropod genus *Radix* (Basommatophora, Pulmonata). *BMC Evolutionary Biology* **6:** 100.

Pickles AR, and Grahame J. 1999. Mate choice in divergent morphs of the gastropod mollusc *Littorina saxatilis* (Olivi): speciation in action? *Animal Behaviour* **58:** 181-184.

Posada D, and Crandall KA. 2001. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). *Molecular Biology and Evolution* **18:** 897-906.

Puillandre N, Samadi S, Boisselier MC, Sysoev AV, Kantor YI, Cruaud C, Couloux A, and Bouchet P. Starting to unravel the toxoglossan knot: Molecular phylogeny of the "turrids" (Neogastropoda: Conoidea). *Molecular Phylogenetics and Evolution***:** In Press. doi:10.1016/j.ympev.2007.11.007.

Ribi G, and Porter AH. 1995. Mating between two hybridizing species, *Viviparus ater* and *V. contectus* (Mollusca: Prosobranchia). *Animal Behaviour* **49:** 1389–1398.

Röckel D, Rolan E, and Monteiro A. 1980. *Cone shells from Cape Verde Islands. A difficult puzzle.* . Feito.

Rohlf FJ. 1996. *TpsDig*. Available at Stony Brook, http://life.bio.sunysb.edu/morph/.

Samadi S, and Barberousse A. 2006. The tree, the network, and the species. *Biological Journal of the Linnean Society* **89:** 509-521.

Samadi S, Roumegoux A, Bargues MD, Mas-Coma S, Yong M, and Pointier JP. 2000. Morphological studies of Lymnaeid snails from the human Fasciolasis endemic zone of Bolivia. *Journal of Molluscan Studies* **2000:** 31-44.

Shuto T. 1974. Larval ecology of prosobranch gastropods and its bearing on biogeography and paleontology. *Lethaia* **7:** 239-256.

Sites JW, and Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* **19:** 462-470.

Sites JW, and Marshall JC. 2004. Operational criteria for delimiting species. *Annual Review of Ecology Evolution and Systematics* **35:** 199-227.

Swofford DL. 2002. PAUP*. Phylogenetic Analysis Using Parsimony* and other methods. . Sunderland: Sinauer Associates.

Sysoev AV. 1988. Ultra-abyssal findings of the family Turridae (Gastropoda, Toxoglossa) in the Pacific Ocean. *Zoologicheskii Zhurnal* **67:** 972.

Sysoev AV, and Ivanov DL. 1985. New taxa of the family Turridae (Gastropoda, Toxoglossa) from the Naska Ridge (south east Pacific). *Zoologicheskii Zhurnal* **64:** 197.

Taylor JD, Kantor YI, and Sysoev AV. 1993. Foregut anatomy, feedings mechanisms and classification of the Conoidea (= Toxoglossa)(Gastropoda). *Bull. Nat. Hist. Mus. Lond* **59:** 125-70.

Tucker JK. 2004. Catalogue of recent and fossil turrids (Mollusca: Gastropoda). *Zootaxa* **682:** 1-1295.

Vogler AP, and Monaghan MT. 2007. Recent advances in DNA taxonomy. *Journal of Zoological Systematics and Evolutionary Research* **45:** 1-10.

Wheeler QD, and Meier R. 2000. *Species concepts and phylogenetic theory*. Columbia University Press, New-York.

Yeap KL, Black R, and Johnson MS. 2001. The complexity of phenotypic plasticity in the intertidal snail *Nodilittorina australis*. *Biological Journal of the Linnean Society* **72:** 63-76.

4. Autres caractères morphologiques : le genre Xenuroturris

Au sein de l'espèce *Xenuroturris cingulifera*, Yuri Kantor, spécialiste du groupe, a détecté la présence chez plusieurs spécimens de deux types de radula très distincts (notées par la suite « semi-enrolled » et « duplex »). Les différences détectées entre ces deux types de radula correspondent habituellement à des différences trouvées entre genres, voire entre sous-familles différentes. De la même façon que l'analyse du gène COI a permis dans les autres genres de remettre en cause les délimitations d'espèces basées sur l'analyse de la coquille, l'utilisation de nouveaux caractères morphologiques, en l'occurrence la radula, permet de proposer de nouvelles hypothèses de délimitation d'espèces. L'espèce *X. cingulifera*, telle qu'elle est définie dans la littérature sur des caractères de la coquille, inclurait en fait deux espèces tel que le suggère l'analyse de la radula. L'examen de la coquille a permis de plus d'identifier une différence (portant sur le nombre de cordons au niveau de l'épaulement) entre les spécimens possédant des radulae différentes. Cependant, cette différence est faible et avait été interprétée comme de la variabilité intraspécifique.

Au sein de la collection « barcode » constituée au MNHN de Paris, seuls quelques spécimens, tous récoltés lors de la mission Santo 2006, possédaient une morphologie correspondant à l'espèce *X. cingulifera*. Parmi ces spécimens, tous possédaient une radula « semi-enrolled », sauf un qui possédait une radula « duplex ». D'autres spécimens, récoltés également lors de la mission Santo 2006, ont été inclus dans l'analyse : ils possèdent tous une radula « semi-enrolled », mais une morphologie différente de l'espèce *X. cingulifera*.

L'analyse des gènes COI et 28S confirme que les spécimens de l'espèce *X. cingulifera* avec des radulae différentes appartiennent en fait à des espèces différentes. De plus, l'analyse a permis de montrer que les spécimens qui avaient une morphologie différente de *X. cingulifera* et une radula « semi-enrolled » appartiennent à deux autres espèces. Enfin, l'analyse phylogénétique montre que les trois espèces avec une radula « semi-enrolled » forme un clade. Les trois espèces qui possèdent une radula « semi-enrolled » ont été placées dans le genre *Iotyrris*. L'espèce dont la morphologie correspond au nom *X. cingulifera* porte donc le nom de *Iotyrris cingulifera*. Les deux autres espèces délimitées ici et possédant une radula « semi-enrolled » sont nouvelles, et ont été décrites dans l'article 8 : *I. musivum* et *I. devoizei*. L'espèce avec une radula « duplex » reste dans le genre *Xenuroturris*, et porte le nom *X. legitima*.

Cet exemple a permis de démontrer qu'une variabilité insoupçonnée peut-être détectée non seulement en utilisant des caractères moléculaires, mais également d'autres caractères morphologiques qui ne sont pas utilisés fréquemment au niveau spécifique, au moins chez les Turridae. Les deux espèces *I. cingulifera* et *X. legitima*, malgré le fait qu'elles soient très différentes moléculairement et au niveau de la radula et qu'elles ne soient pas groupes-frères, possèdent donc une coquille très semblable. Les mécanismes qui peuvent expliquer une telle évolution ne sont pas clairement identifiés. Il pourrait s'agir d'un cas de convergence morphologique, où la coquille serait le résultat de pressions environnementales similaires chez ces deux espèces. Pour tester cette hypothèse, il faudrait pouvoir mettre en évidence que les conditions environnementales (présence de prédateurs, habitat,…) dans lesquelles ces deux espèces vivent sont similaires. La conservation d'une forme ancestrale peut également expliquer ce pattern : les coquilles de ces deux espèces (*I. cingulifera* et *X. legitima*) auraient peu changées, alors que celles des deux autres espèces (*I. musivum* et *I. devoizei*) se seraient fortement différenciées.

Morphological proxies for taxonomic decision in turrids (Mollusca, Neogastropoda): a test of the value of shell and radula characters using molecular data

Yu. I. Kantor*, N. Puillandre**, B. M. Olivera***, P. Bouchet**

*A.N. Severtzov Institute of Ecology and Evolution of Russian Academy of Sciences, Leninski prosp. 33, Moscow 119071, RUSSIA, kantor@malaco-sevin.msk.ru;
**Muséum National d'Histoire Naturelle, 55, Rue Buffon, 75005 Paris, FRANCE, puillandre@mnhn.fr, pbouchet@mnhn.fr;
*** Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112, Utah, USA.

Zoological Science, in press

Abstract

The state of the art of turrid (= Turridae s.l.) systematics is that shells - when it includes the protoconch - are reliable species-level identifiers, but inadequate proxies for allocation to genera or subfamilies. Generally, the radula is used for allocation to a (sub)family, but the hypothesis that the radula is a more adequate proxy than the shell for relationships has not yet been tested by molecular data. Species of *Xenuroturris* may have drastically different radulae, with either "semi-enrolled" or "duplex" marginal teeth, although their shells are very similar or even almost indistinguishable. Molecular data confirm that specimens with different types of radulae constitute different species, but two species of a pair with respectively "semi-enrolled" and "duplex" teeth end up being not closely related. However, it is still unresolved whether species with "semi-enrolled" [= *Iotyrris*] and "duplex" teeth [= *Xenuroturris*] form two supported monophyletic clades. *Iotyrris devoizei* n.sp., and *I. musivum* n.sp., are described from Vanuatu, where they occur sympatrically with *I. cingulifera* and *Xenuroturris legitima*.

Key-words
shell, radula, COI, 28S, convergence, new species.

Running head
Morphological proxies for taxonomic decision in turrids

Introduction

Among the gastropod superfamily Conoidea, "Turridae" (sensu Powell, 1966) is the most speciose group of marine mollusks, with estimates of about 340 Recent valid genera and subgenera (Taylor *et al.*, 1993) and 4,000 named living species (Tucker, 2004). Tropical and deep-water faunas are reservoirs of vast numbers of undescribed species: Bouchet *et al.* (2004) estimated that 58-72% of the 1,726 "turrid" species sampled in New Caledonia is undescribed, and the real number of Recent "turrid" species worldwide may be in the 10-20,000 species range. As is the case with other specialist predators, the vast majority of species occur in low abundance. As a result of a very large number of species and small number of specimens, turrid systematics is intimidating, and turrids are much less studied than other marine gastropods of comparable size. In the history of turrid systematics, shell, radula and anatomy have been given different value and emphasis at different taxonomical levels. Shells - when it includes the protoconch - are usually considered reliable species-level identifiers. However, convergence and homoplasy renders them much less reliable predictors of relationships at higher taxonomic levels (family, genus). For instance, shells of *Toxicochlespira* Sysoev and Kantor, 1990 (Conidae) strongly resemble representatives of *Cochlespira* Conrad, 1865 (Turridae) (Sysoev and Kantor, 1990); shells of *Strictispira* McLean, 1971 (Strictispiridae), are hardly distinguishable from those of many species of *Crassispira* Swainson, 1840 (Turrridae, Crassispirinae) (Tippett, 2006); and the radula-less species *Cenodagreutes aethus* Smith, 1967 is said to be conchologically indistinguishable from the radulate *Raphitoma leufroyi* (Michaud, 1828) (both Conidae, Raphitominae).
At higher taxonomic levels, anatomical characters, especially foregut anatomy, are used to define subfamilies and families (Taylor *et al.*, 1993). However, as these are accessed only after time consuming serial histological sectioning and reconstruction (e.g. Kantor, 1990), the radula is routinely used as a proxy for the rest of foregut anatomy. Radulae have also the advantage that they can be extracted from animals with dried soft parts. Most of the currently recognized (Taylor *et al.*, 1993) subfamilies within Turridae and Conidae were erected (e.g., McLean, 1971), or subsequently redefined, on the basis of radular characters.
The state of the art of turrid systematics is thus based on the shell for species delimitation and/or identification, and on the radula for allocation to a (sub)family. These traditionally held beliefs were shattered when, during the course of routine radular preparations from New Caledonia "turrids", we found two distinct radular types in different specimens of what we initially classified as a single species, generally identified as *Xenuroturris cingulifera* (Lamarck, 1822), a species widely distributed in the Indo-Pacific (Powell 1964). These differences in radulae would be considered to imply at least generic rank when applying conventional taxonomic decision, but in this case the specimens involved have shells so similar that they would hardly be separated as discrete morphs, and even less so as species, even by the most hardy conchologist. This prompted examination of radula and molecular characters of conchologically similar animals from the South and West Pacific, which resulted in evidence of further species "pairs" with similar shells and different radulae.
This is the first time such a case is reported in "turrids", and it brings new questions: is this an isolated case of radula polymorphism? Or is this a textbook case of shell convergence? A molecular analysis confirms that differences indicated by radulae are profound, but it does not rule out that similarities in shell morphologies are the result of conserved patterns rather than the result of convergence of evolved patterns.

Material and methods

Material for this study was collected over the course of several years during expeditions organized by the Muséum National d'Histoire Naturelle, Paris (MNHN) to New Caledonia (Expédition Montrouzier, 1993), the Loyalty Islands (LIFOU 2000), the Philippines (PANGLAO 2004), and Vanuatu (SANTO 2006). Material for molecular studies was preserved in 90 or 100% ethanol by clipping pieces of the head-foot from anesthetized specimens.
Radulae were extracted from the specimens (from those with dried bodies – after rehydration), cleaned with diluted bleach, rinsed in distilled water, mounted on the stubs, air dried, coated with a gold-palladium and investigated with a JEOL JSM 840A Scanning Microscope. Terminology and homology of the central teeth in Turrinae follows Kantor (2006).
The following samples were examined in terms of radular morphology and also, where specifically mentioned, molecular sequences:

***Xenuroturris* identified as *"X. cingulifera"* based on shell characters**
Specimens with "duplex" radular type [= ***Xenuroturris legitima* Iredale, 1929**]
PHILIPPINES, Bohol Province**,** PANGLAO 2004, sta. G1, 09°41.9'N, 123°9,5'E, 100 m (2 spms – Fig. 6B-C). NEW CALEDONIA, Koumac, sta. 913, 20°58'S, 164°32'E, 10-13 m (1 spm – Fig. 6E), sta. 944, 20°35'S, 164°12'E, 14-15 m (1 spm – Fig. 6F). LOYALTY ISLANDS, LIFOU 2000, sta. 1465, 20°47.7'S, 167°07.0'E, 35-45 m (1 spm – Fig. 6D). ESPIRITU SANTO, VANUATU. SANTO 2006, sta. DR87, 15º38.5'S, 167º15.1'E, 13 m (1 specimen sequenced, MNHN 17684) (Fig. 1A, 6A).
Specimens with "semi-enrolled" radular type [= ***Iotyrris cingulifera* Lamarck, 1822**]
NEW CALEDONIA, Nouméa, sta. 255, 22°25'S, 166°20'E, 11 m (1 spm); Nouméa, Grand Récif Aboré, 5 m (1 spm); Touho, sta. 1240 (Expedition Montrouzier), 20°46.5'S, 164°14'E, 0-2 m (1 spm – Fig. 3G); Koumac, sta. 941, 20°39'S, 164°13'E, 15-16 m (1 spm – Fig. 3F), sta. 1303 (Expedition Montrouzier), 20°37.7'S, 164°15.9'E, 0-8 m (1 spm – Fig. 3A); sta. 1304 (Expedition Montrouzier), 20°38.6'S, 164°13.2'E, 12-15 m (1 spm); LAGON, sta. 549, 22°58'S, 166°56'E, 26 m (1 spm – Fig. 3H, 1B); Atoll de Surprise, sta. 445, 18°18'S, 163°02'E, 41 m (1 spm – Fig. 3E), LOYALTY ISLANDS, LIFOU 2000, sta. 1422, 20°47.1'S, 167°07.4'E, 4 m (1 spm) (Fig. – Fig. 3B), sta. 1423, 20°54.0'S, 167°07.3'E, 12 m (2 spms – Fig. 3C-D); sta. 1425, 20°46.8'S, 167°07.2'E, 4-5 m (1 spm). VANUATU, ESPIRITU SANTO. SANTO 2006, sta. FS84, 15º33.6'S, 167º16.6'E, 5 m (2 spms sequenced, MNHN 17685and 17686) (Fig. 3I).
Iotyrris devoizei n.sp.
VANUATU, ESPIRITU SANTO. SANTO 2006, sta. FB72, 15º36.1'S, 166º58.5'E, 16 m (1 specimen, sequenced, MNHN 20014) (Fig. 4 A-B); sta. DS04, 15º31.5'S, 167º14.5'E, depth unknown (2 specimens, sequenced, MNHN 17681 – Fig. 4D, MNHN 17682 – Fig 4E).
Iotyrris musivum n.sp.
VANUATU, ESPIRITU SANTO. SANTO 2006, sta. DS04, 15º31.5'S, 167º14.1'E, 25 m (2 specimens sequenced: MNHN 17683 – Fig. 5C, MNHN 20015 – Fig. 5A).
Xenuroturris kingae (Powell, 1964)
NEW CALEDONIA. Koumac, sta. 1319 (Expedition Montrouzier), 20º44.1'S, 164º15.1'E, 15-20 m (2 specimens – Fig. 5 H-I, 8E).

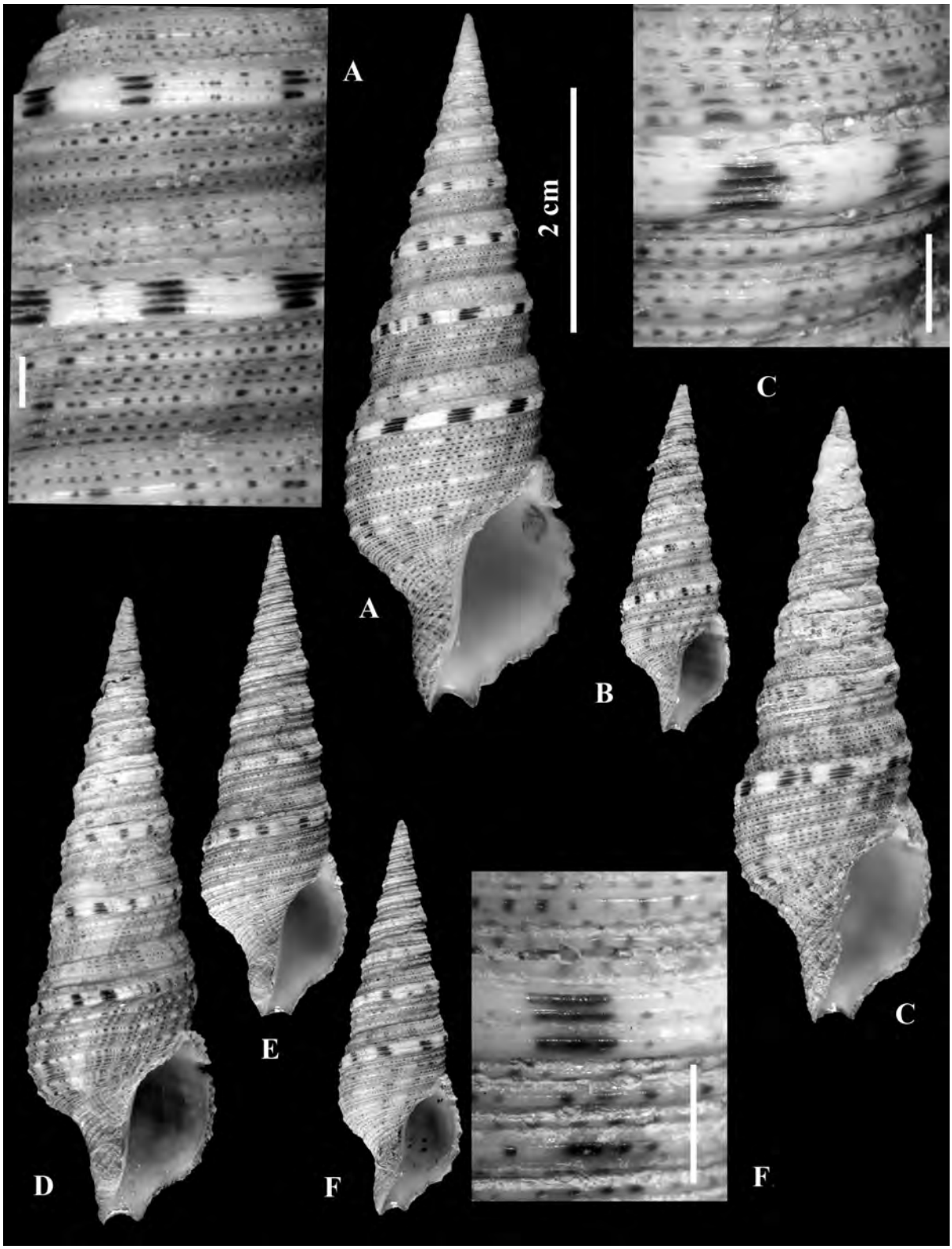


Fig. 1. *Xenuroturris legitima*. A – Vanuatu, SANTO 2006, sta. DR 087 (MNHN 17684), whole shell SL 57.0 mm, and close-up view of sculpture; B and C – Philippines, Panglao 2004, sta. S22, B – SL 28.4 mm, C – whole shell SL 50.5 mm, and close-up view of sculpture; D – Loyalty Islands, LIFOU 2000, sta. 1465, SL 51.5 mm; E – New Caledonia, Koumac, sta. 913, SL 39.5 mm; F – New Caledonia, Koumac, sta. 944, whole shell SL 32.5 mm, and close-up view of sculpture. All shells at the same scale, scale bars of close-ups 2 mm.

Fig. 2. Colour patterns of type or authenticated specimens of the species addressed in this study. A – *Xenuroturris legitima*, SL 57.0 mm, same specimen as Fig. 1A. B – *Iotyrris cingulifera*, SL 44.5 mm, same specimen as Fig. 3H. C – *Iotyrris devoizei*, SL 19.9 mm, same specimen as Fig. 4C. D – *Iotyrris musivum*, SL 24.5 mm, same specimen as Fig. 5D. E – *Xenuroturris kingae*, SL 18.0 mm, same specimen as Fig. 4I. F – *Iotyrris marquensensis*, holotype, SL 28.8 mm. G – *Xenuroturris cerithiformis*, holotype, SL 38.5 mm.

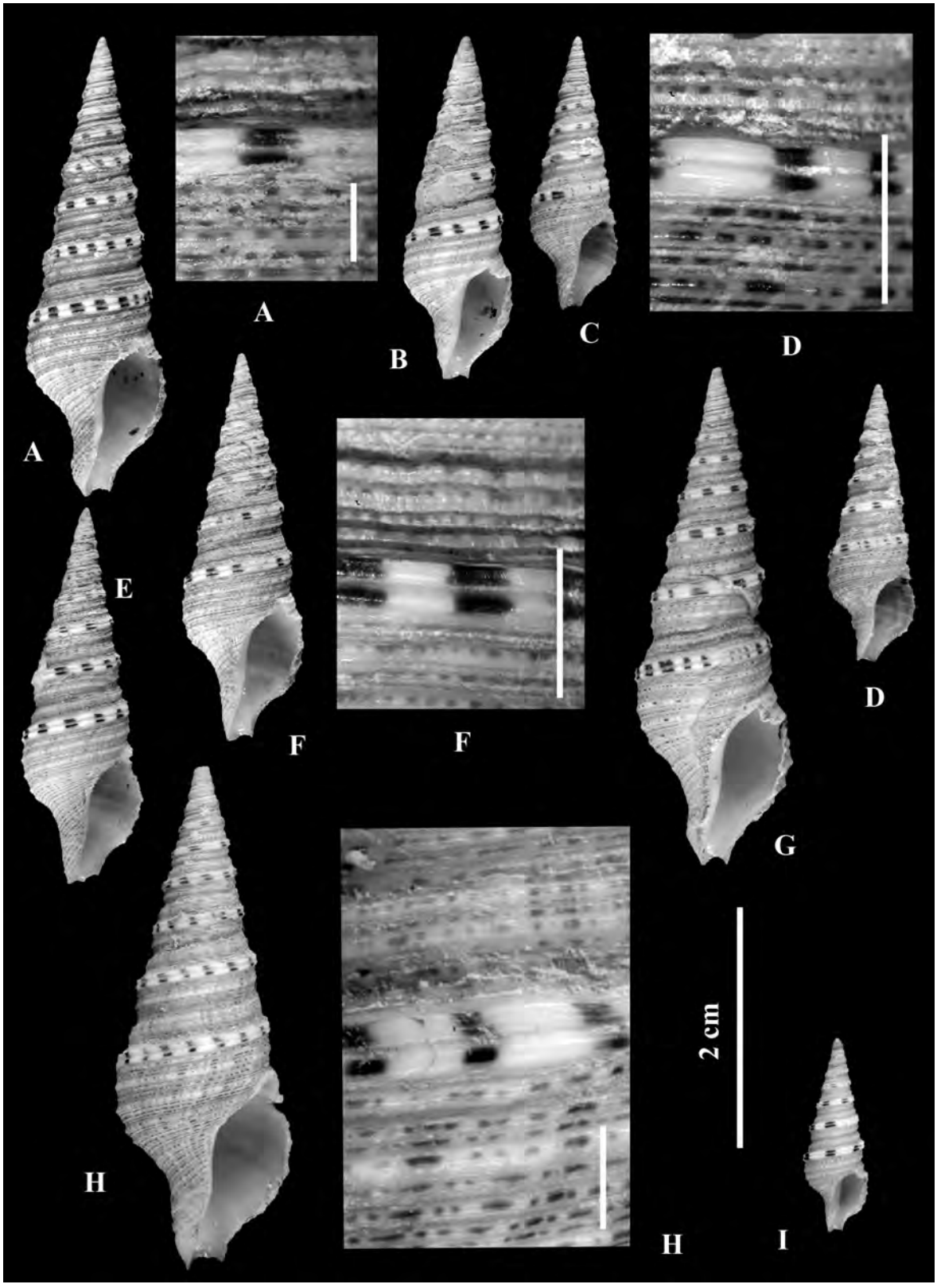


Fig. 3. *Iotyrris cingulifera.* A – New Caledonia, Koumac, sta. 1303, whole shell SL 38.4 mm, and close-up view of sculpture; B – Loyalty Islands, LIFOU 2000, sta. 1422, SL 28.4 mm; C-D – Loyalty Islands, LIFOU 2000, sta. 1423, C – SL 22.4 mm, D – whole shell SL 23.0 mm, and close-up view of sculpture; E – New Caledonia, Surprise Atoll, sta. 445, SL 31.2 mm; F – New Caledonia, Koumac, sta. 941, whole shell SL 32.5 mm, and close-up view of sculpture; G – New Caledonia, Touho, sta. 1240, SL 41.5 mm; H – New Caledonia, sta. 549, whole shell SL 44.5 mm, and close-up view of sculpture; I – Vanuatu, Santo 2006, sta. FS 084, SL 15.6 mm. All shells at the same scale, scale bars of close-ups 2 mm.

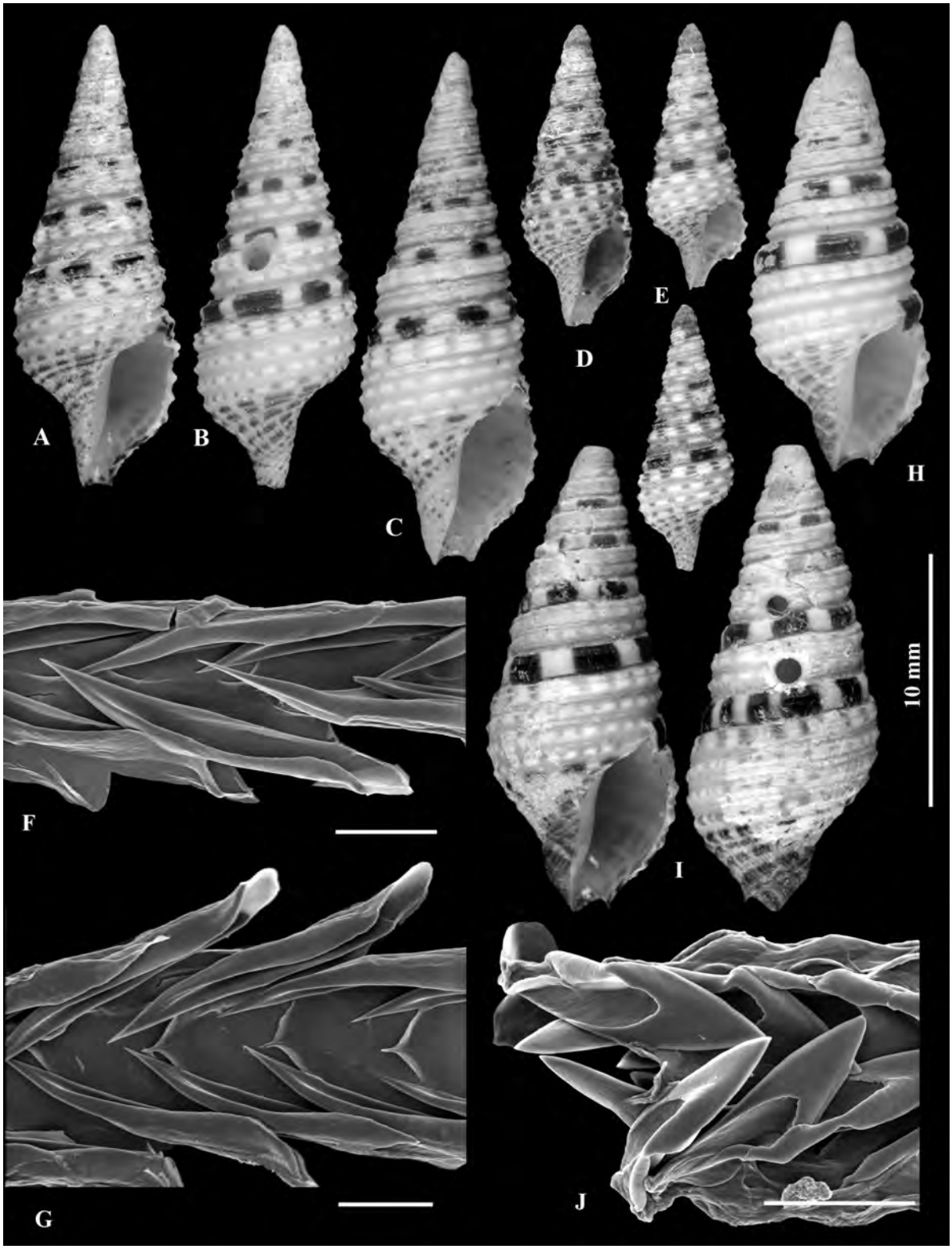


Fig. 4. A-G. *Iotyrris devoizei* n. sp. A-B – holotype, MNHN 20014 (the hole in the back results from drilling to extract the radula). C – Loyalty Islands, LIFOU 2000, sta. 1453, 21-30 m, SL 19.9 mm. D – paratype, MNHN 17681, SL 11.8 mm; E – apertural and dorsal views of paratype, MNHN 17682, SL 10.2 mm. F – radula of holotype. G – radula of the specimen in Fig. 4 C. H-J. *Xenuroturris kingae* Powell, 1964, New Caledonia, Koumac, 15-20 m. H – SL 17.4 mm. I – apertural and dorsal views, SL 18.0 mm. J – radula of the specimen in Fig. 4 I. All shells at the same scale. Scale bars for the radulae 50 µm.

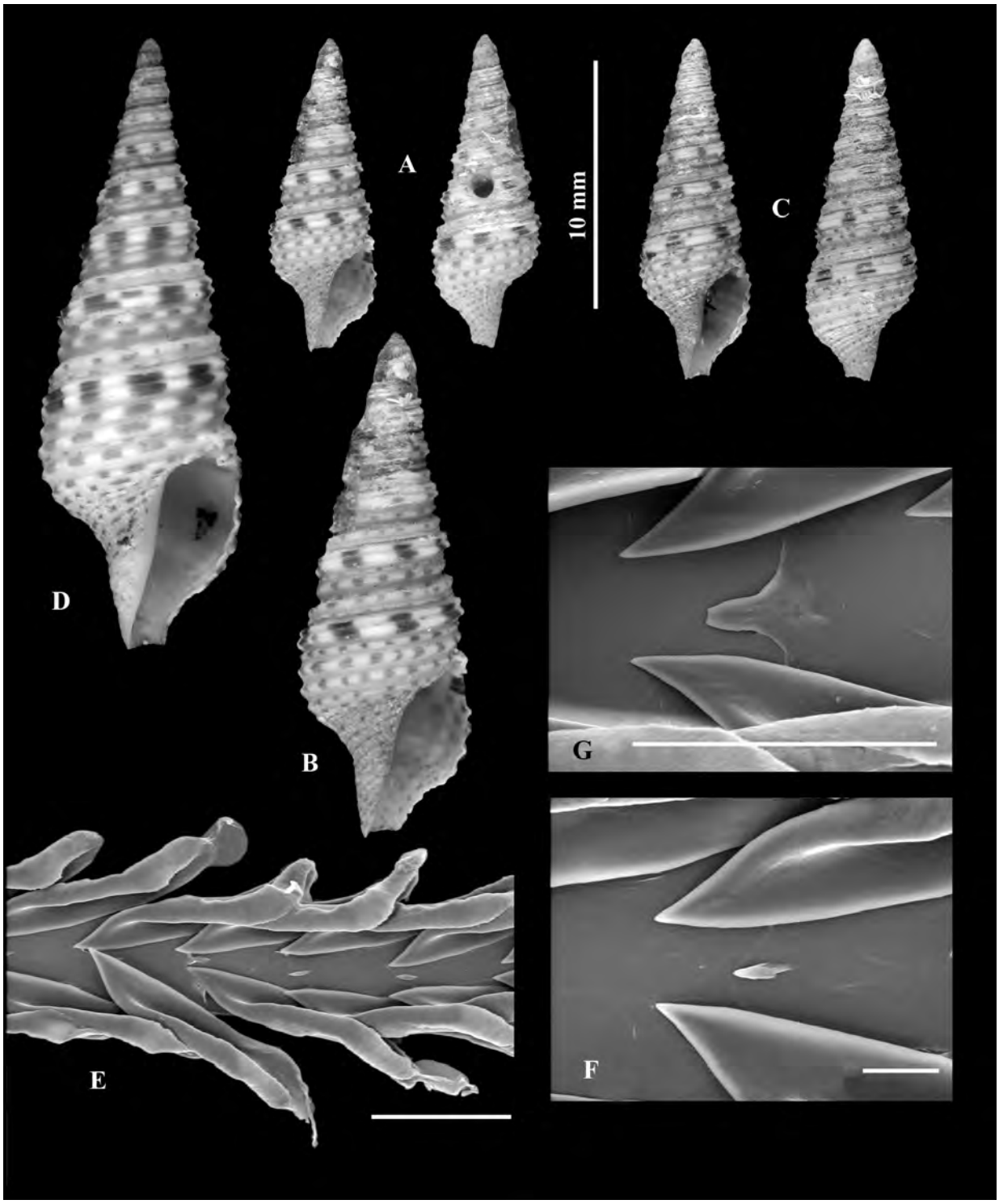


Fig. 5. *Iotyrris musivum* n. sp. A-B – holotype (B – enlarged), MNHN 20015, SL 12.6 mm (the hole in the back results from drilling to extract the radula). C – paratype, MNHN 17683, SL 13.9 mm. D – Loyalty Islands, LIFOU 2000, sta. 1421, 4 m, SL 24.5 mm. E-F – radula of holotype. G – central field of the radula of the specimen in Fig. 5 D. All shells (except B) at the same scale. Scale bars for E and G 50 µm, for F 10 µm.

Molecular study

Sequencing

DNA was extracted from a piece of foot, using 6100 Nucleic Acid Prepstation system (Applied Biosystem). A fragment of 658 bp of Cytochrome Oxidase I (COI) mitochondrial gene was amplified using the universal primers LCO1490 and HCO2198 developed by Folmer *et al.* (1994). One nuclear gene fragment was also analyzed: 900 bp of the rDNA 28S gene, involving D1, D2 and D3 domains (Hassouna *et al.*, 1984), using the primers C1 and D3 (Jovelin and Justine, 2001). All PCR reactions were performed in 25 µl, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq (Qbiogene) for all genes. For 28S gene, termocycles consisted of an initial denaturation step at 94°C for 4', followed by 30 cycles of denaturation at 94°C for 30'', annealing at 52°C and extension at 72°C for 1'. The final extension was at 72°C for 10'. Thermocycles used for COI gene are described in Hebert *et al.* (2003).

PCR products were purified and sequenced by the Genoscope, a French National Sequencing Center. Because of the length of the 28S PCR product, two internal primers (D2 and C2', Dayrat *et al.*, 2001) were used for sequencing, in addition of primers used for PCR. All genes were sequenced for both directions to confirm accuracy of each sequences. Sequences were deposited in GenBank. Genbank accession numbers for corresponding specimens are provided in Tables 1 and 2.

Table 1. Genetic pairwise distances for COI gene. The last line for each entry in first column represents the Genbank accession number.

	17683	20015	17682	17681	17684	20014	17686	17685	17756	17754
MNHN 17683 *Iotyrris musivum* (EU127874)										
MNHN 20015 *Iotyrris musivum* (EU127875)	0.000									
MNHN 17682 *Iotyrris devoizei* (EU127877)	0.132	0.132								
MNHN 17681*Iotyrris devoizei* (EU127876)	0.131	0.131	0.010							
MNHN 17684 *Xenuroturris legitima* (EU127878)	0.821	0.821	0.734	0.717						
MNHN 20014 *Iotyrris devoizei* (EU127879)	0.125	0.125	0.008	0.005	0.721					
MNHN 17686 *Iotyrris cingulifera* (EU127880)	0.080	0.080	0.125	0.163	0.551	0.149				
MNHN 17685 *Iotyrris cingulifera* (EU127881)	0.089	0.089	0.137	0.179	0.551	0.163	0.003			
MNHN 17756 *Lophiotoma albina* (EU127882)	0.259	0.259	0.328	0.281	0.553	0.284	0.329	0.329		
MNHN 17754 *Turris crispa* (EU015677)	0.526	0.526	0.710	0.755	0.678	0.762	0.517	0.563	0.425	
MNHN 17755 *Crassispira* sp. (EU015707)	1.557	1.557	2.024	1.876	2.939	1.835	2.027	2.027	1.419	1.811

Table 2. Genetic pairwise distances for 28S gene. The last line for each entry in first column represents the Genbank accession number.

	17683	20015	17682	17681	17684	20014	17686	17685	17756	17754
MNHN 17683 *Iotyrris musivum* (EU127874)										
MNHN 20015 *Iotyrris musivum* (EU127875)	0.000									
MNHN 17682 *Iotyrris devoizei* (EU127877)	0.002	0.002								
MNHN 17681*Iotyrris devoizei* (EU127876)	0.002	0.002	0.000							
MNHN 17684 *Xenuroturris legitima* (EU127878)	0.004	0.004	0.004	0.004						
MNHN 20014 *Iotyrris devoizei* (EU127879)	0.002	0.002	0.000	0.000	0.004					
MNHN 17686 *Iotyrris cingulifera* (EU127880)	0.005	0.005	0.005	0.005	0.006	0.005				
MNHN 17685 *Iotyrris cingulifera* (EU127881)	0.004	0.004	0.004	0.004	0.005	0.004	0.001			
MNHN 17756 *Lophiotoma albina* (EU127882)	0.011	0.011	0.011	0.011	0.016	0.011	0.015	0.016		
MNHN 17754 *Turris crispa* (EU015677)	0.019	0.019	0.019	0.019	0.024	0.019	0.027	0.024	0.011	
MNHN 17755 *Crassispira* sp. (EU015707)	0.050	0.050	0.050	0.050	0.061	0.050	0.067	0.061	0.053	0.056

Phylogenetic Analyses

Three species of Turridae (each represented by a single specimen) were used as outgroups for phylogenetic analyses: *Lophiotoma albina* (Lamarck, 1822) [Turrinae] (Vanuatu, SANTO 2006 expedition, sta. DR68, 15º22.9'S, 167º13.1'E, 7-27 m, MNHN 17756), as a closely related species (Olivera 2002); *Turris crispa* (Lamarck, 1816) [Turrinae] (Philippines, PANGLAO 2004 expedition, sta. R42, 9°37.1'N, 123°52.6'E, 8-22 m, MNHN 17754) and *Crassispira* sp. [Crassispirinae] (Philippines, PANGLAO 2004 expedition, sta. L46, 9°30.9'N, 123°41.2'E, 90-110 m, MNHN 17755).

COI gene sequence was manually aligned, and 28S was automatically aligned using ClustalW multiple alignment implemented in BioEdit version 7.0.5.3 (Hall, 1999), and the accuracy of the alignment was confirmed by eye. Genetic pairwise distances were calculated with PAUP 4.0b10 (Swofford, 2002), using the best fitting model of nucleotide substitution for each gene as defined by the program Modeltest (Posada & Crandall, 2001), in conjunction with PAUP 4.0b10, following the Akaike information criterion (AIC).

Phylogenetic reconstruction were conducted using Bayesian analysis, consisting of two Markov chains (2000000 generations each with a sampling frequency of one tree each hundred generations) run in six parallel analyses using Mr. Bayes (Huelsenbeck *et al.*, 2001). The same model of substitution used for computing the genetic distances was used for the Bayesian analysis. When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in.

Results

Radulae and Shells

Two morphologically distinct types of radulae were found in the species initially identified as "*Xenuroturris cingulifera*" based on shell characters.

The first radular type is characterized by the so-called "duplex" marginal teeth, typical for the family Turridae (*sensu* Taylor *et al.*, 1993) (Fig. 6A-C). The anterior part of the tooth (closer to the tip), constituting about 0.42-0.5 of total tooth length, is solid, while the posterior part has two thickened edges, a thinner dorsal one and thicker ventral one attached to the radular membrane (Fig. 6B). The central tooth is very small, narrow, and spear shaped. The lateral teeth absent. The ratio between marginal tooth length and shell length varies from 0.47 to 0.57 % (mean 0.52, σ=0.03, n=9), that between tooth length and aperture length varies from 1.67 to 2.05 % (mean 1.90, σ=0.14, n=9).

The second radular type is characterized by "semi-enrolled" marginal teeth (Fig. 6D-F). Both edges of the marginal tooth are elevated and equally developed along the entire tooth length, and they delimit a trough between them. The anterior solid part of the tooth is absent. Central and lateral teeth are absent. The marginal teeth are proportionally slightly longer than in the duplex type: the ratio between marginal tooth length and shell length varies from 0.54 to 0.67 % (mean 0.63, σ=0.05, n=8), that between tooth length and aperture length varies from 1.90 to 2.37 % (mean 2.21, σ=0.18, n=8).

The two forms occur sympatrically, although not syntopically, in Koumac, an intensively sampled site (Bouchet *et al.*, 2002) in New Caledonia; at Lifou in the Loyalty Islands; and on Espiritu Santo in Vanuatu. The shells of both forms are extremely similar and fall within the hitherto accepted (Powell, 1964; Olivera, 2002) range of intraspecific variation of "*Xenuroturris cingulifera*" (Figs. 1-3). However, after we had evidence of the two radular types, close examination of the shells revealed subtle differences.

(i) All specimens of "*Xenuroturris cingulifera*" have paired peripheral cords, the adapical one becoming subdivided into two by a groove with increasing shell height. In

specimens with "semi-enrolled" teeth this subdivision is slightly less obvious (Fig. 3 A, D, F, H) than in specimens with "duplex" teeth. The abapical, initially single and undivided, cord also becomes subdivided at certain shell size, but only in the form with "duplex" teeth. This is shown on the specimen on Fig. 1 A, in which the abapical cord is single on the adapical whorl and becomes subdivided on subsequent whorl. Large specimens of the form with "duplex" tooth thus have a peripheral zone with four nearly even thin cords (Fig. 1 A, C); specimens with the "duplex" marginal teeth tend to have slightly less convex peripheral cords.

(ii) Another difference is the shell base and the canal. Although observations are difficult to quantify and descriptions tend to be somewhat subjective, the form with "semi-enrolled" teeth has slightly more attenuated siphonal canal that is than in the "duplex" form.

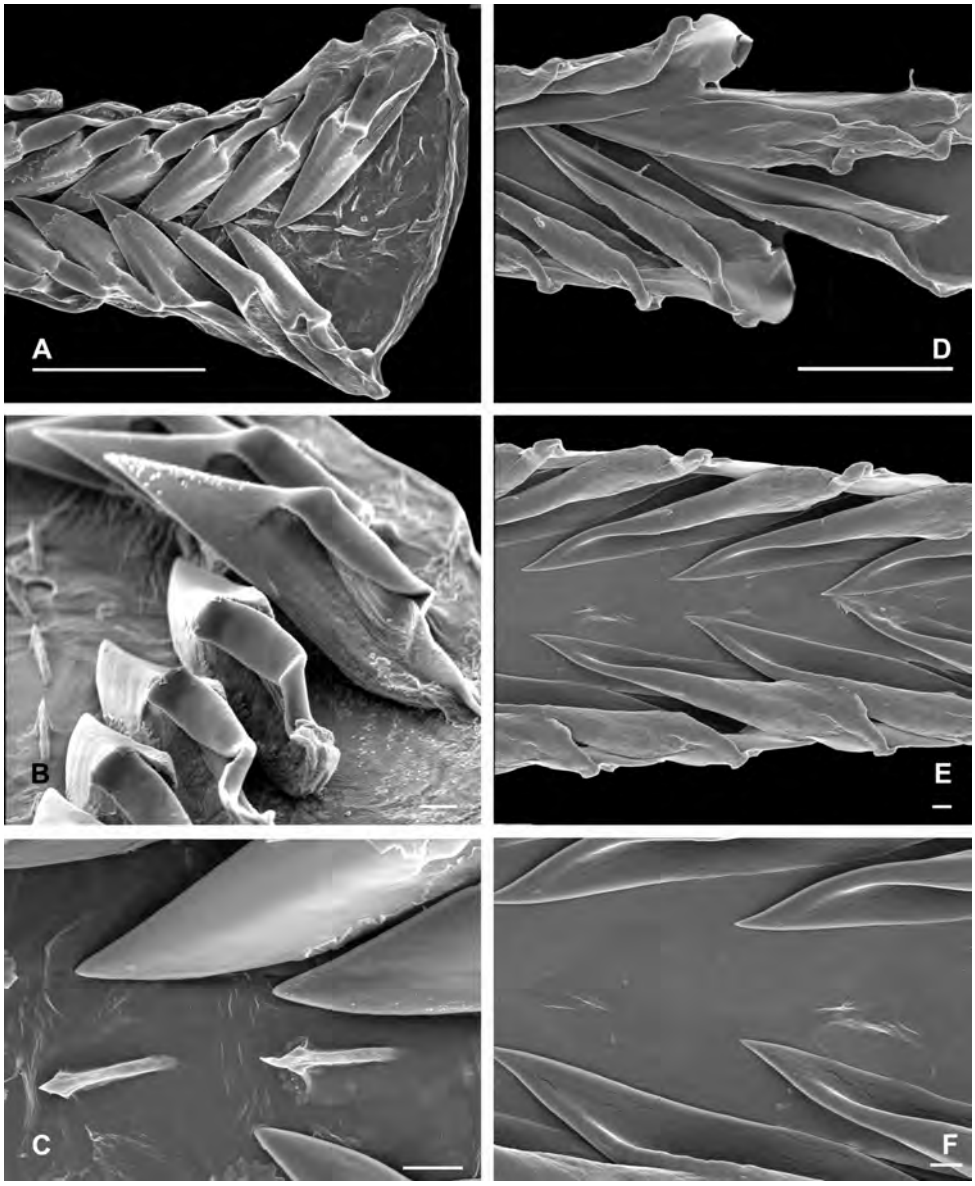


Fig. 6. Two radular types found in specimens originally identified as *"Xenuroturris cingulifera"*. A-C – radula with "duplex" marginal teeth, from specimen (PANGLAO 2004, sta. S22, SL 28.4 mm, shell see Fig. 1 B) now identified as *Xenuroturris legitima*. D-F – radula with "semi-enrolled" marginal teeth, from specimen (LIFOU 2000, sta. 1422, SL 28.4 mm, shell see Fig. 3 B) now identified as *Iotyrris cingulifera*. Scale bars A, D – 100 μm; B, C, E, F – 10 μm.

Both *Iotyrris devoizei* n. sp. and *I. musivum* n. sp. also have radulae with "semi-enrolled" teeth. In *I. devoizei* radular teeth are very similar to the form of "*X. cingulifera"* (Fig. 4 F-G) with the "semi-enrolled" teeth [= true *I. cingulifera*]. The central and lateral teeth were absent in one specimen, but rather well developed in another (Fig. 4 G). The marginal teeth are nearly twice as long as in *I. cingulifera*: the ratio between marginal tooth length and shell length varies from 0.94 to 1.15 % (mean 1.02, σ=0.09, n=5), that between tooth length and aperture length varies from 2.96 to 3.54 % (mean 3.18, σ=0.24, n=5). In *Iotyrris musivum* the shape of marginal teeth is also similar. The central tooth is either very small and sharp (Fig. 5 E-F) or is obtuse and fused with vestigial laterals (Fig. 5G). The marginal teeth are of intermediate length: the ratio between marginal tooth length and shell length varies from 0.75 to 0.87 % (mean 0.79, σ=0.05, n=5), that between tooth length and aperture length varies from 2.37 to 2.76 % (mean 2.58, σ=0.17, n=5).

Molecular study

For COI and 28S genes, alignments resulted in sequences of 658 and 903 bp length respectively. No indel was found for COI. Within ingroups, 97 sites were variable for COI gene, and only 14 for 28S. The Tamura-Nei model (TrN+G, with G = 0.1047) model for the COI gene and the General Time Reversible model (GTR+I, with I = 0.8579) model for the 28S gene were defined as the best fitting models.

Genetic distances are low (0.003, table 1) between specimens specimens of "*Xenuroturris cingulifera*" with "semi-enrolled" teeth [= true *Iotyrris cingulifera*], and are similar to the genetic distances found between specimens of *I. devoizei*, or of *I. musivum* (from 0 to 0.010). On the contrary, genetic distances found between specimens of "*X. cingulifera*" with "semi-enrolled" teeth and the specimen of "*X. cingulifera*" with "duplex" teeth [= *X. legitima*] are very high (0.551), in fact higher than the genetic distances between specimens of *I. devoizei* and *I. musivum* (0.131).

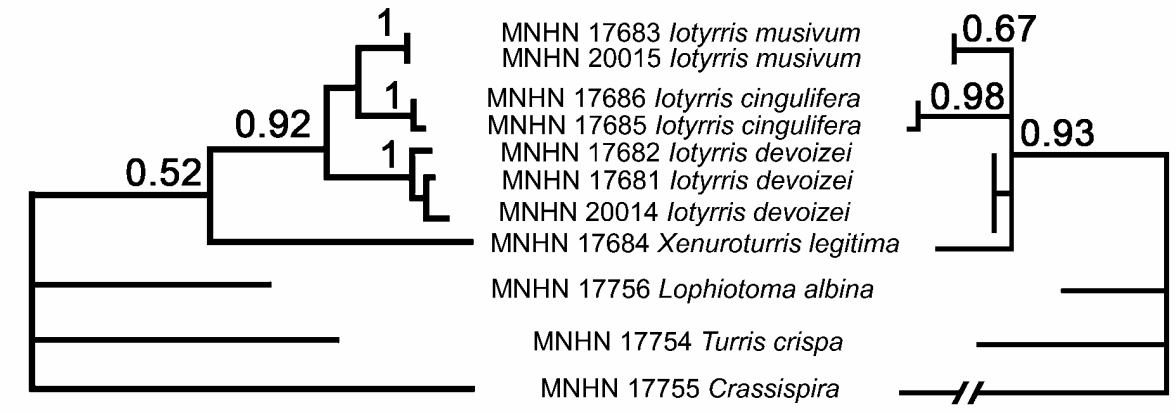


Fig. 7. Trees obtained with COI gene (on the left) and 28S gene (on the right), using Bayesian approach. Posterior probabilities are indicated above nodes.

All specimens of *Xenuroturris-Iotyrris* clustered together for COI, but this result is not supported (Posterior Probability PP = 0.52, Fig. 7). Specimens with a "semi-enrolled" teeth are found to form a monophyletic clade (although weakly supported) that exludes the specimen with the "duplex" teeth. Within the "semi-enrolled" clade, three clades are recognized, each containing two or three specimens. Those three clades each included only specimens hypothesized to be conspecific based on shell and radula characters. Consequently,

on the molecular tree, specimens of "*X. cingulifera"* with "semi-enrolled" teeth appear more closely related to *Iotyrris devoizei* and *I. musivum* than to the specimen of "*X. cingulifera*" with "duplex" teeth.

Similar results are found for 28S gene, with low genetic distances between specimens with "semi-enrolled teeth" (0.001, table 2) and higher genetic distances (0.005, table 2) between specimens of "*X. cingulifera*" with "semi-enrolled" teeth [true *Iotyrris cingulifera*] and "duplex" teeth [= *X. legitima*]. Monophyly of the *Xenuroturris-Iotyrris* clade is weakly supported (0.93). Within "semi-enrolled" clade, the three clades defined with the COI gene are found again, but only one of them is strongly supported. Relationships between those three clades and the specimen of "*X. cingulifera*" with "duplex" teeth [= *X. legitima*] are not resolved.

Discussion

Taxonomic rank of the different forms

The genetic distances between specimens of "*Xenuroturris cingulifera*" and the phylogenetic relationships between those specimens provide two strong arguments that confirm the hypothesis based on the observation of the radulae:

(a) with both COI and 28S gene, genetic distances between specimens displaying different radulae are very high, corresponding at least to distances found between specimens of different species,

(b) with COI gene, specimens of "*X. cingulifera*" with "semi-enrolled" teeth are found to be more closely related to other species of *Iotyrris* than to the specimens of "*X. cingulifera*" with duplex teeth. As specific status of *I. devoizei* and *I. musivum* is not questioned, two species must be recognized within "*X. cingulifera*".

The species level status of the two other species, *Iotyrris devoizei* and *I. musivum* also leaves no doubt. All four species under consideration were found sympatrically in Santo, and *I. devoizei* and *I. musivum* were even found syntopically in a single underwater vacuum-cleaning collecting effort. *I. devoizei* and *I. musivum* are described as new species.

Taxonomic and nomenclatural implications

Although the species considered in this paper are not unambiguously recognizable from their shells (compare, e. g., Fig. 1 F and 3 E, or 1 B and 3 C), we examined the type material of the nominal species currently subsumed in "*Xenuroturris cingulifera*", to ascertain which names, if any, might be applicable. Two available names are applicable to two of our species, the other two are described as new.

The type material of *Xenuroturris legitima* Iredale, 1929 has a short recurved siphonal canal and both peripheral cords are subdivided (Fig. 8 C, E). The name is thus applicable to the form with "duplex" radular teeth.

The syntypes of *Pleurotoma cingulifera* Lamarck, 1822 are badly worn specimens that render their identification difficult. The only significative character is in the peripheral cords, of which the abapical one is not subdivided, even in the largest syntype (shell length 65 mm). In our specimens of similar size with "duplex" radular teeth, that cord is always subdivided. Although some ambiguity persists, we apply the name *cingulifera* to the form with "semi-enrolled" marginal radular teeth. We are of the opinion that this is nomenclaturally more stable than leaving *cingulifera* as a *nomen dubium* and describing the form with "semi-enrolled" marginal radular teeth as a new species.

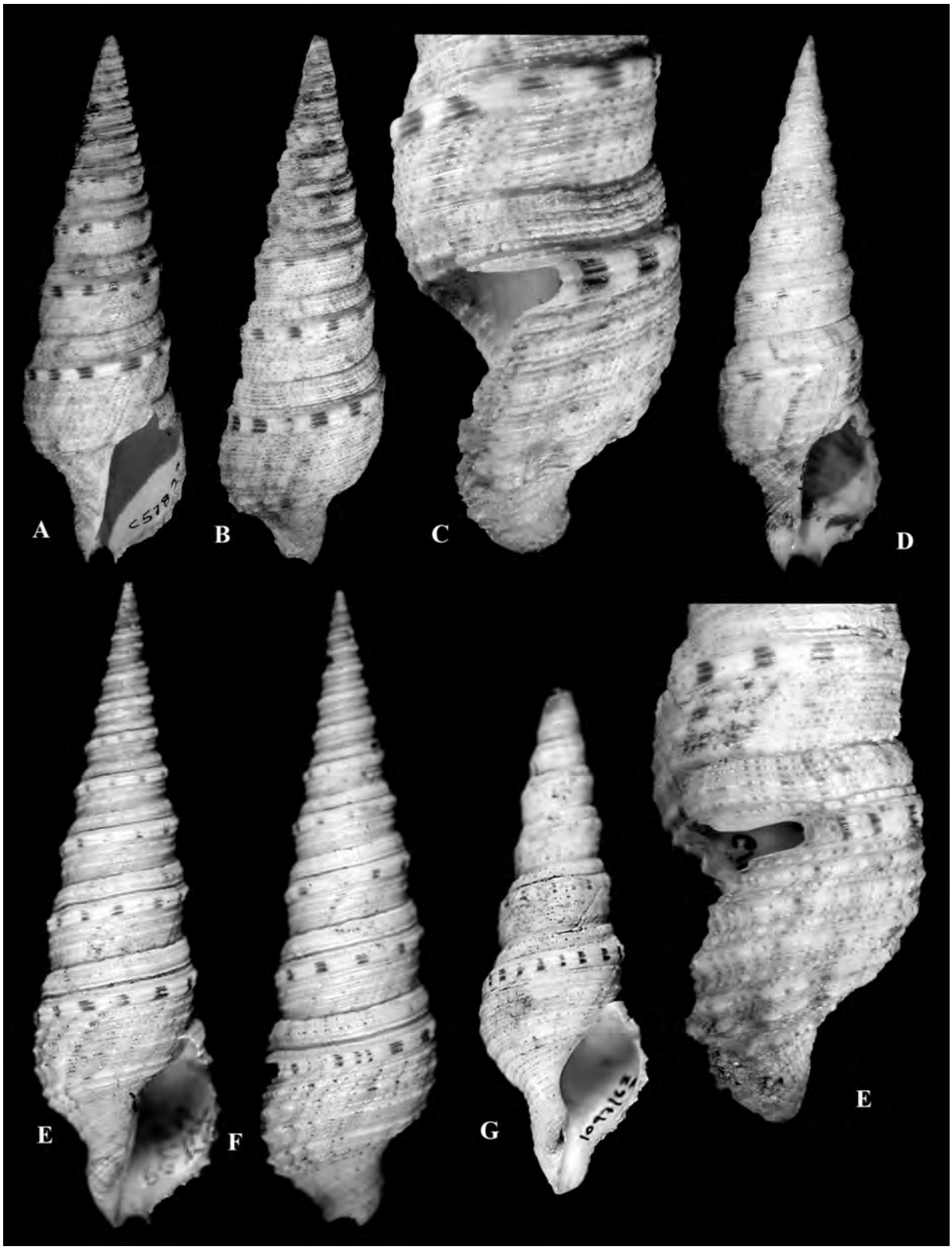


Fig. 8. A-E – syntypes of *Xenuroturris legitima* Iredale, 1929: A-C – AMS C57823, SL 55.7 mm; D-E – AMS C220457, SL 72 mm. E-G – syntypes of *Pleurotoma cingulifera* Lamarck, 1822: E-F – MHNG 1097/62/1, SL 65.0 mm, G – MHNG 1097/62/3, SL 54.0 mm. A-E, photos by A.C. Miller, copyright: Australian Museum. E-G, photos courtesy of Y. Finet, MHNG.

Ideally, the current name-bearing types of *Pleurotoma cingulifera* Lamarck, 1822 and *Xenuroturris legitima* Iredale, 1929 should be replaced by live-taken neotypes with known radular and molecular characteristics. This can be done only by a decision of the International Commission on Zoological Nomenclature, but we are of the opinion that this is unnecessary as long as the systematic and nomenclatural conclusions of this paper are accepted by zoologists.

The type species of *Xenuroturris*, by monotypy, is *X. legitima* Iredale, 1929, and this name is thus applicable to *Xenuroturris*-type shells with "duplex" teeth. Medinskaya and Sysoev (2001) proposed the monotypic new genus *Iotyrris* for what they had identified as *Xenuroturris cerithiformis* Powell, 1964 (misidentified; = *Iotyrris marquesensis* Sysoev, 2002), a species very similar to other species of *Xenuroturris* in shell characters and possessing a radula with "semi-enrolled" teeth. Although we did not include the type species of *Iotyrris* in our analysis, we suggest that the name *Iotyrris* is applicable to these turrines with *Xenuroturris*-type shells and "semi-enrolled" radular teeth.

Xenuroturris Iredale, 1929

Iredale, 1929: 283.
Type species (by monotypy): *Xenuroturris legitima* Iredale, 1929

Xenuroturris legitima Iredale, 1929

Iredale, 1929: 283, pl. 31, figs. 3-4.
Selected synonymy: *Xenuroturris cingulifera cingulifera* – Powell, 1964 (part.): plate 175, figs. 19-20, non fig. 12. *Xenuroturris cingulifera* – Medinskaya and Sysoev, 2002: 9, fig. 3?, fig. 6A. *Lophiotoma cingulifera* – Olivera, 2002, figs. 1 F, 1 G, H1, H2.
Syntypes: Two type specimens of *Xenuroturris legitima* housed in the Australian Museum are labeled as holotype (C.57823) and paratypes (C.110457). However, in the original description, Iredale did not fix a holotype and therefore all type specimens should be considered syntypes.
Type locality: Michaelmas Cay, Queensland, Australia.
Distribution: The distribution ascertained by us includes New Caledonia, the Loyalty Islands, Vanuatu, and the Philippines. The species is also recorded from South Africa (Kilburn 1983: 564-565, figs 30, 32) and Japan (Okutani 2000: 631, pl. 314, fig. 57), but this now require confirmation based on radula and/or molecular characters.

Iotyrris Medinskaya et Sysoev, 2001

Medinskaya and Sysoev, 2001: 12.
Type species: *Iotyrris marquesensis* Sysoev, 2002 (Fig. 2F), by fixation by Sysoev (2002) under Art. 70.3 of the Code of Nomenclature.

Iotyrris cingulifera (Lamarck, 1822)

Pleurotoma cingulifera Lamarck, 1822: 94.
Selected synonymy: *Xenuroturris cingulifera cingulifera* – Powell, 1964 (part.): plate 175, fig. 12, non figs. 19-20. *Lophiotoma cingulifera* – Olivera, 2002 (part.), fig. H3.

Syntypes: 3 possible syntypes of *Pleurotoma cingulifera* are housed in the Muséum d'histoire naturelle de la ville de Genève (MHNG 1097/62) (Fig. 8 E-G). The largest specimen, possibly illustrated by Kiener (1839), has a shell length of 65 mm (MHNG 1097/62/1), which more or less matches Lamarck's indication of a size of "2 pouces 4 lignes", i.e. 63 mm.
Type locality: not stated. Powell (1964) restricted it to Mauritius, a restriction that has no standing under the ICZN.
Distribution: The distribution ascertained by us includes New Caledonia, the Loyalty Islands, Vanuatu, and the Philippines. Records from Zanzibar (Powell 1964) and Guam (Olivera (2002: fig. H3) require confirmation based on radula and/or molecular characters.

Iotyrris devoizei n. sp.

(Fig. 4 A-G)

TYPE MATERIAL. Holotype MNHN 20014 and 2 paratypes MNHN 17681, 17682.
TYPE LOCALITY. Vanuatu, south of Espiritu Santo, Elia Island, 15º36.1'S, 166º58.5'E, 16 m [SANTO 2006 expedition, sta. FB72].
ETYMOLOGY. The species is named after Patrice Petit-Devoize, biologist and diver, and a pillar of the MNHN expeditions diving team.
OTHER MATERIAL EXAMINED. Vanuatu, Espiritu Santo, SANTO 2006 expedition, sta. DS04, wreck of *President Coolidge*, 15º31.5'S, 167º14.1'E, 25 m (2 lv, paratypes).
New Caledonia, LAGON sta. 888, 20º22'S, 164º38'E, 20 m, 1 dd; Expedition Montrouzier, sta. 1269, 20º35.1'S, 165º08.1'E, 15-20 m, 1 dd; sta. 1271, 20º52.7'S, 165º19.5'E, 5-25 m, 4 dd; sta. 1311, 20º40.4'S, 164º14.9'E, 10-60 m, 1 dd; sta. 1318, 20º41.4'S, 164º14.8'E, 20-30 m, 1 dd. Loyalty Islands, LIFOU 2000 expedition, sta. 1421, 20º52.4'S, 167º08.5'E, 4 m, 2 dd; sta. 1422, 20º47.1'S, 167º07.4'E, 4 m, 2 dd; sta. 1423, 20º54.0'S, 167º07.3'E, 12 m, 1 dd; sta. 1429, 20º47.5'S, 167º07.1'E, 8-18 m, 4 dd; sta. 1435, 20º55.2'S, 167º00.7'E, 5-30 m, 2 dd; sta. 1442, 20º46.4'S, 167º02.0'E, 47 m, 1 dd; sta. 1448, 20º45.8'S, 167º01.7'E, 20 m, sta. 1450, 20º45.8'S, 167º01.7'E, 27-31 m, 1 dd; sta. 1451, 20º47.3'S, 167º06.8'E, 10-21 m, 3 dd; sta. 1453, 20º54.6'S, 167º02.1'E, 21-30 m, 3 dd, 1 lv (Fig. 2C, 4C); sta. 1455, 20º56.8'S, 167º02.7'E, 15-20 m, 2 dd; sta. 1457, 20º46.8'S, 167º02.8'E, 5-10 m, 3 dd.
DISTRIBUTION. New Caledonia, Loyalty Islands, Vanuatu, shallow water, alive in 16 to 30 m, shells from 4 to 47 m. It is possible that the species also occurs in Hawaii and the Marquesas.
DESCRIPTION (holotype). Shell conical, consisting of 7 teleoconch whorls, with high spire, shell diameter to shell height 0.35, aperture height to shell height 0.33, spire height to shell height 0.49. Protoconch overgrown and eroded, light brown, of more than 3 whorls, sculptured with closely spaced axial threads. Teleoconch whorls slightly convex, spire whorls are slightly angulated on the periphery. Last whorl sharply narrowing towards attenuated, nearly straight, siphonal canal. Suture shallowly adpressed, slightly wavy. Subsutural ramp narrow, convex, with single bulging keel subdivided by two narrow and shallow grooves. The keel is strongly marked by rectangular, unevenly sized, dark red-brown spots. Immediately below the ramp are two narrow but tall, closely spaced, cords, forming the shoulder. Below them are three broadly spaced cords, equal to those on the shoulder, separated by smooth interspaces, except one narrow thread between the lower shoulder cord and upper periphery one. There are 5 more closely spaced cords on the shell base and 5 narrower and closer spaced threads on the canal. Axial sculpture absent except for inconspicuous growth lines. Aperture narowly oval. Outer lip broken, judging from the growth lines anal sinus deep, situated on the shoulder. Inner lip nearly straight, columellar part straight, callus very narrow not extending on the parietal wall. Color creamy white, with dark strong spots on the subsutural ramp, much

ligher yellow narrow spots, covering all the cords of the periphery, tessellate dark closely spaced spots on the cords of the shell base, smaller and slightly ligher spots on the threads of the canal.
Shell height (holotype) 17.8 mm, shell diameter 6.3 mm, last whorl height 9.0 mm, aperture height 5.9 mm.
REMARKS. Paratypes are similar to the holotype in shell outline and sculpture, and appear to be juveniles. However, one of them (MNHN 17681, Fig. 4 D), with a shell length of only 11.8 mm, is already a mature male with a long penis equal in length to that of the mantle cavity. That specimen has more pronounced color spots on the shoulder and shell periphery, so that the contrast between the shell base and periphery is much less pronounced. The protoconch of the second paratype (Fig. 4 E) is best preserved, but its nucleus missing; the number of whorls is thus at least 4.
The largest specimen we attribute to *I. devoizei* has a shell length of 27.1 mm.
Iotyrris devoizei is conchologically very similar to *X. kingae* (Powell, 1964) in terms of sculpture and colour pattern, down to the finest details. The only differences between the two is the more elongate shell of *I. devoizei*, with a clearly longer siphonal canal, and the slightly less bulging subsutural rim. The two species have a radula of a different type – while *I. devoizei* possesses semi-enrolled marginal radular teeth, *X. kingae* has duplex teeth (Fig. 4 J).

Iotyrris musivum n. sp.

(Fig. 5)

TYPE MATERIAL. Holotype MNHN 20015 and 1 paratype MNHN 17683.
TYPE LOCALITY. Vanuatu, Espiritu Santo, wreck of *President Coolidge*, 15º31.5'S, 167º14.1'E, 25 m [SANTO 2006 expedition, sta. DS04].
ETYMOLOGY. *Musivum*, Latin meaning mosaic, with reference to the colour pattern; used as a noun in opposition.
OTHER MATERIAL EXAMINED. New Caledonia, Expedition Montrouzier, Koumac, sta. 1286, 20º38'-20º39'S, 164º16'-164º17'E, intertidal, 1 lv. Loyalty Islands, LIFOU 2000 expedition, sta. 1421, 20º52.4'S, 167º08.5'E, 4 m, 1 lv (Fig. 2D, 5D); sta. 1424, 20º54.9'S, 167º03.0'E, 4 m.
DISTRIBUTION. New Caledonia, Loyalty Islands, Vanuatu, shallow water, intertidal to 16 m.
DESCRIPTION (holotype). Shell conical, consisting of 5.75 teleoconch whorls, with high spire, shell diameter to shell height 0.34, aperture height to shell height 0.31, spire height to shell height 0.48. Protoconch brown, consists of 4.25 convex whorls, sculptured with closely spaced axial threads, which are slightly prosocline on most whorls, but turned to strongly opisthocline on the posteriormost part of the last protoconch whorl. There are around 25 threads on the last protoconch whorl. Teleoconch whorls nearly flat, slightly angulated at the shoulder. Last whorl sharply narrowing towards attenuated nearly straight siphonal canal. Suture shallowly impressed, slightly wavy. Subsutural ramp narrow, with single strong sharp keel, followed by narrow thread. Immediately below the ramp there are are two narrow but tall triangular in section closely spaced cords, forming the shoulder. Below there are four broadly spaced subequal cords, separated by smooth interspaces. There are 5 thinner and more closely spaced cords on the shell base and 8 narrower and closer spaced threads on the canal. Axial sculpture absent except for inconspicuous growth lines. Aperture narrowly oval. Outer lip thin, evenly rounded. Anal sinus deep, situated on the shoulder. Inner lip slightly convex, columellar part straight, callus very narrow not extending on the parietal wall. Color creamy white, with regularly spaced light brown spots, covering cords of the entire shell

surface and much more pronounced and more dark strong subrectangular spots on the shoulder cords.
Operculum yellow, transparent, leaf-shaped, with terminal nucleus.
Shell height (holotype) 12.6 mm, shell diameter 4.3 mm, last whorl height 6.6 mm, aperture height 3.9 mm.
REMARKS. The holotype is a juvenile specimen. We nevertheless chose to fix it as the name-bearing type, since it is clearly recognizable, because we know its radula and DNA sequence. The paratype (Fig. 5C), which is very similar to the holotype in shell outline and sculpture, has a length of 13.9 mm. The largest specimen has a shell length of 29.1 mm. In large specimens the subsutural ramp has additional thin but sharp spiral thread immediately below the suture. There are additional thin threads between the major cords on the shell periphery.
Iotyrris musivum resembles *Iotyrris marquesensis* Sysoev, 2002 (which also has "semi-enrolled" marginal radular teeth) but differs in having a broader shell with relatively shorter canal. Besides, the color marks on the shoulder keels are much less pronounced in *I. marquesensis* and do not differ from the spots on the rest of the shell. The subsutural ramp in *I. marquesensis* is separated from the shoulder keels by a deep cleft, while in *X. sp.* only by narrow and shallow groove. From small specimens of *I. cingulifera* the new species differs in much larger and color spots of the shell, brighter color spots on the shoulder cords and more coarse spiral sculpture. Another species which is remotely similar to *I. musivum* is *Lophiotoma olangoensis* Olivera, 2002. The new species differs in having a much more brightly colored subsutural rim and nearly straight canal, which in *L. olangoensis* is clearly curved to the left in apertural view. We examined the radula of a specimen from Lifou (Loyalty Islands) which we identified as *L. olangoensis* and it appeared to be similar to that of *Xenuroturris legitima*, that is consisting of duplex marginal teeth, very small and narrow central one and vestigial laterals.

Radula as a taxonomic character in "Turridae"
Since McLean's (1971) classification of the Turridae *s.l.*, the radula has routinely been used for subfamily and genus level allocation within Conoidea. However, relatively little correlation between foregut anatomy and shell and radular characters was found in many groups (e. g., Kantor *et al.*, 1997; Kantor and Taylor, 2002). For instance, several genera of Crassispirinae (Turridae) having very similar shells (*Inquisitor* Hedley, 1918, *Funa* Kilburn, 1988, and *Ptychobela* Thiele, 1925) differ significantly in radular teeth shape. In practice, similarity in shell together with dissimilarity in radula was interpreted as shell homeomorphy. For instance, Bouchet and Sysoev (2001) established the new genus *Leiosyrinx* (Conidae, Raphitominae), conchologically very similar to *Typhlosyrinx* Thiele, 1925, based on differences in their radulae. The subfamily Zonulispirinae McLean, 1971 (Turridae) was proposed mainly on the basis of its radula with "semi-enrolled" marginal teeth only, while in shell characters it is rather similar to the subfamily Crassispirinae (Turridae). However, the hypothesis that the radula better reflects relationships than the shell had never been tested by molecular data.
Kantor and Taylor (2000) showed that very different marginal radular teeth in Turridae *s.s.* are formed in a similar way by thickening of the tooth edges, longitudinal folding and progressive elevation of the posterior edge of the tooth from the membrane in the course of tooth maturation. They illustrated this process for what they called *Xenuroturris cingulifera* with "duplex" teeth [= *X. legitima*] (Kantor and Taylor, 2000: fig. 3 g-h). Comparison of the process of maturation in both teeth types revealed that, at some point, the semi-matured teeth are rather similar in shape, and the final differences are the result of extra sclerotization of the duplex teeth that is responsible for the long anterior solid tip. The difference between

"duplex" and "semi-enrolled" teeth might therefore be less profound than it seems, and its phylogenetic value was only speculative. An important result of this study is thus that species such as *Iotyrris cingulifera*, *I. devoizei* and *I. musivum*, with rather different shells but with similar radulae (in this case, "semi-enrolled" marginal teeth) form a clade, thus confirming that radular type is indeed reliable for revealing relationships.
The radulae of two further species with *Xenuroturris*-type shells were examined from animals with dried soft parts, and reveal further species pairs with "duplex" and "semi-enrolled" teeth. One case is *Xenuroturris kingae* (Powell, 1964) [with "duplex" teeth] and *Iotyrris devoizei* [with "semi-enrolled" marginal teeth] (Fig. 4). Both are conchologically very similar in terms of sculpture and even identical in the finest details of coloration, the only significant difference being a more elongate shell, with distinctly longer siphonal canal, in *X. devoizei*. Such differences were earlier considered to be of intraspecific value. Thus, pending confirmation, we think that Olivera's (2002) figs 2B1 and 2B4 represent *Iotyrris devoizei*, while figs 2B2-3 and 2B5 represent *X. kingae*. The two species co-occur in Hawaii and the Marquesas (based on Olivera's illustrations) and in New Caledonia (this paper). Another case is *Iotyrris marquesensis* Sysoev, 2002 [with "semi-enrolled" marginal teeth] (Fig. 2F) and *Xenuroturris cerithiformis* Powell, 1964 [with "duplex" teeth] (Fig. 2G), which differ so little in shell shape that Medinskaya and Sysoev (2001) initially confused them.
It is remarkable that within the subfamily Turrinae a "semi-enrolled" radular tooth is found only in this clade, and it raises the question of whether they originated once in the genus - and *Iotyrris* is indeed monophyletic, and conchologically similar species with different radular types are the result of convergence or homoplasy -, or they appeared several times - and *Iotyrris* is polyphyletic and conchologically similar species with different radular types are indeed sister taxa. Regrettably, our study includes only one case where both members of the pair - *X. legitima* and *I. cingulifera* - were included in the analysis, and they ended up being not closely related in the molecular tree. Thus, although the molecular analysis confirms that differences indicated by radulae are profound, it does not rule out that similarities in shell morphologies are the result of conserved patterns rather than the result of convergence of evolved patterns.
Currently, *Xenuroturris* is considered a mere subgenus of *Lophiotoma* Casey, 1904 (Kilburn, 1983; Taylor *et al.*, 1993; Higo *et al.*, 1993), the main difference being the truncated base of *Xenuroturris*. Olivera (2002) even expressed serious doubts that *Xenuroturris* should be recognized as a valid subgenus, pointing out the strong similarities in shell sculpture between *Lophiotoma albina* (Lamarck, 1822) and "*X. cingulifera*". Our study included a single species conchologically attributable to *Lophiotoma*, and is not relevant on the phylogenetic relationships between the two genera.
In conclusion, and despite these reservations, our working hypothesis is that *Lophiotoma*, *Xenuroturris* and *Iotyrris* are all valid genera, a view that could not have been predicted from shells alone. Currently, most "turrid" genera are defined purely based on shell characters and there are anatomical or radula data for fewer than 10% of the described species. As long as molecular data exist only for a desperately small proportion of that diversity, we must be prepared that many of the currently accepted taxa are poly- or paraphyletic, at all taxonomic levels: specific, generic, and subfamilial.

Acknowledgements

The present research was carried out during a visiting curatorship of YK in MNHN and the senior author wants to express his thanks to Philippe Maestrati, Virginie Héros, and especially Pierre Lozouet, for expertise and assistance during his stay in Paris. Key material for

molecular studies originates from the MNHN-IRD-PNI Santo 2006 expedition, and we thank Patrice Petit de Voize, Laurent Albenga and Jacques Pelorce for collecting the specimens during daily dives during 6 weeks of sustained bad weather. Preliminary sorting of the material to morphospecies was accomplished by Alexander Sysoev and Richard Kilburn, as part of the New Caledonia "turrid factory" project. We express our thanks to Dr. Jason Biggs who shared with us unpublished information and photographs of different species of *Xenuroturris*, and to two anonymous reviewers for the valuable comments and suggestions. The Total Foundation supported the LIFOU 2000, PANGLAO 2004 and SANTO 2006 expeditions. The work was partially supported by a grant of the Russian Foundation for Basic Research, No. 06-04-48462.

References

Bouchet P, Sysoev A (2001) *Typhlosyrinx*-like tropical deep-water turriform gastropods (Mollusca, Gastropoda, Conoidea). J Nat Hist 35(11):1693-1715

Bouchet P, Lozouet P, Maestrati P, Heros V (2002) Assessing the magnitude of species richness in tropical marine environments: exceptionally high numbers of molluscs at a New Caledonia site. Biol J Linn Soc 75: 421-436

Bouchet P, Sysoev A, Lozouet P (2004) An inordinate fondness for turrids. *Molluscan Megadiversity: Sea, Land and Freshwater. World Congress of Malacology (Perth, Western Australia, 11-16 July 2004), Abstracts*: 12

Dayrat B, Tillier A, Lecointre G, Tillier S (2001) New clades of Euthyneuran Gastropods (Mollusca) from 28S rRNA sequences. Mol Phylogenetics Evol 19: 225-235

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. Mol Mar Biol Biotechnol 3: 294-299

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98

Hassouna M, Michot B, Bachellerie J-P (1984) The complete nucleotide sequence of mouse 28SrRNA gene. Implications for the process of size increase of the large subunit rRNA of higher eukaryotes. Nucleic Acids Res 12: 3563-3583

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA Barcodes. Proc R Soc Lond B 270: 313-321

Higo S, Callomon P, Gotō Y (1999) Catalogue and bibliography of the marine shell-bearing Mollusca of Japan. Elle Scientific Publications, Osaka

Huelsenbeck JP, Ronquist F, Hall B (2001) MBayes: bayesian inference of phylogeny. Bioinformatics 17: 754-755

Iredale T (1929) Queensland molluscan notes, No. 1. Mem Queensland Mus 9: 261-295

Jovelin R, Justine J-L (2001) Phylogenetic relationships within the Polyopisthocotylean monogeneans (Plathyhelminthes) inferred from partial 28S rDNA sequences. Int J Parasitol 31: 393-401

Kantor YuI (1990) Anatomical basis for the origin and evolution of toxoglossan mode of feeding. Malacologia 32(1): 3-18

Kantor YuI (2006) On the morphology and homology of the “central tooth” in the radulae of Turrinae (Conoidea: Turridae). Ruthenica 16(1-2): 47-52

Kantor YuI, Medinskaya AI, Taylor JD (1997) Foregut anatomy and relationships of the Crassispirinae (Gastropoda, Conoidea). Bull Nat Hist Mus, London (Zoology) 63(1): 55-92

Kantor YuI, Taylor JD (2000) Formation of marginal radular teeth in Conoidea (Neogastropoda) and the evolution of the hypodermic envenomation mechanism. J Zool, London 252: 251-262

Kantor YuI**,** Taylor JD **(**2002) Foregut anatomy and relationships of raphitomine gastropods (Gastropoda: Conoidea: Raphitominae). In "Systematics, Phylogeny and Biology of the Neogastropoda" Ed by M Oliverio, R Chemello. Boll Malacologico, Suppl 5: 161-174

Kilburn RN (1983) Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 1. Subfamily Turrinae. Ann Natal Mus 25(2): 549-585

Kiener LC (1839) Genre Pleurotome. (*Pleurotoma*, Lam.). In "Spécies général et Iconographie des coquilles vivantes" 5: 17-84

Lamarck JB (1822) Histoire naturelle des animaux sans vertèbres, 6(2). Paris, Verdière. 1-232

McLean J (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, Sysoev AV (2001) The foregut anatomy of the genus *Xenuroturris* (Gastropoda, Conoidea, Turridae), with a description of a new genus. Ruthenica 11(1): 7-14

Okutani T (2000) (ed.) Marine mollusks in Japan. Tokai University Press

Olivera BM (2002) The gastropod genus *Xenuroturris* (Iredale, 1929) evaluated and a new turrid *Lophiotoma olangoensis* described from the central Philippines. *Science Dillman* 14(2): 40-50

Posada D, Crandall KA (2001) Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). *Molecular Biology and Evolution* 18: 897-906

Powell AWB (1964) The family Turridae in the Indo-Pacific. Part 1, The subfamily Turrinae. Indo-Pacific Mollusca 1(5): 227-345

Sowerby GB (1908) Descriptions of new species of *Terebra, Pleurotoma, Trochus, Tellina, Dosinia,* and *Modiola*. Proc Malac Soc London 8(4): 198-201

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony* and other methods. Sunderland: Sinauer Associates

Sysoev AV (2002) On the type species of *Iotyrris* Medinskaya et Sysoev, 2001 (Gastropoda, Turridae). Ruthenica 12(2): 169-171

Sysoev AV, Kantor YuI (1990) A new genus and species of "*Cochlespira*-like" turrids (Gastropoda, Toxoglossa, Turridae). Apex 5(1-2): 1-6

Taylor JD, Kantor YuI, Sysoev AV (1993) Foregut anatomy, feeding mechanisms, relationships and classification of Conoidea (Toxoglossa) (Gastropoda). Bull Nat Hist Mus, London (Zoology) 59: 125-169

Tippett DL (2006) The genus *Strictispira* in the western Atlantic (Gastropoda: Conoidea). Malacologia 48(1-2): 43-64

Tucker JK (2004) Catalog of Recent and fossil turrids (Mollusca: Gastropoda). Zootaxa 682: 1-1295

Wilson B (1994) Australian Marine Shells, Vol. 2—Prosobranch Gastropods. Part two (Neogastropoda). Odyssey Publishing, Kallaroo, Australia

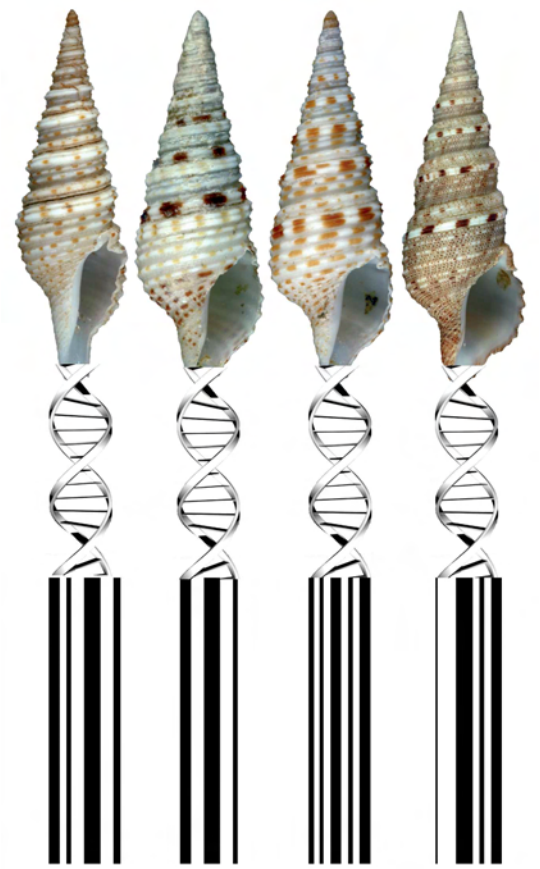
5. Identifier d'autres caractères

Dans la plupart des cas, les caractères et les critères additionnels permettent donc de confirmer ou d'infirmer les hypothèses primaires de délimitation d'espèces proposées par l'analyse exploratoire de la variabilité du gène COI. Dans le cas de la sous-famille des Turrinae notamment, il sera tout de même nécessaire de compléter l'analyse par le séquençage d'un gène nucléaire plus variable que le 28S. Le gène ITS2 n'a pas pu être séquencé avec succès pour un grand nombre de spécimens de Turrinae, mais il pourrait constituer un gène intéressant pour cette sous-famille. Les introns de gènes nucléaires sont également des cibles intéressantes, car leur fort taux de mutation leurs permettrait d'accumuler des différences même entre espèces proches. Cependant, ces loci sont très peu utilisés chez les gastéropodes et leur utilisation nécessiterait un travail de recherche de loci et de définition d'amorces au préalable.

Dans le genre *Benthomangelia*, un tel gène nucléaire permettrait notamment de tester les hypothèses alternatives définies par le COI, où un groupe inclut trois populations différentes génétiquement (pour le gène COI) et morphologiquement. Dans ces cas, où l'analyse de différents caractères permet de proposer différentes hypothèses de délimitation d'espèces, les populations analysées pourraient se situer entre le moment de la spéciation et le moment où les deux espèces sont totalement différenciées, comme illustré sur la figure 6 (page 55). Certains caractères ou critères indiquent déjà que les deux groupes sont différents, alors que d'autres, se mettant en place plus tard dans le processus de spéciation, ne différencient pas encore les deux espèces.

Les analyses réalisées permettent également l'exploration d'une variabilité morphologique jusque là non utilisée. Il est ainsi possible de trouver des différences morphologiques importantes qui viennent corroborer les résultats moléculaires. Le jeu de données correspondant à la sous-famille des Turrinae nécessiterait une analyse morphologique similaire à celle qui a été mise en place pour le genre *Benthomangelia*, notamment dans les cas où aucune différence morphologique n'a pu être mise en évidence avec les approches morphologiques traditionnelles (espèces pseudo-cryptiques).

CONCLUSION



L’obsolescence est un destin qu’il faut sincèrement espérer, si l’on ne veut pas que la science stagne et dépérisse.

S. J. Gould

A. Vers une nouvelle taxonomie

1. Délimitation et description d'espèces

L'analyse taxonomique des Turridae aura non seulement permis de proposer de nouvelles hypothèses phylogénétiques et de délimitation d'espèces, mais aura été surtout l'occasion de mettre au point et de tester les nouveaux outils et les nouvelles méthodes de la taxonomie. La méthodologie associe une étape exploratoire, au cours de laquelle des méthodes récentes mais également une nouvelle méthode mise au point dans le cadre de la thèse ont été appliquées, et une étape de tests des hypothèses de délimitation d'espèces proposées. Au sein des Turridae, cette méthodologie aura permis de délimiter au minimum 111 espèces. Parmi elles, environ 70 n'ont pas pu être associées à un nom disponible dans la littérature : les 2 nouvelles espèces du genre *Iotyrris* et les 11 du genre *Bathytoma* ont été décrites et nommées (en collaboration avec des spécialistes du groupe : Yuri Kantor et Alexander Sysoev), mais les 4 du genre *Benthomangelia* et les quelques 50 de la sous-famille des Turrinae ne le sont pas encore.

Plus généralement, le rythme de découverte de nouvelles espèces est en augmentation grâce à des approches de taxonomie intégrative, notamment basée sur une analyse moléculaire (Paquin & Hedin 2004). En revanche, dans beaucoup d'articles, ces délimitations d'espèces ne sont pas suivies de descriptions formelles (Miller 2007) : il n'existe pas pour l'instant de nom pour ces espèces, et elles ne peuvent pas, par exemple, être ajoutées à une liste d'espèces menacées, dans le cadre de la mise en place d'éventuels plans de sauvegarde.

Dans plusieurs cas, les hypothèses de délimitation d'espèces ne sont pas suffisamment robustes, notamment lorsque des contradictions existent entre différents caractères (Monaghan *et al*. 2006), pour permettre de décrire cette diversité. Cependant, même lorsque ces hypothèses sont robustes, elles peuvent ne pas être décrites et associées à un nouveau nom pour des raisons liées au processus de description en lui-même :

1) une description formelle d'une espèce doit suivre des règles de nomenclature strictes, dictées par le Code International de Nomenclature Zoologique (ICZN, 1999 ; Minelli 1999 ; Knapp *et al*. 2004). Rédiger une description suffisamment complète, tout en suivant ces règles, nécessitent un savoir-faire que la plupart des biologistes moléculaires auteurs d'analyses taxonomiques basées sur l'ADN n'ont pas.

2) la description de nouvelles espèces nécessite de se plonger dans une littérature parfois très ancienne et rédigée dans une langue autre que l'anglais (Godfray 2002 ; Balakrishnan 2005). Cette revue bibliographique est indispensable pour tenir compte des descriptions passées réalisées dans le groupe étudié, et notamment des différents noms qui existent (Minelli 2003 ; Bertrand *et al*. 2006). En effet, pour décrire une nouvelle espèce et lui attribuer un nouveau nom, il faut s'assurer qu'il n'existe aucun nom dans la littérature qui pourrait s'appliquer à cette espèce. A ce titre, l'inclusion des spécimens-types à l'analyse permet de lier de façon univoque une espèce délimitée à un nom déjà existant (Chapitre 2 page 63).

3) les descriptions d'espèces sont dans la majorité des cas publiées dans des revues à faibles facteurs d'impacts, et ne sont donc pas valorisées en tant que productions scientifiques (Minelli 2003). Les évaluations des laboratoires et des chercheurs étant en majorité basées sur la qualité des publications en terme de facteurs d'impacts, l'intérêt de publier des descriptions d'espèces est faible de ce point de vue. Agnarsson & Kuntner (2007) expliquent que ce problème est lié au fait que les auteurs de descriptions d'espèces ne sont jamais (ou très rarement) cités dans la bibliographie des articles. Cependant, même dans les revues avec un bon facteur d'impact et qui acceptent de publier des descriptions d'espèces, la délimitation de nouvelles espèces à l'aide notamment de l'outil moléculaire n'est pas systématiquement

suivie de descriptions formelles. Parmi ces revues, Zoologica Scripta (ZS) et Zoological Journal of the Linnean Society (ZJLS) comptent parmi celles ayant un facteur d'impact élevé (2.364 et 2.290 respectivement, en 2007). ZS et ZJLS ont publié respectivement 27 et 51 articles en 2008. Six pour ZS et cinq pour ZJLS incluent des délimitations de nouvelles espèces basées en partie ou totalement sur l'ADN, dont deux pour ZS (Fritz *et al*. 2008 ; Rezac *et al*. 2008) et quatre pour ZJLS (Malaquias & Reid 2008 ; Mariaux *et al*. 2008 ; Poulakakis & Sfenthourakis 2008 ; Savage & Watling 2008) associent les descriptions d'espèces. Les auteurs d'un des articles publiés dans ZS et qui n'inclut pas les descriptions d'espèces précisent qu'elles seront publiées prochainement dans un autre article (Moura *et al*. 2008).

2. Taxonomie intégrative et gestion des collections

Malgré ces réticences ou ces difficultés à concrétiser ces délimitations d'espèces en leur attribuant des noms, les projets barcoding joue un rôle de moteur pour les sciences taxonomiques (Mace 2004 ; Raven 2004). Ils sont notamment à l'origine du regain d'intérêt pour la taxonomie, que ce soit au niveau conceptuel ou méthodologique. Les changements dans les pratiques taxonomiques concernent notamment la quantité de données accumulées : comme on l'a vu, délimiter des espèces et créer des bases de données de référence dans le cadre des projets barcoding nécessitent la récolte de nombreux spécimens, chacun associé à un grand nombre d'informations (Janzen *et al*. 2005).

La gestion de ces spécimens, comme ceux conservés dans les collections du MNHN, nécessite la mise en place de nouveaux outils, notamment informatique. Dans le cas des collections « mollusques » du MNHN, les pratiques liées à la récolte des spécimens ont évolué (Chapitre 3 page 94), chaque spécimen collecté vivant étant conservé spécifiquement pour les analyses moléculaires. Pour permettre une analyse intégrative de ces spécimens, faisant appel aux différentes informations qui lui sont liées, il est indispensable que chaque individu soit associé à toutes ces informations par l'intermédiaire d'un numéro unique. Les nouvelles données qui seront par la suite acquises (notamment les séquences d'ADN) seront également liées au spécimen par l'intermédiaire de ce numéro (Encadré 6). Cette traçabilité permet au chercheur d'accéder facilement aux collections (coquille, tissu, ADN,…) et à toutes les informations qui lui sont associées (informations de collecte, identification taxonomique, images,…) (Brooke 2000 ; Graham *et al*. 2004 ; De Ley *et al*. 2005 ; Ward *et al*. 2005 ; Jones & Blaxter 2006 ; Guralnick *et al*. 2007).

La taxonomie entre donc dans l'ère de l'informatique. Les taxonomistes, en tant que descripteurs des unités de base du vivant, doivent permettre un accès facile à toutes les informations produites (article 7 de la Convention sur la Diversité Biologique, http://www.cbd.int), pour le scientifique spécialiste d'un groupe mais aussi pour tous les utilisateurs finaux, du gestionnaire de plans de sauvegarde au simple curieux de la nature. Plusieurs outils de diffusion des données taxonomiques sont également en train d'être mis en place, notamment par le biais d'Internet (voir par exemple : Encyclopedia of Life, http://www.eol.org). Certains auteurs n'hésitent plus à parler d'e-types, des spécimens-types numérisés et accessibles via internet (Wilson 2004), et de web-taxonomie, où les connaissances taxonomiques qui ont été et qui seront accumulées seraient également publiées sur la toile (Scoble 2004 ; Godfray 2007 ; Wiens 2007).

ENCADRE 6

Base de données moléculaire

L'un des aspects importants de ma thèse était la gestion des spécimens et de l'information qui y était associée, sous toutes ces formes. La principale difficulté est en effet de ne pas perdre le lien entre les différentes parties du spécimen de référence (voucher) : le spécimen en lui-même (la coquille), le tissu, l'ADN, les informations liées à la collecte, les séquences d'ADN,… A titre personnel, l'utilisation du tableur Excel m'a permis de gérer cette information pendant ma thèse, mais les problèmes liés à la quantité d'informations et au nombre de personnes susceptibles d'utiliser, de gérer ou de compléter ces informations nécessitent la mise en place d'autres outils. De plus, le nombre de projets barcoding au MNHN est en constante augmentation. L'objectif est donc de proposer un outil de gestion des spécimens et de toute l'information associée dans le cadre de projets barcoding, utilisable par tous les gestionnaires de ce type de projet, quel que soit l'organisme sur lequel ils travaillent.
J'ai pour ma part participé à une réflexion qui a été menée sur la mise au point d'outils informatiques pour la gestion de ces données. Le premier résultat obtenu est une base de données qui a été construite sous un système de gestion Oracle, associé à une interface Jacim (Figure 14). Une discussion organisée avec plusieurs chercheurs et gestionnaires de collection du muséum nous a amené à proposer un autre schéma de fonctionnement, actuellement en cours de réalisation, où les informations sont dissociées en deux parties :
- une base de données « spécimens » : elle inclut toutes les informations liées au spécimen (récolte, identification, lieu de stockage,…). Au MNHN, chaque groupe d'organismes possède une base de données « spécimens » qui lui est propre (une pour les mollusques, une autre pour les crustacés,…). Ces bases de données sont parfois construites sous des modèles différents, et chaque base de données « spécimens » restera donc indépendante. De plus, certaines bases de données « spécimens » sont déjà en place depuis plusieurs années, alors que d'autres sont en cours de réalisation. Dans plusieurs cas, une des difficultés principales a été de passer d'un système centré sur un lot (plusieurs spécimens correspondent à une seule entrée dans la base de données) à un système centré sur le spécimen (une entrée par spécimen), chaque spécimen pouvant être associé à des informations différentes (comme une séquence ADN, par exemple).
- une base de données « moléculaire » : elle inclut toutes les informations liées à l'analyse moléculaire d'un spécimen, de l'extraction de l'ADN à l'analyse des séquences. Cette base de données sera commune à tous les taxa, et permettra de proposer un modèle standardisé de gestion des données moléculaires. Le lien entre ces deux bases de données se fait par l'intermédiaire d'un numéro d'inventaire unique attribué à chaque spécimen.

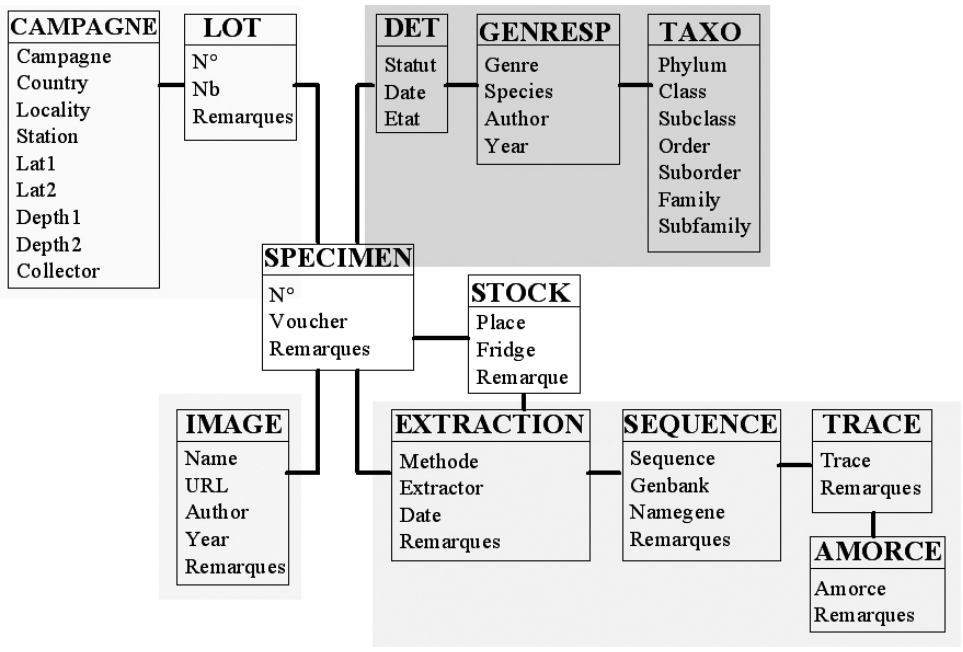


Figure 14 : schéma simplifié de la base de données « barcode ». Au centre se trouve le spécimen, auquel sont rattachées les images, les informations liées à la récolte, à l'identification taxonomique et à l'analyse moléculaire.

B. Perspectives : au-delà de la délimitation d'espèces

Les différents résultats obtenus au cours de cette thèse, qu'ils concernent les nouvelles hypothèses phylogénétiques ou les délimitations d'espèces qui ont été proposées, permettent non seulement d'enrichir notre connaissance de la biodiversité, mais fournissent également une matière indispensable pour l'analyse des processus qui sont à l'origine de cette biodiversité. Je présente ici deux perspectives offertes par l'analyse des patterns décrits au cours de cette thèse : (i) l'estimation de la richesse spécifique et ce qu'elle nous permet de déduire sur la structure des écosystèmes et (ii) l'analyse du processus de spéciation, des facteurs qui en sont à l'origine et des modalités d'évolution des caractères au cours de ce processus.

1. Distribution des Turrinae et richesse spécifique

Pour la sous-famille des Turrinae, le jeu de données analysées représente une grande quantité de données (1000 spécimens, 88 espèces – article 5 page 106). Son analyse, présentée dans l'article 9, a permis de fournir des informations sur la distribution géographique et bathymétrique des espèces, et également sur la richesse spécifique de ce groupe.

La plupart des espèces qui ont été délimitées au sein de la sous-famille des Turrinae ont des aires de distribution géographique et bathymétrique qui paraissent le plus souvent très limitées. Ces résultats peuvent être interprétés à la fois comme la conséquence de phénomènes d'endémisme ou de spécialisation écologiques, mais également comme un biais dû à l'échantillonnage réalisé. Plusieurs missions dont sont issus les spécimens analysés étaient restreintes à une zone bathymétrique donnée, comme la mission Panglao 2004, uniquement littorale, ou à un environnement particulier, comme les missions EBISCO et Norfolk 2, dédiées à l'exploration des monts sous-marins.

Les courbes d'accumulation d'espèces indiquent également que quel que soit l'aire géographique ou la section bathymétrique considérée, un nombre important d'espèces n'a pas encore été échantillonné. Ce résultat confirme la difficulté d'échantillonner ce groupe, à la fois à cause de sa grande diversité, de sa rareté, et au moins pour les milieux profonds de la difficulté de récolter des spécimens. Une grande partie des espèces délimitées n'a pas pu être associée à un nom d'espèce disponible dans la littérature, particulièrement pour celles collectées en milieu profond. Les résultats permettent ainsi d'estimer qu'il existerait plus de 500 espèces de Turrinae dans le monde, et confirment également les estimations du nombre d'espèces pour le groupe des Turridae qui avaient été faites précédemment : il existerait plus de 10000 espèces de Turridae dans le monde.

La validité de telles estimations de la diversité spécifique réside essentiellement dans la façon dont les hypothèses de délimitation d'espèces utilisées pour faire ces estimations ont été proposées. Il a été montré que les estimations de richesses spécifiques basées non pas sur des hypothèses de délimitation d'espèces telles qu'elles ont pu être proposées pour les Turridae, mais sur des morphoespèces ou des RTUs (Recognizable Taxonomic Units), peuvent présenter des taux d'erreurs importants (Krell 2004 ; Caesar *et al*. 2006). De plus, l'utilisation de RTUs ou de morphoespèces n'est en général pas répétable : il s'agit d'un avis subjectif donné par un taxonomiste, avis qui pourra être différent pour un autre taxonomiste. Utiliser des caractères et des critères clairement définis permet d'assurer que les hypothèses de délimitation d'espèces proposées ne dépendent que des données observées, et pas de l'observateur.

Les conséquences de ces erreurs d'estimations de diversité biologique ne sont pas négligeables. En effet, la diversité biologique est l'un des paramètres régulièrement utilisés pour étudier la structure et le fonctionnement d'un écosystème. Comparer la diversité biologique entre deux environnements permet notamment d'estimer l'influence de différents facteurs sur ces environnements. Par exemple, la richesse spécifique permet d'évaluer l'impact des activités humaines (McKinney 2002), notamment lié à l'introduction d'espèces invasives (Gurevitch & Padilla 2004) sur l'écosystème. Ce type d'estimation permet également d'identifier les milieux riches en espèces, et potentiellement ceux dont la diversité est encore largement inconnue (comme les milieux profonds dans le cas des Turridae). Ces milieux pourront ensuite être étudiés prioritairement, notamment dans le cadre des projets barcoding (comme par exemple le projet d'exploration de la diversité marine, MarBoL, pour Marine Barcode of Life, qui vient de démarrer, notamment au MNHN). Elles peuvent également permettre d'identifier les milieux qui doivent être préservés en priorité, en mettant par exemple en évidence des milieux pour lesquels la diversité biologique est en déclin (Moritz 2002).

Diversity and distribution of the Turrinae (Gastropoda; Conoidea) in South-West Pacific

N. Puillandre[1,2,*], Philippe Bouchet[2], S. Samadi[1]

[1] Université Pierre et Marie Curie (UPMC), UMR 7138, Systématique, adaptation, évolution (UPMC/IRD/MNHN/CNRS), Paris, France.

[2] Muséum National d'Histoire Naturelle (MNHN), Departement Systematique et Evolution, UMR 7138 Systématique, Adaptation, Evolution, 55 rue Buffon, 75231 Paris Cedex 05, France

* Corresponding author: UMR 7138, Systématique, adaptation, évolution, UPMC; IRD; MNHN; CNRS, Service de systématique moléculaire (CNRS, IFR 101), Département systématique et évolution, Muséum National d'Histoire Naturelle, CP26, 57 rue Cuvier, 75231 Paris Cedex 05, France.; puillandre@mnhn.fr; +33 1 40 79 37 43

En préparation ; pour soumission à Malacologia

Abstract

The subfamily Turrinae is a hyperdiverse group of marine gastropods, including a large amount of undescribed species. Recently, an integrative taxonomy approach, including the analysis of molecular characters, was conducted on a dataset comprising 1000 specimens collected in South-West Pacific, leading to the recognition of 88 species. We propose here to analyse the geographical and bathymetrical distribution of these species. A large number includes only few specimens, collected in a limited geographical and bathymetrical range. Even if endemism or ecological specialization can be involved, the results are also linked to bias sampling, as all the cruises during which the specimens were collected are not equivalent in terms of environment prospected and of numbers of specimens collected. Furthermore, the results also indicate that the species richness of the Turrinae in the sampled area is underestimated, whatever the geographical or bathymetrical scale considered.

Introduction

The subfamily Turrinae (Conoidea, Turridae) is a hyperdiverse group of predatory marine gastropods. It is worldwide distributed, present from shallow water to abysses (Powell 1966). More than one hundred species are recognized within the subfamily in the South-West Pacific, but a large amount of diversity remains undescribed. Actually, morphological variability observed in material collected in the last decades by several cruises set up by the Muséum National d'Histoire Naturelle (MNHN) and the Institut de Recherche pour le Développement (IRD) suggests that the total number of species may be far higher than the actual number of described species. However, because of plasticity of shell characters and convergence phenomena (Kantor *et al*. unpublished results), delimiting species based on morphological characters is particularly difficult in this group.
Puillandre et al. (unpublished results) proposed an integrative approach, combining molecular (COI and 28S genes), morphological and geographical data, to delimit species in the subfamily Turrinae in the South-West Pacific. The COI gene was used to propose primary hypotheses of species delimitation, using two exploratory methods based on phylogenetic reconstructions. Then, several other characters (a nuclear gene, the 28S, but also morphological and ecological characters) were used to discuss these hypotheses and to define species boundaries in this group. Eighty-eight species were recognized. Among them, seven potentially includes two or three different species: because of a lack of evidence for reproductive isolation, they will be each considered only one species in the present analysis.
This survey of species diversity of Turrinae (Puillandre et al. unpublished results) has revealed that the diversity in this group could be even more important than stated by looking only at morphological characters. Actually, a single morphospecies can include several molecular entities. This result highlights the need to use new tools, such as DNA-taxonomy (Vogler & Monaghan 2007), to complement traditional morphologically-based taxonomy, especially in hyperdiverse groups. Only a robust taxonomical framework will allow a correct evaluation of species richness and distribution. Furthermore, this framework is also a necessary pre-requisite for all other biological or ecological studies (Agapow *et al*. 2004).

This study was based on the analysis of one thousand specimens collected in various places in South-West Pacific (Philippines, Solomon Islands, Vanuatu…). This extensive sampling is here used to examine distribution patterns of the delimited species of Turrinae. The main objective is to describe the abundance and the geographic and bathymetric distributions of the species. Species diversity is also discussed, and total species richness is compared between different bathymetric ranges and the different geographic regions prospected. By confronting the number of described and undescribed species, an extrapolation of the total diversity of the Turrinae in the South-West Pacific is proposed. However, the bathymetric range prospected and the sampling effort was not identical between the different cruises, and this potential bias has to be integrated when comparing the results obtained in different geographic regions. Such global analysis of diversity and distribution also provides several useful information to discuss potential endemicity or ecological specialization.

Material and Methods

Dataset used
All specimens of Turrinae were collected between 2004 and 2007 in several deep-sea cruises conducted by the MNHN and the IRD in Taiwan, Philippines, Solomon Islands, Vanuatu, Chesterfield Islands and Norfolk ridge (South of New-Caledonia). Collection data for each specimen used in this study are available, including geographic coordinates and depth of

collection (see details in Table 1). Species delimitation defined by the integrative taxonomy analysis are presented also in Table 1. Specimens collected in the field were analysed without taking into account *any prior morphospecies* sorting. This allowed a non biased analysis by including in the analysis potential cryptic species that would have been missed with a preliminary morphological sorting. One thousand specimens were thus analysed, collected in 232 stations.

Table 1: Geographic and bathymetric distribution of the Turrinae species. Species name or acronym, number of species, occurrences in the ten cruises (A07 = Aurora 07, Philippines; BOA = Santo, Vanuatu; EB = EBISCO, Chesterfield Islands; NOR = Norfolk ridge; P04 = Panglao 2004, Philippines; P05 = Panglao 2005, Philippines; SAL2 = Salomon 2, Solomon Islands; SAL3 = Salomon 3, Solomon Islands; S06 = Santo 2006, Vanuatu; T04 = Taiwan 2004), and bathymetric range.

Species	Nb	A07	BOA	EB	NOR	P04	P05	SAL2	SAL3	S06	T04	Depth
Gemmula cf. acuta	3					x				x		9-99m
Gemmula cf. cosmoi 1	71	x				x	x					75-364m
Gemmula cf. cosmoi 2	24		x							x		228-366m
Gemmula cf. hombroni 1	1					x						3m
Gemmula cf. hombroni 2	18									x		0-99m
Gemmula cf. hombroni 3	1						x					87m
Gemmula cf. martini	7							x	x			120-125m
Gemmula cf. monolifera 1	1					x						109m
Gemmula cf. monolifera 2	8					x	x	x		x		16-176m
Gemmula cf. monolifera 3	9									x		25-102m
Gemmula cf. monolifera 4/5	9									x		25-99m
Gemmula cf. rarimaculata 1	1			x								420m
Gemmula cf. rarimaculata 2	3			x	x							188-365m
Gemmula cosmoi	1										x	216m
Gemmula diomedea	30					x	x					230-1761m
Gemmula kieneri	10					x	x					75-400m
Gemmula lisajoni	2									x		20-25m
Gemmula sikatunai	6						x					230-295m
Gemmula sp. 1	1									x		274m
Gemmula sp. 10	8							x				155-176m
Gemmula sp. 11	20	x										103-356m
Gemmula sp. 12	24		x						x	x		226-584m
Gemmula sp. 13/14	42	x	x	x					x	x		309-584m
Gemmula sp. 15	1			x								684m
Gemmula sp. 16	1			x								331m
Gemmula sp. 17	77	x				x	x			x		40-155m
Gemmula sp. 18	1	x										350m
Gemmula sp. 19/20	2			x				x				569-743m
Gemmula sp. 2	1					x						120m
Gemmula sp. 21	6	x					x		x	x		421-774m
Gemmula sp. 22	3			x				x				493-743m
Gemmula sp. 23/24/25	11	x	x	x						x		271-520m
Gemmula sp. 26	1	x										427m
Gemmula sp. 27	2		x							x		366-375m
Gemmula sp. 28	1								x			305m
Gemmula sp. 29	1						x					87m
Gemmula sp. 3	1						x					87m
Gemmula sp. 30	5	x				x						33-96m
Gemmula sp. 31	5									x		575m
Gemmula sp. 32	1	x										593m
Gemmula sp. 33	15	x		x	x				x			270-684m
Gemmula sp. 34	24			x	x							440-890m
Gemmula sp. 35	4							x	x			977-1005m
Gemmula sp. 4	1								x			276m
Gemmula sp. 5	3				x							389m
Gemmula sp. 6	1							x				355m
Gemmula sp. 7/8	24	x						x	x	x		420-469m
Gemmula sp. 9	1		x									46m
Gemmula speciosa	9						x	x				65-155m
Iotyrris cingulifera	3									x		9-25m
Iotyrris devoizei	4									x		16-20m
Iotyrris musivum	2									x		20m
Lophiotoma abbreviata	11									x		2-25m
Lophiotoma acuta	101					x				x		0-99m
Lophiotoma albina	4									x		0-25m
Lophiotoma bisaya	7	x					x					75-140m
Lophiotoma brevicaudata	23					x				x		3-38m
Lophiotoma cf. friedrichbonhoefferi 2	2									x		102-342m
Lophiotoma cf. indica 1	7							x		x		95-155m
Lophiotoma cf. indica 2/3	8	x										43-79m
Lophiotoma cf. indica 4/cf. friedrichbonhoefferi 1	4							x		x		116-155m
Lophiotoma cf. indica 5	8		x					x				155-228m
Lophiotoma cf. indica 6	8		x									220-228m

Lophiotoma cf. indica 7/8	94		x					x	x	x	375-584m
Lophiotoma cf. jickelli1/2	10					x				x	0-44m
Lophiotoma cf. panglaonensis 1	4	x				x					255-400m
Lophiotoma cf. panglaonensis 2	10	x									194-311m
Lophiotoma cf. unedo 1	8			x							267-375m
Lophiotoma cf. unedo 2/3	32	x	x		x		x		x	x	152-389m
Lophiotoma cf. unedo 4	68		x						x	x	220-482m
Lophiotoma hastula	4	x								x	67-102m
Lophiotoma panglaonensis/ cf. panglaonensis 3	16		x			x	x		x	x	220-400m
Lophiotoma picturata	14									x	0-25m
Lophiotoma polytropa	6	x				x					0-3m
Lophiotoma ruthveniana	3									x	188-231m
Lophiotoma tabayensis	1	x									158m
Turridrupa anbofasciata	1									x	25m
Turridrupa astricta	4					x				x	80-188m
Turridrupa bijubata	1					x					7m
Turridrupa sp. 1	3			x							361-407m
Turridrupa sp. 2	2									x	20-25m
Turridrupa sp. 3	1									x	25m
Turridrupa sp. 4	4					x				x	25m
Turris babylonia	1					x					15m
Turris garnonsii	15									x	0-35m
Turris spectabilis	1									x	130m
Xenuroturris gemmuloides	2			x							388-392m
Xenuroturris legitima	1									x	0m

However, the sampling strategy was not equivalent in all cruises. The cruise Panglao 2004 (Philippines) was exclusively littoral (almost all specimens were collected between 0 and 100 m deep). On the contrary, cruises Aurora 2007 (Philippines), BOA (Vanuatu), EBISCO (Chesterfield Islands), Norfolk 2 (South of New-Caledonia), Panglao 2005 (Philippines) and Salomon 2 and 3 (Solomon Islands) were dedicated to the collection of deep species. The Santo 2006 cruise (Vanuatu) was combining littoral and deep collection events. The ecological range of prospecting was limited to the seamounts of Lord How and Norfolk ridges for the cruises EBISCO and Norfolk 2 respectively. Finally, if all the collected specimens were conserved for several cruises, only a few specimens by species were specifically prepared in the field for molecular analyses during Norfolk 2, Panglao 2004, Panglao 2005 and EBISCO cruises.

Diversity, distribution and species richness

For several analyses, results obtained for the different cruises are grouped in five different regions: Philippines (Panglao 2004, Panglao 2005 and Aurora 2007 cruises), Vanuatu (BOA1 and Santo 2006 cruises), Solomon Islands (Salomon2 and Salomon 3 cruises), Chesterfield Islands (EBISCO cruise) and Norfolk ridge (Norfolk 2 cruise). Species diversity in each of the sampled geographic region is first described, especially looking at the number of species shared by different regions. Species diversity is also described at different depth of collection. Numerical (number of specimens by species), geographical (number of stations where the species is found) and bathymetrical (range of depth where the species is found) abundances are also analysed.

Species accumulation curves are estimated using EstimateS software (Colwell 1997), running the program for 100 drawings of the sampled stations. Bootstrap, Jacknife 1 and Jacknife 2 estimators are used to infer the total number of species that are potentially present in the sampled area. Species accumulation curves are calculated for the whole dataset, but also for each of the five geographic regions sampled and for different ranges of depth.

Estimation of total species richness

Of the 88 species delimited, only 28 were linked unequivocally to a species name. The ratio between the number of described species and the total number of species is calculated for different bathymetric ranges. Furthermore, this ratio can be used to infer the total number of species in the sampled area. By considering that this ratio of new species calculated for the sampled area is constant for the whole group, it can be applied to the total number of species described: 92 species of Turrinae are described in Indo-Pacific and 172 in the world (Tucker 2004).

Results

Distribution and abundance

Among the 88 species delimited, 45 are present in the Vanuatu, 40 in the Philippines, 21 in the Solomon Islands, 13 in the Chesterfield and 5 in the Norfolk ridge. Sixty-one are present in only one of these five regions, 19 are present in two of them, five in three of them and three in four of them. The average number of species by station for each region ranges from 0.41 (Vanuatu) to 1.05 (Solomon Islands) (Table 2).

Table 2: Species diversity in each geographic region. The number of stations corresponds to the stations where at least one specimen of Turrinae was collected.

	Nb species	Nb specimens	Nb stations	Nb species/ Nb stations
Philippines	40	281	70	0.57
Vanuatu	45	600	111	0.41
Solomon	21	54	20	1.05
Chesterfield	13	32	18	0.72
Norfolk ridge	5	32	12	0.42

As shown in figure 1A, the number of species found in each range of depth is variable. Forty-seven species are found in shallow water (between 0 and 100m deep); around twenty species are found for each section of 100m until 500m deep, and then the number of species decreases. This result has to be related to the total number of sites that have been prospected for each section of depth: there are almost as many collecting events in the shallow water section than in all the other sections taken together (Figure 1 B).

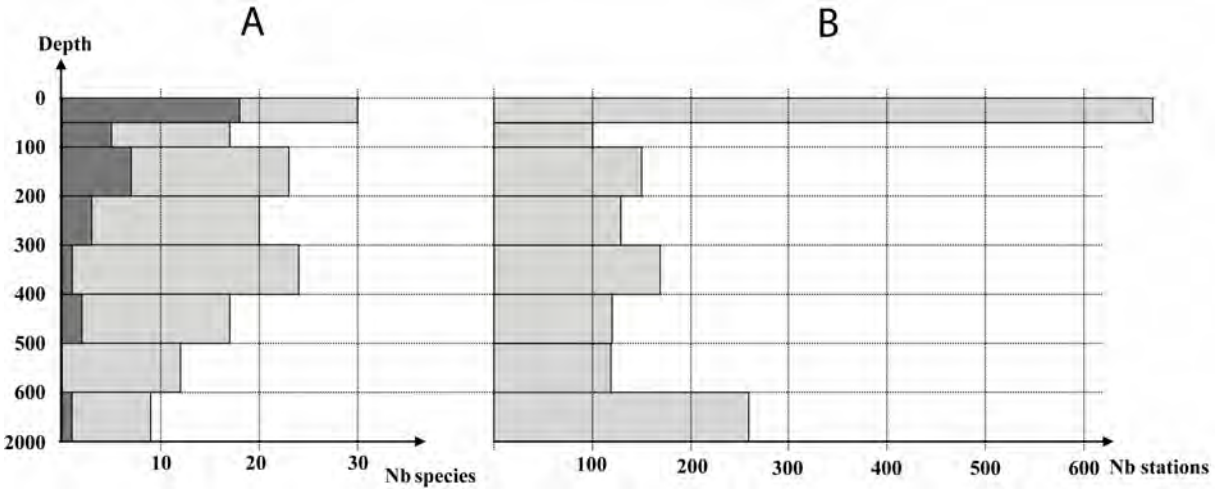


Figure 1: A. Total number of species collected (grey) and number of previously described species (dark grey) in each depth section. B. Total number of stations prospected (with or without specimen of Turrinae) in each depth section.

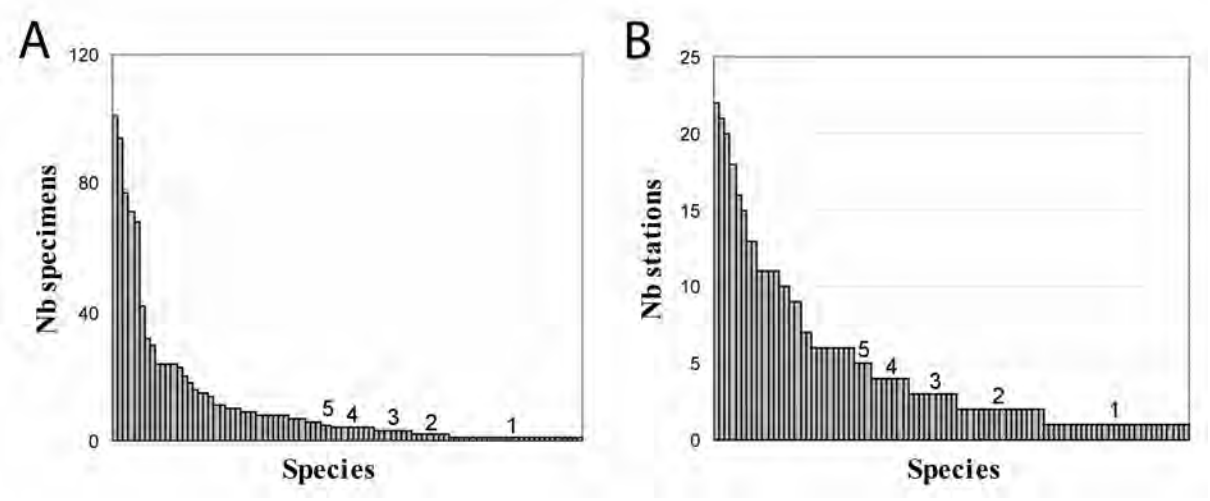


Figure 2: Number of specimens for each species (A) and number of stations where each species were present (B).

The average number of specimens by species is 11, but displays a great variability (from 1 to 101 specimens by species; $\sigma = 19.51$). Twenty-five species are represented by only one specimen, and 47 are represented by less than 5 specimens (Figure 2 A). More than half of the species (59) are present in less than 5 different stations (Figure 2 B). The average range of bathymetric distribution for all the species represented by at least two specimens from different stations is 135m (Table 1). However, deep species seems to be less restricted bathymetrically than littoral species, as shown in Figure 3.

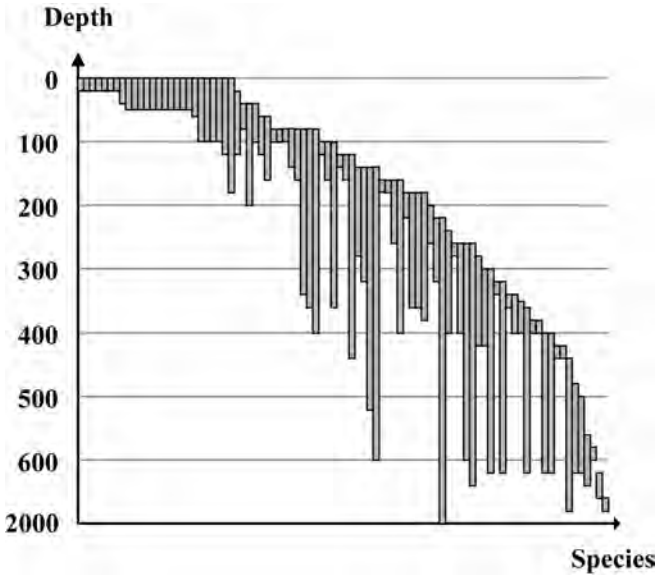


Figure 3: Range of depth where each species were found.

Species richness

The total species accumulation curve shows up that the plateau is not reached (Figure 4). This result, associated to the fact that most of the species are rare numerically and geographically, indicates that we didn't sample all the species present in the prospected area. Indeed, the estimated number of species present in this area ranges from 100 with the bootstrap estimator to 126 with the Jack 2 estimator. Even if the number of species differs from one geographic region to another, shapes of species accumulation curves are similar, with a non-saturated sampling (Figure 5). However, the rate of species discovery in Chesterfield Islands (EBISCO cruise) and Norfolk ridge (Norfolk 2 cruise) is lower than for the others geographic regions. The slope of the curve is also very important for the Solomon Islands, indicating that the weak number of stations prospected (20) is far from allowing a good estimation of the diversity in the archipelago.

Species accumulation curves calculated for each bathymetric section are once again not saturated (Figure 6). This is particularly true for sections deeper than 50m, where collecting effort is weaker (Figure 1 B).

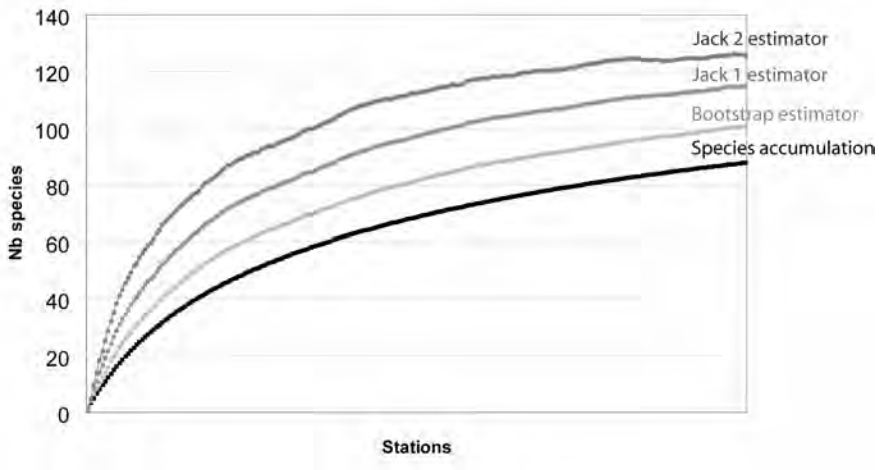


Figure 4: Species accumulation curve for the whole dataset, with estimation of total number of species with Bootstrap, Jack1 and Jack 2 estimators.

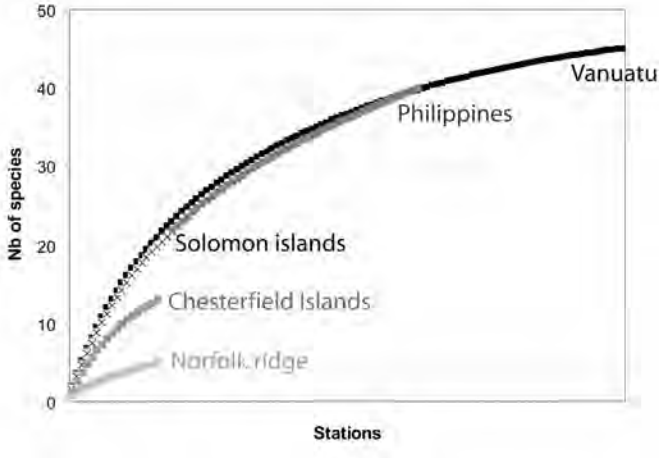


Figure 5: Species accumulation curves for species collected in each geographic region.

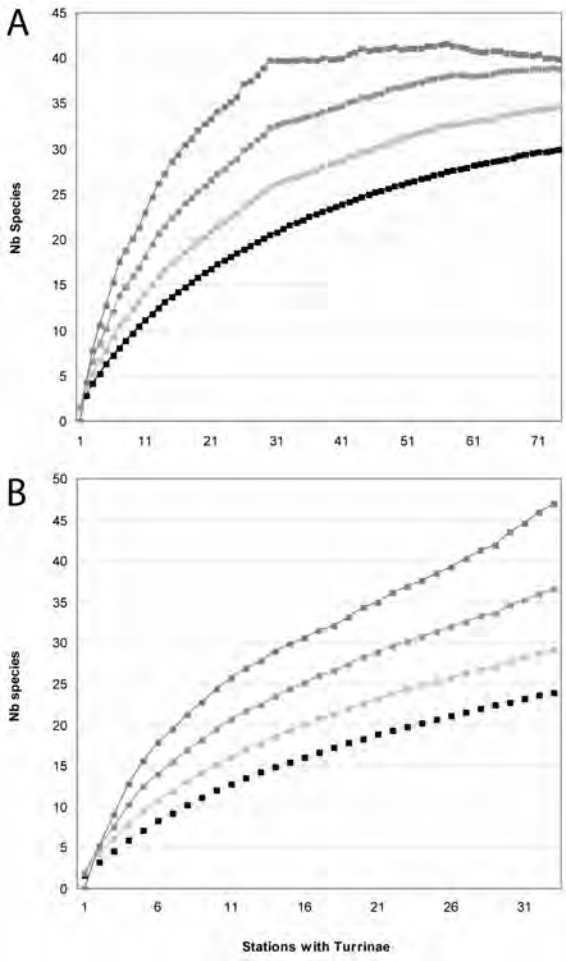


Figure 6: Species accumulation curves with estimation of total number of species with Bootstrap, Jack1 and Jack 2 estimators for species collected between 0 and 50 m deep (A) and between 300 and 400 m deep (B).

Undescribed diversity

Most of these described species are found in shallow waters: the ratio between the number of described species and the total number of species is 0.6 between 0 and 50m, but is only 0.16 after 50m deep (Figure 1). Between 500 and 600m deep, none of the 12 species has been linked to a species name.

The number of new species (i.e. not linked to an available species name) represents an increase of 314% of the species diversity. By applying this ratio to the total number of species described of Turrinae, potentially 289 species occurs in the Indo-Pacific and 540 in the world.

Discussion

Species diversity, abundance and distribution

Even if some species were collected in various places, for example from Philippines to Vanuatu for the species *Gemmula sp. 17*, most of them are represented by few specimens collected in a limited geographical range. The numerical rarity seems to be linked to a limited geographic and bathymetric distributions: as shown in table 1, a high proportion (10/13) of the species represented by more than 20 specimens are present in different geographic regions, contrary to the species represented by less specimens (16/75).

Several hypotheses can be proposed to explain such results. First, such rarity can reflect endemicity, where species occur in a limited area (Meyer et al. 2005). In these cases, a species can be found in large numbers, but only at some stations in a limited area. This is for example the case of the species *G. diomedea*, represented by 30 specimens, but collected only during the cruise Panglao 2004 and 2005, between 9°29'N/123°16'E and 9°39'N/124°03E. Second, it can reflect ecological specialisation, where species are encountered only in one environment. Such a specialized species can be found in various places, but only in a particular environment. This can be the case of *Gemmula sp. 34*, found during EBISCO and Norfolk 2 cruise, i.e. exclusively on seamounts, and not in other cruises where similar depth where prospected.

Finally, such a numerical, geographical or ecological rarity could only be the results of a sampling bias. A limited distribution can be only apparent, resulting from an artifact due to undersampling (Samadi *et al*. 2006; McClain 2007). This hypothesis is supported by the fact that different sampling strategies were followed in each cruise. Actually, this can greatly influence the diversity collected at a given place. During the Panglao 2004 cruise, the only littoral cruise performed in Philippines, specimens were sorted by morphospecies in the field, and only a few specimens by morphospecies were conserved for molecular analyses. During the Santo 2006 cruise in Vanuatu, all specimens collected were conserved for molecular analyses, and several species were found within a single morphospecies. Some littoral cryptic species found in Vanuatu could thus have been missed in Philippines. The same way, only seamounts were prospected in Chesterfield and New-Caledonia (EBISCO and Norfolk 2 cruises): if other environment had been prospected, Vanuatu or Philippines species could have been discovered in these regions.

Importance of sampling strategy

Sampling strategies appear to be crucial for a correct estimation of species diversity. It should privilege the variability of the prospected places, at local scale (different environments, but also different depth) and at large scale (different geographic regions). Furthermore, if specimens of Turrinae are generally large, and thus easy to collect with traditional methods (hand-picking, trawl...), a large number of species of molluscs are actually small species (less than one cm), and thus require specific sampling method (Bouchet *et al*. 2002). Back in the laboratory, the methodology applied to DNA-based species delimitation must also take into account the weak number of specimens by species. Most of the methodologies available are clearly sensitive to such bias (Pons *et al*. 2006, Hebert *et al*. 2003). Efforts are now made by using more genes or coalescent-based approaches (Knowles & Carstens 2007).

Samples used in this study were collected during 10 different cruises, but only one of them fulfils completely these requirements. The Santo 2006 cruise allowed the realisation of almost 700 collection events, prospecting in a large range of environments, differing in terms of habitats (soft to hard bottoms, mangroves, coral reefs, deep ecosystems,...) and sampling methodologies (hand-picking, brushing, suction, trawl, dredge...). Furthermore, a team of four people were dedicated to the processing of molluscs specimens for DNA analyses. Thus, all living specimens were processed, without any kind of prior morphological sorting.
However, even the Santo 2006 sampling does not allow recovering the total diversity of the prospected place, as shown by the species accumulation curves (Fig. 5). If a biased sampling could mislead us on the real distribution of the species, it could also mislead us when comparing the species richness in different geographic regions. For example, the species accumulation curves for Vanuatu and Philippines are superposed, but sampling effort was superior in Vanuatu, especially for littoral species. The same way, less species have been discovered during EBISCO and Norfolk 2 cruises, but only seamounts were prospected during these cruises. Actually, whatever the geographic and bathymetric range considered, new species were discovered, and species accumulation curves are not saturated. Even in shallow waters, a supposedly well-known environment, new species were found. A striking result is the large number of species encountered in deep ecosystems, largely unexplored but thought to be less diverse than shallow water. On the contrary, our results indicate that relatively to the number of prospected stations, deep waters could be more speciose than shallow waters for the subfamily Turrinae (Fig. 2). Furthermore, these environments are more difficult to sample, and thus potentially more subject to under- or biased sampling than shallow waters, suggesting that a higher level of diversity in deep waters has been missed in this analysis. In addition, the diversity in deep waters is proportionally less known, as most of the deep species are still undescribed.

Conclusion

This analysis constitutes the first attempt to estimate species diversity in this group, based on an objective and reproducible species delimitation methodology. Our results confirm previous results obtained for other groups, especially in molluscs: a large amount of species seems rare and localized, and the actual number of species described in the subfamily Turrinae is definitely a large underestimation of the total number of species. Furthermore, this hypothesis is certainly more generally true, especially for the family Turridae (that includes the subfamily Turrinae), that is thought to be the more speciose group of marine molluscs. Applying the same type of estimation as for the subfamily Turrinae, and based on the actual number of described species in the Turridae (4000; Tucker 2004), this family may include 12560 species, thus constituting a potentially inexhaustible source of species discovery.

Acknowledgments

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the Tropical Deep Sea Benthos programme on board R/V Alis, that generated the deep-sea samples used in this study. The Total Foundation supported the Panglao 2004 and Santo 2006 biodiversity surveys.

References

Agapow, P. M., O. R. P. Bininda-Edmonds, K. A. Crandall, J. L. Gittleman, G. M. Mace, J. C. Marshall, and A. Purvis. 2004. The impact of species concept on biodiversity studies. The quarterly review of biology 79:161-179.

Bouchet, P., P. Lozouet, P. Maestrati, and V. Héros. 2002. Assessing the magnitude of species richness in tropical marine environments: exceptionnally high numbers of molluscs at a new Caledonian site. Biological Journal of the Linnean Society 75:421-36.

Colwell, R. K. 1997. EstimateS: *statistical estimation of species richness and shared species from samples*. Version 8.0.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identifications through DNA Barcodes. Proceedings of the Royal Society B 270:313-321.

Knowles, L. L., and B. C. Carstens. 2007. Delimiting species without monophyletic gene trees. Systematic Biology 56:887-895.

McClain, C. R. 2007. Seamounts: identity crisis or split personality? Journal of Biogeography 34:2001-2008.

Meyer, C. P., J. B. Geller, and G. Paulay. 2005. Fine scale endemism on coral reefs: archipelagic differentiation in Turbinid gastropods. Evolution 59:113-125.

Pons, J., T. G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S. Kamoun, W. D. Sumlin, and A. P. Vogler. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55:595-609.

Powell, A. W. B. 1966. The molluscan families Speightiidae and Turridae. An evaluation of the valid taxa, both recent and fossil, with lists of characteristics species. Bulletin of the Auckland. Institute and Museum 5:5-184.

Samadi, S., L. Bottan, E. McPherson, B. Richer De Forges, and M. C. Boisselier. 2006. Seamount endemism questioned by the geographic distribution and population genetic structure of marine invertebrates. Marine Biology 146:1463-1475.

Tucker, J. K. 2004. Catalogue of recent and fossil turrids (Mollusca: Gastropoda). Zootaxa 682:1-1295.

Vogler, A. P., and M. T. Monaghan. 2007. Recent advances in DNA taxonomy. Journal of Zoological Systematics and Evolutionary Research 45:1-10.

2. Le processus de spéciation

Les données accumulées sur les Turridae permettent d'identifier plusieurs groupes d'espèces qui pourraient être utilisés comme modèle pour l'analyse des processus de spéciation. Par exemple, les patterns mis en évidence pour le genre *Bathytoma* ont permis de proposer plusieurs hypothèses préliminaires quant au processus à l'origine de ces patterns (article 6 page 128). La distribution géographique des espèces a notamment permis d'identifier des cas compatibles avec les hypothèses du modèle de spéciation allopatrique. Pour d'autres paires d'espèces, collectées à la même station, le modèle de spéciation sympatrique (ou parapatrique) semble plus adapté.

L'analyse des jeux de données obtenus apporte également plusieurs informations sur les facteurs qui peuvent être à l'origine d'évènements de spéciation. Ainsi, la distribution géographique des espèces délimitées montre qu'un isolement par la distance peut être un facteur intervenant dans les phénomènes de spéciation. Un tel isolement implique l'existence de barrières à la dispersion, mais l'identification de telles barrières en milieu marin reste une question controversée (Palumbi 1994 : Palumbi et Lessios 2005). A l'inverse, en sympatrie, l'apparition de nouvelles espèces est liée à des facteurs environnementaux : deux populations au sein d'une espèce vont diverger si elles s'adaptent chacune à des conditions environnementales différentes (Dieckmann & Doebelli 1999 ; Rissler & Apodaca 2007). Dans le cas des Turridae, l'identification de ces facteurs environnementaux est difficile : la grande profondeur à laquelle sont collectés les spécimens ne nous permet pas en général d'accéder à ces informations (température, pH, composition chimique,…).

Une étude plus approfondie de la diversité de quelques espèces modèles au sein des Turridae permettrait également d'analyser des caractères morphologiques ou moléculaires soumis à sélection (Fontaneto *et al*. 2007). Ces caractères peuvent notamment être directement impliqués dans un processus de spéciation induit par des facteurs environnementaux.

Identifier de tels caractères pourrait permettre de détecter des différences entre espèces récentes, notamment celles qui sont encore en cours de spéciation (« incipient species », Gomez *et al*. 2007). En effet, ces caractères soumis à la sélection peuvent ainsi fixer rapidement des différences entre espèces divergentes (Chapitre 2 page 54). Reconnaître ces espèces récentes constitue très certainement la principale difficulté pour les taxonomistes (Tautz *et al*. 2003 ; Hebert *et al*. 2005 ; Hickerson *et al*. 2006 ; Monaghan *et al*. 2006). Les solutions qui peuvent être envisagées consistent soit à augmenter le nombre de gènes analysés (Hudson & Coyne 2002 ; Felsenstein 2006 ; Shaffer & Thomson 2007), soit à identifier des caractères coalescant plus rapidement (Wang *et al*. 2008), notamment comme on vient de le voir ceux impliqués dans le processus de spéciation. Ces nouveaux marqueurs peuvent également être associés à l'utilisation de nouvelles méthodes basées sur la théorie de la coalescence qui semblent prometteuses (Rosenberg 2003 ; Maddison & Knowles 2006 ; Carstens & Knowles 2007 ; Knowles & Carstens 2007).

Chez les Conoidea, un caractère en particulier intervient potentiellement dans les processus de spéciation. La plupart des Conoidea produit des peptides toxiques (appelés « Conotoxines » chez les Cônes) qui lui permettent de capturer ses proies (Letourneux 2004). Les taux d'évolution de ces séquences peptidiques, analysés chez plusieurs espèces du genre *Conus*, semblent extrêmement rapides (Duda & Palumbi 1999). Cette évolution rapide

pourrait correspondre à une réponse de l'organisme à son environnement, notamment à l'apparition d'une nouvelle proie, d'autant que l'analyse de quelques espèces de *Conus* montre que chacune possède un ensemble de toxines qui lui est propre (Stocklin & Olivera, com. pers.). Par conséquent, cette évolution rapide des toxines, associée à l'occupation d'une nouvelle niche écologique, pourrait être à l'origine d'évènements de spéciation. Autrement dit, des groupes de spécimens possédant des toxines différentes pourraient correspondre à des espèces différentes.

L'objectif serait alors de comparer la diversité des toxines et la diversité spécifique, afin de tester si l'évolution de ces toxines est directement corrélée à l'apparition de nouvelles espèces. Cette question sera l'objet de mon post-doc, au sein du laboratoire du Prof. Baldomero Olivera, à l'Université de l'Utah : elle permettra de valider l'utilisation de la diversité de ces toxines pour une approche de délimitation d'espèces, notamment lorsqu'elles sont récentes.

BIBLIOGRAPHIE

Abdo, Z., and B. G. Golding. 2007. A step toward barcoding life: a model-based, decision-theoretic method to assign genes to preexisting species groups. Systematic Biology 56:44-56.

Adams, B. J. 1998. Species concepts and the evolutionary paradigm in modern nematology. Journal of Nematology 30:1-21.

Agapow, P. M., O. R. P. Bininda-Edmonds, K. A. Crandall, J. L. Gittleman, G. M. Mace, J. C. Marshall, and A. Purvis. 2004. The impact of species concept on biodiversity studies. The quarterly review of biology 79:161-179.

Agnarsson, I., and M. Kuntner. 2007. Taxonomy in a Changing World: Seeking Solutions for a Science in Crisis. Systematic Biology 56:531-539.

Ahrens, D., M. T. Monaghan, and A. P. Vogler. 2007. DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae). Molecular Phylogenetics and Evolution 44:436–449.

Andjic, V., G. E. S. Hardy, M. N. Cortinas, M. J. Wingfield, and T. I. Burgess. 2007. Multiple gene genealogies reveal important relationships between species of *Phaeophleospora* infecting *Eucalyptus* leaves. FEMS Microbiology Letters 268:22-33.

Armstrong, K. F., and S. L. Ball. 2005. DNA barcodes for biosecurity: invasive species identification. Philosophical Transactions of the Royal Society B 360:1813-1823.

Arnot, D. E., C. Roper, and R. A. Bayoumi. 1993. Digital codes from hypervariable tandemly repeated DNA sequences in the *Plasmodium falciparum* circumsporozoite gene can genetically barcode isolates. Molecular and Biochemical Parasitology 61:15-24.

Arzanlou, M., J. Z. Groenewald, R. A. Fullerton, E. C. A. Abeln, J. Carlier, M. F. Zapater, I. W. Buddenhagen, A. Viljoen, and P. W. Crous. 2008. Multiple gene genealogies and phenotypic characters differentiate several novel species of *Mycosphaerella* and related anamorphs on banana. Persoonia 20:19-37.

Baker, C. S., M. L. Dalebout, S. Lavery, and H. A. C. Ross. 2003. www.DNA-surveillance: applied molecular taxonomy for species conservation and discovery. Trends in Ecology and Evolution 18:271-272.

Balakrishnan, R. 2005. Species concepts, species boundaries and species identification: a view from the Tropics. Systematic Biology 54:689-693.

Balastegui, A., P. Ruiz-Lapuente, and R. Canal. 2001. Reclassification of gamma-ray bursts. Monthly Notices of the Royal Astronomical Society 328:283-290.

Ball, S. L., P. D. N. Hebert, S. K. Burian, and J. M. Webb. 2005. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. Journal of the North American Benthological Society 24:508-524.

Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439:719-723.

Barrett, R. D. H., and P. D. N. Hebert. 2005. Identifying spiders through DNA barcodes. Canadian Journal of Zoology 83:481-491.

Baum, D. A. 1998. Individuality and the existence of species through time. Systematic Biology 47:641-653.

Bazin, B., S. Glemin, and N. Galtier. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570-572.

Beaumont, A. R., G. Turner, A. R. Wood, and D. O. F. Skibinski. 2004. Hybridisations between *Mytilus edulis* and *Mytilus galloprovincialis* and performance of pure species and hybrid veliger larvae at different temperatures. Journal of Experimental Marine Biology and Ecology 302:177-188.

Begossi, A., M. Clauzet, J. L. Figueiredo, L. Garuana, R. V. Lima, P. F. Lopes, M. Ramires, A. L. Silva, and R. A. M. Silvano. 2008. Are biological species and higher-ranking categories real? Fish folk taxonomy on Brazil's Atlantic forest coast and in the Amazon. Current Anthropology 49:291-306.

Bensasson, D., D. X. Zhang, and G. M. Hewitt. 2000. Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. Molecular Biology and Evolution 17:406-415.

Bertrand, Y., and M. H. Arlin. 2006. Stability and universality in the application of taxon names in phylogenetic nomenclature. Systematic Biology 55:848-858.

Bertrand, Y., F. Pleijel, and G. W. Rouse. 2006. Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks. Systematics and Biodiversity 4:149-159.

Besansky, N. J., D. W. Severson, and M. T. Ferdig. 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don’t know can hurt you. Trends in Parasitology 19:545-546.

Bichain, J. M. 2007. La systématique des bythinelles revisitéeMuséum National d'Histoire Naturelle, Paris.

Bichain, J. M., M. C. Boisselier, P. Bouchet, and S. Samadi. 2007. Delimiting species in the genus *Bythinella* (Mollusca: Caenogastropoda: Rissooidea): molecular and morphometric approachs. Malacologia 49:291-311.

Blaxter, M. 2003. Counting angels with DNA. Nature 421:122-124.

Blaxter, M. 2004. The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society B 359:669-679.

Blaxter, M., J. Mann, T. Chapman, F. Thomas, C. Whitton, R. Floyd, and E. Abebe. 2005. Defining operationnal taxonomy units using DNA barcode data. Philosophical Transactions of the Royal Society B 360:1935-43.

Bogdanov, I. P. 1987. Kompleksnij podkhod k izucheniiu taksonomichkikh priznakov morskhik pastropodov na primere molliuskov podsemejstva Oenopotinae subfam. nov. Vsesoiuznoe soveshchanie po izucheniiu molliuskov 8:35-37.

Bouchet, P. 1990. Turrid genera and mode of development: the use and abuse of protoconch morphology. Malacologia 32:69-77.

Bouchet, P., P. Lozouet, P. Maestrati, and V. Héros. 2002. Assessing the magnitude of species richness in tropical marine environments : exceptionnally high numbers of molluscs at a new Caledonian site. Biological Journal of the Linnean Society 75:421-36.

Bouchet, P., and J.-P. Rocroi. 2005. Classification and nomenclator of Gastropods families. Malacologia 47.

Bouchet, P., and A. Waren. 1980. Revision of the north-east Atlantic bathyal and abyssal Turridae (Mollusca, Gastropoda). Journal of Molluscan Studies SuppI. 8:1-19.

Bourlat, S. J., H. Nakano, M. Akerman, M. J. Telford, M. C. Thorndyke, and M. Obst. 2008. Feeding ecology of *Xenoturbella bocki* (phylum Xenoturbellida) revealed by genetic barcoding. Molecular Ecology Resources 8:18-22.

Bozikova, E., P. Munclinger, K. C. Teeter, P. K. Tucker, M. Macholan, and J. Pialek. 2005. Mitochondrial DNA in the hybrid zone between *Mus musculus musculus* and *Mus musculus domesticus*: a comparison of two transects. Biological Journal of the Linnean Society 84:363-378.

Bradley, R. D., and R. J. Baker. 2001. A test of the genetic species concept: cytochrom-b sequences and mammals. Journal of Mammalogy 82:960-973.

Brooke, M. L. 2000. Why museums matter. Trends in Ecology and Evolution 15:136-137.

Brookes, J. I., and R. Rochette. 2007. Mechanism of a plastic phenotypic response: predator-induced shell thickening in the intertidal gastropod *Littorina obtusata*. Journal of Evolutionary Biology 20:1015-1027.

Brower, A. V. Z. 2006. Problems with DNA barcodes for species delimitation: 'ten species' of *Astraptes fulgerator* reassessed (Lepidoptera: Hesperiidae). Systematics and Biodiversity 4:127-132.

Caesar, R. M., M. Sörensson, and A. I. Cognato. 2006. Integrating DNA data and traditional taxonomy to streamline biodiversity assessment: an example from edaphic beetles in the Klamath ecoregion, California, USA. Diversity and Distributions 12:483-489.

Calinski, T., and J. Harabasz. 1974. A dendrite method for cluster analysis. Communications in Statistics 3:1-27.

Campbell, D. C., P. D. Johnson, J. D. Williams, A. K. Rindsberg, J. M. Serb, K. K. Small, and C. Lydeard. 2008. Identification of 'extinct' freshwater mussel species using DNA barcoding. Molecular Ecology Resources:10.1111/j.1755-0998.2008.02108.x.

Carstens, B. C., and L. L. Knowles. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. Systematic Biology 56:400-411.

Chacon, M., and O. Luci. 2003. Patients classification by risk using cluster analysis and genetic algorithms. Pages 350-358 *in* Progress in pattern recognition, speech and image analysis (A. Sanfeliu, and J. Ruiz-Shulcloper, eds.). Springer-Verlag, Berlin.

Chase, M. W., N. Salamin, M. Wilkinson, J. M. Dunwell, R. P. Kesanakurthi, N. Haidar, and V. Savolainen. 2005. Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society B 360:1889-1895.

Clare, E. L., B. K. Lim, M. D. Engstrom, J. L. Eger, and P. D. N. Hebert. 2006. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Molecular Ecology Notes 7:184-190.

Costa, F. O., J. R. deWaard, J. Boutillier, S. Ratsnasingham, R. T. Dooh, M. Hajibabaei, and P. D. N. Hebert. 2007. Biological identifications through DNA barcodes: the case of the Crustacea. Canadian Journal of Fisheries and Aquatic Science 64:272-295.

Dayrat, B. 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society 85:407-415.

De Haan, S., M. Bonierbale, M. Ghislain, J. Núñez, and G. Trujillo. 2007. Indigenous biosystematics of andean potatoes: folk taxonomy, descriptors and nomenclature. Acta Horticulturae 745:89-134.

De Ley, P., I. T. Tandingam De Ley, K. Morris, E. Abebe, M. Mundo-Ocampo, M. Yoder, J. Heras, D. Waumann, A. Rocha-Olivares, A. H. Jay Burr, J. G. Baldwin, and W. K. Thomas. 2005. An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. Philosophical Transactions of the Royal Society B 360:1945-1958.

De Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. Pages 57-75 *in* Endless forms. Species and speciation (D. J. Howard, and S. H. Berlocher, eds.). Oxford University Press, Oxford.

De Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879-886.

Degnan, J. H., and N. A. Rosenberg. 2006. Discordance of species trees with their most likely gene trees. PLoS Genetics 2:762-768.

DeSalle, R., M. G. Egan, and M. Sidall. 2005. The unholy trinity : taxonomy, species delimitation and DNA Barcoding. Philosophical Transactions of the Royal Society B 360:1905-16.

Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. Nature 400:354-357.

Dubois, A. 1988. Le genre en zoologie: essai de systématique théorique. Mémoires du Muséum National d'Histoire Naturelle, Paris

Duda, T. F., and A. J. Kohn. 2005. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Molecular Phylogenetics and Evolution 34:257-272.

Duda, T. F., and S. R. Palumbi. 1999. Molecular genetics of ecological diversification: Duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. Proceedings of the National Academy of Sciences 96:6820-6823.

Ebach, M. C., and C. Holdrege. 2005. DNA barcoding is no substitute for taxonomy. Nature 434:697.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Molecular Ecology 17:1170-1188.

Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Clustering analysis of display of genome-wide expression patterns. Proceedings of the National Academy of Sciences 95:14863-14868.

Elias, M., R. L. Hill, K. R. Willmott, K. K. Dasmahapatra, A. V. Z. Brower, J. Mallet, and C. D. Jiggins. 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proceedings of the Royal Society B 274:2881–2889.

Erpenbeck, D., J. N. A. Hooper, and G. Wördheide. 2005. CO1 phylogenies in diploblasts and the 'Barcoding of Life' — are we sequencing a suboptimal partition? Molecular Ecology Notes 1:1-4.

Evans, K. M., A. H. Wortley, and D. G. Mann. 2007. An assessment of potential Diatom ''Barcode'' genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158:349-364.

Fang, S. G., Q. H. Wan, and N. Fijihara. 2002. Formalin removal from archival tissue by critical point drying. BioTechniques 33:604-611.

Fedosov, A., and Y. Kantor. 2008. Toxoglossan gastropods of the subfamily Crassispirinae (Turridae) lacking a radula, and a discussion of the status of the subfamily Zemaciinae. Journal of Molluscan Studies 74:27-35.

Felsenstein, J. 2006. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? Molecular Biology and Evolution 23:691-700.

Ferguson, J. W. H. 2002. On the use of genetic divergence for identifying species. Biological Journal of the Linnean Society 75:509-516.

Flot, J.-F., and S. Tillier. 2007. The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: the putative D-loop and a novel ORF of unknown function. Gene 401:80-87.

Floyd, R., E. Abebe, A. Papert, and M. L. Blaxter. 2002. Molecular barcodes for soil nematode identification. Molecular Ecology 11:839-850.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299.

Fontaneto, D., E. A. Herniou, C. Boschetti, M. Caprioli, G. Melone, C. Ricci, and T. G. Barraclough. 2007. Independently evolving species in asexual Bdelloid rotifers. PLoS Biology 5:1-8.

Fraley, C., and A. E. Raftery. 1998. How many clusters? Which clustering method? Answers via model-based cluster analysis. The computer journal 41:578-588.

Fritz, U., D. Guicking, M. Auer, R. S. Sommer, M. Wink, and A. K. H. Hundsdörfer. 2008. Diversity of the Southeast Asian leaf turtle genus *Cyclemys*: how many leaves on its tree of life? Zoologica Scripta 37:367-390.

Fu, J., and X. Zeng. 2008. How many species are in the genus *Batrachuperus*? A phylogeographical analysis of the stream salamanders (family Hynobiidae) from southwestern China. Molecular Ecology 17:1469-1488.

Fukami, H., A. F. Budd, D. R. Levitan, J. Jara, R. Kersanach, and N. Knowlton. 2004. Geographic differences in species boundaries among members of the *Montastraea annularis* complex based on molecular and morphological markers. Evolution 58:324-337.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology Evolution and Systematics 34:397-423.

Gandolfi, A., P. Bonilauri, V. Rossi, and P. Menozzi. 2001. Intraindividual and intraspecies variability of ITS1 sequences in the ancient asexual *Darwinula stevensoni* (Crustacea: Ostracoda). Heredity 87:449-455.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature 406:220-227.

Gay, L., G. Neubauer, M. Zagalska-Neubauer, C. Debain, J.-M. Pons, P. David, and P.-A. Crochet. 2007. Molecular and morphological patterns of introgression between two large white-headed gull species in a zone of recent secondary contact. Molecular Ecology 16:3215-3227.

Giraud, T., G. Refrégier, M. Le Gac, D. M. de Vienne, and M. E. Hood. 2008. Speciation in fungi. Fungal Genetics and Biology 45:791-802.

Godfray, H. C. J. 2002. Challenges for taxonomy. Nature 417:17-19.

Godfray, H. C. J. 2006. To boldly sequence. Trends in Ecology and Evolution 21:603-604.

Godfray, H. C. J. 2007. Linnaeus in the information age. Nature 446:259-260.

Gomez, A., P. J. Wright, D. H. Lunt, J. M. Cancino, G. R. Carvamho, and R. N. Hughes. 2007. Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon. Proceedings of the Royal Society B 274:199-207.

Gompert, Z., C. C. Nice, J. A. Fordyce, M. L. Forister, and A. M. Shapiro. 2006. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. Molecular Ecology 15:1759-1768.

Gotelli, N. J. 2004. A taxonomic wish-list for community ecology. Philosophical Transactions of the Royal Society B 359:585-597.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. Townsend Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology and Evolution 19:497-503.

Guralnick, R. P., A. W. Hill, and M. Lane. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. Ecology Letters 10:663-672.

Gurevitch, J., and D. K. Padilla. 2004. Are invasive species a major cause of extinctions? Trends in Ecology and Evolution 19 470-474.

Hajibabaei, M., J. R. deWaard, N. V. Ivanova, S. Ratnasingham, R. T. Dooh, S. L. Kirk, P. M. Mackie, and P. D. N. Hebert. 2005. Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of the Royal Society B 360:1959-1967.

Hajibabaei, M., D. H. Janzen, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences 103:968-971.

Harrison, R. G. 1998. Linking evolutionary pattern and process: the relevance of species concepts for the study of speciation. Pages 19-31 *in* Endless Forms: Species and Speciation (D. J. Howard, and S. H. Berlocher, eds.). Oxford University Press, Oxford.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003a. Biological identifications through DNA Barcodes. Proceedings of the Royal Society B 270:313-321.

Hebert, P. D. N., and T. R. Gregory. 2005. The promise of DNA Barcoding for taxonomy. Systematic Biology 54:852-859.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004a. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences 101:14812-14817.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Philosophical Transactions of the Royal Society B 270:S96-S99.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis. 2004b. Identification of birds through DNA barcodes. PLoS Biology 2:1657-1663.

Hedley, C. 1922. A revision of the Australian Turridae. Records of the Australian Museum 13:213-359.

Heralde, F. M., M. Watkins, J.-P. Ownby, P. K. Bandyopadhyay, A. D. Santos, G. P. Concepcion, and B. M. Olivera. 2007. Molecular phylogeny of some Indo-Pacific genera in the subfamily Turrinae H. Adams and A. Adams, 1853 (1838) (Gastropoda: Neogastropoda). Nautilus 121:131-138.

Hey, J. 2006. On the failure of modern species concepts. Trends in Ecology and Evolution 21:447-450.

Hickerson, J. H., C. P. Meyer, and C. Moritz. 2006. DNA Barcoding will often fail to discover new animal species over broad parameter space. Systematic Biology 55:729-739.

Hogg, I. A., and P. D. N. Hebert. 2004. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. Canadian Journal of Zoology 82:749-754.

Holland, B. S., M. N. Dawson, G. L. Crow, and D. K. Hofmann. 2004. Global phylogeography of *Cassiopea* (Scyphozoa: Rhizostomeae): molecular evidence for cryptic species and multiple invasions of the Hawaiian Islands. Marine Biology 145:1119-1128.

Hollander, J., M. L. Collyer, D. C. Adams, and K. Johannesson. 2006. Phenotypic plasticity in two marine snails: constraints superseding life history. Journal of Evolutionary Biology 19:1861–1872.

Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557-1565.

Hunter, S. J., T. I. Goodall, K. A. Walsh, R. Owen, and J. C. Day. 2008. Nondestructive DNA extraction from blackflies (Diptera: Simuliidae): retaining voucher specimens for DNA barcoding projects. Molecular Ecology Resources 8:56-61.

Jansen, M. P., J. A. Foekens, I. L. van Staveren, M. M. Dirkzwager-Kiel, K. Ritstier, M. P. Look, M. E. Meijer-van Gelder, A. M. Sieuwerts, H. Portengen, L. C. J. Dorssers, J. G. M. Klijn, and E. M. Berns. 2005. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. Journal of Clinical Oncology 23:732-740.

Janzen, D. H. 2004. Now is the time. Philosophical Transactions of the Royal Society B 359:731-732.

Janzen, D. H., M. Hajibabaei, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2005. Wedding biodiversity inventory of a large and complex lepidopteria fauna with DNA Barcoding. Philosophical Transactions of the Royal Society B 360:1835-45.

Jones, M., and M. Blaxter. 2006. TaxMan: a taxonomic database manager. BMC Bioinformatics 7:536.

Kantor, Y. I., and A. V. Sysoev. 1989. On the morphology of toxoglossan gastropods lacking a radula, with a description of new species and genus of Turridae. Journal of Molluscan Studies 55:537-549.

Kantor, Y. I., and J. D. Taylor. 2002. Foregut anatomy and relationships of raphitomine gastropods (Gastropoda: Conoidea: Raphitominae). Pages 161-174 *in* Systematics, Phylogeny and Biology of the Neogastropoda (M. Oliverio, and R. Chemello, eds.).

Kelly, R. P., I. Neil Sarkar, D. J. Eernisse, and R. De Salle. 2007. DNA Barcoding using chitons (genus *Mopalia*). Molecular Ecology Notes 7:177-183.

Kerr, K. C. R., M. Y. Stoeckle, C. J. Dove, L. A. Weigt, C. M. Francis, and P. D. N. Hebert. 2007. Comprehensive DNA barcode coverage of North American birds. Molecular Ecology Notes 7:535-543.

Kilburn, R. N. 1983. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 1. Subfamily Turrinae. Annals of the Natal Museum 25:549-585.

Kilburn, R. N. 1985. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 2. Subfamily Clavatulinae. Annals of the Natal Museum 26:417-470.

Kilburn, R. N. 1986. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 3. Subfamily Borsoniinae. Annals of the Natal Museum 27:633-720.

Kilburn, R. N. 1988. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 4. Subfamily Drilliinae, Crassispirinae and Strictispirinae. Annals of the Natal Museum 29:167-320.

Kilburn, R. N. 1991. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 5. Subfamily Taraninae. Annals of the Natal Museum 32:325-339.

Kilburn, R. N. 1992. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 6. Subfamily Mangeliinae, section 1. Annals of the Natal Museum 33:461-575.

Kilburn, R. N. 1993. Turridae (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 6. Subfamily Mangeliinae, section 2. Annals of the Natal Museum 34:317-367.

Kilburn, R. N. 1994. Turridae [*s.l.*] (Mollusca: Gastropoda) of southern Africa and Mozambique. Part 7. Subfamily Crassispirinae, section 2. Annals of the Natal Museum 35:177-228.

Kilburn, R. N. 1995. Turridae (*s.l.*) of southern Africa and Mozambique (Mollusca: Gastropoda: Conoidea) Part 8. Conidae: subfamily Mangeliinae, section 3. Annals of the Natal Museum 36:261-269.

Knapp, S., G. Lamas, E. N. Lughadha, and G. Novarino. 2004. Stability or stasis in the names of organisms: the evolving codes of nomenclature. Philosophical Transactions of the Royal Society B 359:611-622.

Knowles, L. L., and B. C. Carstens. 2007. Delimiting species without monophyletic gene trees. Systematic Biology 56:887-895.

Knowlton, N. 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420:73-90.

Köhler, F. 2007. From DNA taxonomy to barcoding — how a vague idea evolved into a biosystematic tool. Mitteilungen aus dem Museum für Naturkunde in Berlin - Zoologische 83 Supplement:44-51.

Kohonen, T. 1995. Self-Organizing Maps. Springer-Verlag, Berlin.

Krell, F. T. 2004. Parataxonomy vs. taxonomy in biodiversity studies – pitfalls and applicability of 'morphospecies' sorting. Biodiversity and Conservation 13:795-812.

Kress, J. K., E. A. Zimmer, L. A. Weigt, and D. H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences 102:8369-8374.

Kress, W. J., and D. L. Erickson. 2007. A two-locus global DNA Barcode for land plants: the coding rbcL gene complements the non-coding trnHpsbA spacer region. PLoS One 6(e508):1-10.

Lahaye, R., M. van der Bank, D. Bogarin, J. Warner, F. Pupulin, G. Gigot, O. Maurin, S. Duthoit, T. Barraclough, and V. Savolainen. 2008. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences 105:2923-2928.

Lambert, D. M., A. Baker, L. Huynen, P. Haddrath, P. D. N. Hebert, and C. D. Millar. 2005. Is a large-scale DNA-based inventory of ancient life possible? Journal of Heredity 93:1-6.

Lanzone, C., R. A. Ojeda, and M. H. Gallardo. 2007. Integrative taxonomy, systematics and distribution of the genus *Eligmodontia* (Rodentia, Cricetidae, Sigmodontinae) in the temperate Monte Desert of Argentina. Mammalian Biology 72:299-312.

Lecointre, G., and H. Le Guyader. 2001. La classification phylogénétique du vivant. Belin, Paris.

Le Gac, M. & Giraud, T. in press. Existence of a pattern of reproductive character displacement in Basidiomycota but not in Ascomycota. Journal of Evolutionary Biology doi:10.1111/j.1420-9101.2008.01511.x.

Letourneux, Y. M. 2004. Les mollusques du genre *Conus* et les applications thérapeutiques de leurs veninsFaculté de pharmacie, Nantes.

Lherminier, P., and M. Solignac. 2000. L'espèce : définitions d'auteurs. Comptes Rendus de l'Académie des Sciences, Paris 323:153-165.

Lipscomb, D., N. Platnick, and Q. D. Wheeler. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18:65-66.

Little, D. P., and D. W. Stevenson. 2007. A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. Cladistics 23:1-21.

Lorenz, J. G., W. E. Jackson, J. C. Beck, and R. Hanner. 2005. The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. Philosophical Transactions of the Royal Society B 360:1869-1877.

Mace, G. M. 2004. The role of taxonomy in species conservation. Philosophical Transactions of the Royal Society B 359:711-719.

Maddison, W. P., and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55:21-30.

Madeira, S. C., and A. L. Oliveira. 2005. An evaluation of discretization methods for non-supervised analysis of time-series gene expression dataINESC-ID Technical Report 42.

Malaquias, M. A. E., and D. G. Reid. 2008. Systematic revision of the living species of Bullidae (Mollusca: Gastropoda: Cephalaspidea), with a molecular phylogenetic analysis. Zoological Journal of the Linnean Society 153:453-543.

Mallet, J. 2001. The speciation revolution. Journal of Evolutionary Biology 14:887-888.

Mallet, J., N. J. B. Isaac, and G. M. Mace. 2005. Response to Harris and Froufe, and Knapp et al.: Taxonomic inflation. Trends in Ecology and Evolution 20:8-9.

Mallet, J., and K. R. Willmott. 2003. Taxonomy: renaissance or Tower of Babel? Trends in Ecology and Evolution 18:57-59.

Mariaux, J., N. Lutzmann, and J. Stipala. 2008. The two-horned chamaeleons of East Africa. Zoological Journal of the Linnean Society 152:367-391.

Markmann, M., and D. Tautz. 2005. Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences. Philosophical Transactions of the Royal Society B 360:1917-1924.

Marshall, J. C. 2006. Delimiting species: comparing methods for mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. Evolution 60:1050-1065.

Maslov, D. A., S. J. Westenberger, X. Xu, D. A. Campbell, and N. R. Sturm. 2007. Discovery and Barcoding by analysis of spliced leader RNA gene sequences of new isolates of Trypanosomatidae from Heteroptera in Costa Rica and Ecuador. Journal of Eukaryotic Microbiology 54:57-65.

Matz, M. V., and R. Nielsen 2005. A likelihood ratio test for species membership based on DNA sequence data. Philosophical Transactions of the Royal Society B 360:1969-1974.

May, R. M. 2004. Tomorrow's taxonomy: collecting new species in the field will remain the rate-limiting step. Philosophical Transactions of the Royal Society B 359:733-734.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. Pages 381-424 *in* Species: The Units of Biodiversity (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Chapman and Hall, London.

Mayr, E. 1982. The growth of biological thought: Diversity, evolution, and inheritance. Belknap Press of Harvard University Press, Cambridge, Massachusetts.

McKinney, M. L. 2002. Do human activities raise species richness? Contrasting patterns in United States plants and fishes. Global Ecology & Biogeography 11:343-348.

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., and A. Sysoev. 2003. The anatomy of *Zemacies excelsa*, with a description of a new subfamily of Turridae (Gastropoda, Conoidea). Ruthenica 13:81-87.

Meier, R., K. Shiyang, G. Vaidya, and P. K. L. Ng. 2006. DNA Barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Systematic Biology 55:715-728.

Meyer, P. C., and G. Paulay. 2005. DNA Bracoding: error rates based on comprehensive sampling. PLoS Biology 3:1-10.
Mikkelsen, N., C. Schander, and E. Willassen. 2007. Local scale DNA barcoding of bivalves (Mollusca): a case study. Zoologica Scripta 36:455-463.
Miller, B. A. 1970. Studies on the biology of Indo-Pacific *Terebra*University of New Hampshire, Durham.
Miller, B. A. 1971. Feeding mechanisms of the family Terebridae. Annual Reports of the American Malacological Union 1970:72-74.
Miller, S. E. 2007. DNA barcoding and the renaissance of taxonomy. Proceedings of the National Academy of Sciences 104:4775-4776.
Minelli, A. 1999. The names of animals. Trends in Ecology and Evolution 14:462-463.
Minelli, A. 2000. The ranks and the names of species and higher taxa, or a dangerous inertia of the language of natural history. Pages 339-351 *in* Cultures and institutions of Natural History. Essays in the history and philosophy of science (M. T. Ghiselin, and A. E. Leviton, eds.). California Academy of Sciences, San Francisco.
Minelli, A. 2003. The status of taxonomic literature. Trends in Ecology and Evolution 18 75-76.
Monaghan, M. T., M. Balke, T. R. Gregory, and A. P. Vogler. 2005. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. Philosophical Transactions of the Royal Society B 360:1925-33.
Monaghan, M. T., M. Balke, J. Pons, and A. P. Vogler. 2006. Beyond barcodes: complex DNA taxonomy of a South Pacific Island radiation. Proceedings of the Royal Society B 273:887-893.
Morando, M., L. A. Avila, and J. W. Sites. 2003. Sampling strategies for delimiting species: genes, individuals, and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaemidae) in Andean–Patagonian South America. Systematic Biology 52:159-185.
Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51:238-254.
Moritz, C., and C. Cicero. 2004. DNA Barcodes : promise and pitfalls. PLoS Biology 2:1529-1531.
Morrison, J. P. E. 1965. On the families of Turridae *in* The Thirty-First Annual Meeting of the American Malacologial Union, New-York:1-2.
Moura, C. J., D. J. Harris, M. R. Cunha, and A. D. Rogers. 2008. DNA barcoding reveals cryptic diversity in marine hydroids (Cnidaria, Hydrozoa) from coastal and deep-sea environments. Zoologica Scripta 37:93-108.
Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. 1994. Extinction rates can be estimated from molecular phylogenies Philosophical Transactions of the Royal Society B 344:77-82.
Nee, S., A. O. Mooers, and P. H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. The Proceedings of the National Academy of Sciences USA 89:8322-8326.
Newmaster, S. G., A. J. Fazekas, and S. Ragupathy. 2006. DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. Canadian Journal of Botany 84:335-341.
Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16:358-364.
Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6:211-223.
Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin.
Olivera, B. M. 1997. *Conus* venom peptides, receptor and ion channel 385 targets, and drug design: 50 million years of neuropharmocology. Molecular Biology of the Cell 8:2101-2109.
Olivera, B. M. 2002. *Conus* venom peptides: reflections from the biology of clades and species. Annual Review of Ecology and Systematics 33:25-47.
Olivera, B. M. 2006. *Conus* peptides: Biodiversity-based discovery and exogenomics. Journal of Biological Chemistry 42:31173-31177.
Olivera, B. M., and R. W. Teichert. 2007. Diversity of the neurotoxic *Conus* peptides: a model for concerted pharmacological discovery. Molecular Interventions 7:251-260.
Page, T. J., S. C. Choy, and J. M. Hughes. 2005. The taxonomic feedback loop: symbiosis of morphology and molecules. Biology Letters 1:139-142.
Palumbi, S. R. 1994. Genetic divergence, reproductive isolation and marine speciation. Annual Review of Ecology and Systematics 25:547-572.
Palumbi, S. R., and H. A. Lessios. 2005. Evolutionary animation: How do molecular phylogenies compare to Mayr's reconstruction of speciation patterns in the sea? Proceedings of the National Academy of Sciences 102:6566-6572.
Paquin, P., and M. Hedin. 2004. The power and perils of 'molecular taxonomy': a case study of eyeless and endangered *Cicurina* (Araneae: Dictynidae) from Texas caves. Molecular Ecology 13:3239-3255.
Payne, R. B., and M. D. Sorenson. 2007. Integrative systematics at the species level: plumage, songs and molecular phylogeny of quailfinches *Ortygospiza*. Bulletin of the British Ornithologists' Club 127:4-26.

Pfenninger, M., Cordellier, M. & Streit, B. 2006. Comparing the efficacy of morphologic and DNA-based taxonomy in the freshwater gastropod genus *Radix* (Basommatophora, Pulmonata). BMC Evolutionary Biology 6: 100.

Pimm, S., P. Raven, A. Peterson, C. H. Sekercioglu, and P. R. Ehrlich. 2006. Human impacts on the rates of recent, present, and future bird extinctions. Proceedings of the National Academy of Sciences 103:10941-10946.

Pons, J., T. G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S. Kamoun, W. D. Sumlin, and A. P. Vogler. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55:595-609.

Poulakakis, N., and S. Sfenthourakis. 2008. Molecular phylogeny and phylogeography of the Greek populations of the genus *Orthometopon* (Isopoda, Oniscidea) based on mitochondrial DNA sequences. Zoological Journal of the Linnean Society 152:707-715.

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 2:1-192.

Powell, A. W. B. 1966. The molluscan families Speightiidae and Turridae. An evaluation of the valid taxa, both recent and fossil, with lists of characteristics species. Bulletin of the Auckland Institute and Museum 5:5-184.

Powers, T. 2004. Nematode molecular diagnostics: from bands to barcodes. Annual Review of Phytopathology 42:367-383.

Proudlove, G., and P. J. Wood. 2003. The blind leading the blind: cryptic subterranean species and DNA taxonomy. Trends in Ecology and Evolution 18:272-273.

Rach, J., R. DeSalle, I. N. Sarkar, B. Schierwater, and H. Hadrys. 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proceedings of the Royal Society B 275:237-247.

Rash, S., and D. Gusfield. 2002. String Barcoding: uncovering optimal virus signatures *in* Proceedings of the sixth annual international conference on Computational biology, Washington, DC:254 - 261.

Ratnasingham, S., and P. D. N. Hebert. 2007. BOLD : The Barcode of Life Data System. Molecular Ecology Notes 7:355-364.

Raven, P. H. 2004. Taxonomy: where are we now? Philosophical Transactions of the Royal Society B 359:729–730.

Remigio, E. A., and P. D. N. Hebert. 2003. Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. Molecular Phylogenetics and Evolution 29:641-647.

Rezac, M., S. Pekar, and J. Johannesen. 2008. Taxonomic review and phylogenetic analysis of central European *Eresus* species (Araneae: Eresidae). Zoologica Scripta 37:263-287.

Rissler, L. J., and J. J. Apodaca. 2007. Ecological niche models and phylogeography help uncover cryptic biodiversity. Systematic Biology 56:924-942.

Roca, A. L., N. Georgiadis, J. Pecon-Slattery, and S. J. O'Brien. 2001. Genetic evidence for two species of elephant in Africa. Science 293:1473-1477.

Roe, A. D., and F. A. H. Sperling. 2007. Population structure and species boundary delimitation of cryptic *Dioryctria* moths: an integrative approach. Molecular Ecology 16:3617-3633.

Rohlf, F. J. 1996. TpsDig. State University of New-York at Stony Brook, http://life.bio.sunysb.edu/morph/.

Rosenberg, N. A. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly and polyphyly in a coalescent model. Evolution 57:1465-1477.

Rosenberg, N. A., and R. Tao. 2008. Discordance of species trees with their most likely gene trees: the case of five taxa. Systematic Biology 57:131-140.

Ross, H. A., S. Murugan, and W. L. S. Li. 2008. Testing the reliability of genetic methods of species identification via simulation. Systematic Biology 57:216-230.

Rowley, D., J. A. Coddington, M. W. Gates, A. L. Norrbom, R. A. Ochoa, N. J. Wandenberg, and M. H. Greenstone. 2007. Vouchering DNA-barcoded specimens: test of a nondestructive extraction protocol for terrestrial arthropods. Molecular Ecology Notes 7:915-924.

Rubinoff, D. 2006. DNA Barcoding evolves into the familiar. Conservation Biology 20:1548-1549.

Rubinoff, D., S. Cameron, and K. P. Will. 2006. Are plant DNA barcodes a search for the Holy Grail? Trends in Ecology and Evolution 21:1-2.

Sáez, A. G., and E. Lozano. 2005. Body doubles - Cryptic species: as we discover more examples of species that are morphologically indistinguishable, we need to ask why and how they exist. Nature 433:111.

Salter Kubatko, L., and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56:17-24.

Samadi, S. 2008. La spéciation *in* Biologie Evolutive (F. Thomas, T. Lefèvre, and R. M. De Boeck, eds.). in press.

Samadi, S., and A. Barberousse. 2006. The tree, the network, and the species. Biological Journal of the Linnean Society 89:509-521.

Sanders, K. L., A. Malhorta, and R. S. Thorpe. 2006. Combining molecular, morphological and ecological data to infer species boundaries in a cryptic tropical pitviper. Biological Journal of the Linnean Society 87:343-364.

Saunders, G. W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Philosophical Transactions of the Royal Society B 360:1879-1888.

Savage, J. M., and J. I. Watling. 2008. Not so rare snakes: a revision of the *Geophis sieboldi* group (Colubridae: Dipsadinae) in lower Central America and Colombia. Zoological Journal of the Linnean Society 153:561-599.

Savolainen, V., R. S. Cowan, A. P. Vogler, G. K. Roderick, and R. Lane. 2005. Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society B 360:1805-1811.

Schander, C., and E. Willassen. 2005. What can biological barcoding do for marine biology? Marine Biology Research 1:79-83.

Scheffer, S. J., M. L. Lewis, and R. C. Joshi. 2006. DNA Barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. Annals of the Entomological Society of America 99:204-210.

Scicluna, S. M., B. Tawari, and G. C. Clark. 2006. DNA barcoding of blastocystis. Protist 157:77-85.

Scoble, M. J. 2004. Unitary or unified taxonomy? Philosophical Transactions of the Royal Society B 359:699-710.

Seberg, O., C. H. Humphries, S. Knapp, D. W. Stevenson, G. Petersen, N. Scharff, and N. M. Andersen. 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. Trends in Ecology and Evolution 18:63-65.

Seifert, K. A., R. A. Samson, J. R. deWaard, J. Houbraken, A. Lévesque, J. M. Monclavo, G. Louis-Seize, and P. D. N. Hebert. 2007. Prospects for fungus identification using CO1 DNA barcodes, with *Penicillium* as a test case. Proceedings of the National Academy of Sciences 104:3901-3906.

Shaffer, H. B., and R. C. Thomson. 2007. Delimiting species in recent radiations. Systematic Biology 56:896-906.

Shearer, T. L., and M. A. Coffroth. 2008. Barcoding corals: limited by interspecific divergence, not intraspecific variation. Molecular Ecology Resources 8:247-255.

Shuto, T. 1974. Larval ecology of prosobranch gastropods and its bearing on biogeography and paleontology. Lethaia 7:239-256.

Siddall, M. E., and R. B. Budinoff. 2005. DNA-barcoding evidence for widespread introductions of a leech from the South American *Helobdella triserialis* complex. Conservation Genetics 6:467-472.

Sites, J. W., and J. C. Marshall. 2003. Delimiting species: a renaissance issue in systematic biology. Trends in Ecology and Evolution 19:462-470.

Skage, M., and C. Schander. 2007. DNA from formalin-fixed tissue: extraction or repair? That is the question. Marine Biology Research 3:289 295.

Smith, A. M., B. L. Fisher, and P. D. N. Hebert. 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society B 360:1825-1834.

Smith, M. A., N. A. Poyarkov Jr., and P. D. N. Hebert. 2008a. CO1 DNA barcoding amphibians: take the chance, meet the challenge. Molecular Ecology Resources 8:235-246.

Smith, M. A., N. E. Woodley, D. H. Janzen, W. Hallwachs, and P. D. N. Hebert. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). Proceedings of the National Academy of Sciences 103 3657-3662.

Smith, P. J., S. M. McVeagh, and D. Steinke. 2008b. DNA barcoding for the identification of smoked fish products. Journal of Fish Biology 72:464-471.

Spies, I. B., S. Gaichas, D. E. Stevenson, J. W. Orr, and M. F. Canino. 2006. DNA-based identification of Alaska skates (Amblyraja, Bathyraja and Raja:Rajidae) using cytochromec oxidase subunit I (coI) variation. Journal of Fish Biology 69 (Supplement B):283-292.

Stahls, G., and E. Savolainen. 2008. MtDNA COI barcodes reveal cryptic diversity in the *Baetis vernus* group (Ephemeroptera, Baetidae). Molecular Phylogenetics and Evolution 46:82-87.

Statnikov, A., C. F. Aliferis, I. Tsamardinos, D. Hardin, and S. Levy. 2005. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics 21:631-643.

Steinke, D., M. Vences, W. Salzburger, and A. Meyer. 2005. TaxI: a software tool for DNA barcoding using distance methods. Proceedings of the Royal Society B 360:1975-1980.

Steinley, D. 2003. Local optima in K-Means clustering: what you don't know may hurt you. Psychological Methods 8:294-304.

Steinley, D. 2007. Initializing K-means batch clustering: a critical evaluation of several techniques. Journal of Classification 24:99-121.

Stoeckle, M. 2003. Taxonomy, DNA, and the Bar Code of Life. BioScience 53:2-3.

Summerbell, R. C., C. A. Lévesque, K. A. Seifert, M. Bovers, J. W. Fell, M. R. Diaz, T. Boekhout, G. S. de Hoog, J. Stalpers, and P. W. Crous. 2005. Microcoding: the second step in DNA barcoding. Philosophical Transactions of the Royal Society B 360:1897-1903.

Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Vogler. 2003. A plea for DNA taxonomy. Trends in Ecology and Evolution 18:70-74.

Taylor, J. D., Y. I. Kantor, and A. V. Sysoev. 1993. Foregut anatomy, feedings mechanisms and classification of the Conoidea (= Toxoglossa)(Gastropoda). Bulletin of the Natural History Museum of London (Zoology) 59:125-170.

Terryn, Y. 2007. A collectors guide to recent Terebridae (Mollusca: Neogastropoda). ConchBooks, Hackenheim.

Thalmann, O., J. Hebler, H. N. Poinar, S. Pääbo, and L. Vigilant. 2004. Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. Molecular Ecology 13:321-335.

Thiele, J. 1929. Handbuch der systematischen weichtierkunde. G. Fischer, Jena.

Thomas, C. D. 2007. A sixth mass extinction? Nature 450:349.

Trewick, S. A. 2008. DNA Barcoding is not enough: mismatch of taxonomy and genealogy in New Zealand grasshoppers (Orthoptera: Acrididae). Cladistics 24 240-254.

Tryon, G. W. 1884. Manual of Conchology; Structural and systematic, with illustrations of the species. Vol. VI. Conidae, Pleurotomidae. Tryon, G. W., Philadelphia.

Tucker, J. K. 2004. Catalogue of recent and fossil turrids (Mollusca: Gastropoda). Zootaxa 682:1-1295.

van Velzen, R., F. T. Bakker, and J. J. A. van Loon. 2007. DNA barcoding reveals hidden species diversity in *Cymothoe* (Nymphalidae). Proceedings of the Netherlands Entomological Society Meeting 18:95-103.

Vences, M., M. Thomas, R. M. Bonett, and D. R. Vieites. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philosophical Transactions of the Royal Society B 360:1859-1868.

Vogler, A. P., and M. T. Monaghan. 2007. Recent advances in DNA taxonomy. Journal of Zoological Systematics and Evolutionary Research 45:1-10.

Wang, J. Y., Frasier, T. R., Yang, S. C. & White, B. N. 2008. Detecting recent speciation events: the case of the finless porpoise (genus *Neophocaena*). Heredity, 101: 145-155.

Ward, R. D., B. H. Holmes, T. W. William, and P. R. Last. 2008. DNA barcoding Australasian chondrichthyans: results and potential uses in conservation. Marine and Freshwater Research 59:57-71.

Ward, R. D., T. S. Zemlak, B. H. Innes, P. R. Last, and P. D. N. Hebert. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B 360:1847-1857.

Wheeler, Q. D. 2005. Losing the plot: DNA ''barcodes'' and taxonomy. Cladistics 21:405-407.

Wheeler, Q. D., and R. Meier. 2000. Species concepts and phylogenetic therory: a debate. Columbia University Press, New-York.

Whiteman, N. K., D. Santiago-Alarcona, K. P. Johnson, and P. G. Parker. 2004. Differences in straggling rates between two genera of dove lice (Insecta: Phthiraptera) reinforce population genetic and cophylogenetic patterns. International Journal for Parasitology 34:1113-1119.

Wicker, N., D. Dembele, W. Raffelsberger, and P. Poch. 2002. Density of points clustering, application to transcriptomic data analysis. Nucleic Acids Research 30:3392-4000.

Wicker, N., G. R. Perrin, J. C. Thierry, and O. Poch. 2001. Secator: a program for inferring protein subfamilies from phylogenetic trees. Molecular Biology and Evolution 18:1435-1441.

Wiemers, M., and K. Fiedler. 2007. Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). Frontiers in Zoology 4:1-16.

Wiens, J. J. 2007. Species delimitation: new approaches for discovering diversity. Systematic Biology 56:875-878.

Will, K. P., B. D. Mishler, and Q. D. Wheeler. 2005. The perils of DNA Barcoding and the need for integrative taxonomy. Systematic Biology 54:844-851.

Will, K. P., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20:47-55.

Wilson, E. O. 2004. Taxonomy as a fundamental discipline. Philosophical Transactions of the Royal Society B 359:739.

Yassin, A. 2007. Ecological genetics and systematics of the genus *Zaprionus* (Diptera: Drosophilidae): molecular and morphological approachesUniversité Paris XI, Orsay.

Zhang, A. B., D. S. Sikes, C. Muster, and S. Q. Li. 2008. Inferring species membership using DNA sequences with back-propagation neural networks. Systematic Biology 57:202-215.

Zink, R. M., and G. F. Barrowclough. 2008. Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology 17:2107-2121.

ANNEXE

		Aurora 2007	BOA 1 (2005)	EBISCO (2005)	Norfolk 2 (2003)	Panglao 2004	Panglao 2005	Salomon 2 (2004)	Salomon 3 (2007)	Santo 2006	Taïwan 2004	Autres [1]	
Durée (jours)		16	15	18	17	40	10	19	16	45	8		
Profondeur moyenne (min-max)		580 (42-2253)	510 (72-1273)	470 (50-1147)	540 (86-1434)	30 (0-400)	530 (64-2323)	510 (94-1218)	510 (90-1545)	200 (0-722)	434(48-1689)		
Effort de collecte		+++	++	++	+	+	+	++	++	+++	+		
Spécimens récoltés		788	185	141	92	498	368	462	361	1681	7	145	**4728** [2]
ADN extraits		311	163	94	61	265	300	214	133	783	7	83	**2414**
Séquences obtenues	**COI**	292	143	78	39	207	225	128	98	693	4	25	**1932**
	28SD1	260	82	69	15	138	175	123	33	531	2	3	**1431**
	16S	25	7	23		72	22	20	1	141		12	**323**
	12S	27	6	26	2	83	32	25	1	144		14	**360**
	H3		11	18	1	50	25	21		4			**130**
	28SD7		11	16		44	23	16		4			**114**
	18S1		19	59	1	128	67	81		4			**359**
	18S2		11	22	1	51	28	24		4			**141**
	18S3		11	22	1	51	28	24		4			**141**
	ITS2	18	1	2		66	5	11	11	2			**116**
											TOTAL		**5047** [3]

Description des 10 missions qui ont permis la récolte de la majorité des spécimens utilisés : durée, profondeur et effort de collecte (+ : seul un ou quelques spécimens par morphoespèces ont été conservés pour les analyses moléculaires ; ++ : tous les spécimens ont été conservés ; +++ : tous les spécimens ont été conservés, et chacun a été fixé séparément, après extraction du tissu de la coquille pour une meilleure fixation de l'ADN) et nombre de spécimens récoltés. Pour chaque mission est également précisé le nombre de spécimens dont l'ADN a été extrait, et le nombre de séquences obtenues pour chaque gène. [1] : spécimens récoltés lors de missions plus anciennes (Bathus 4, Musorstom 3, 4 et 8, SUVA 4 et Taïwan 2001) ou fournis par plusieurs collaborateurs : Serge Gofas (spécimens collectés en Afrique), Bruce Marshall (Muséum de Nouvelle-Zélande), Jon-Arne Sneli (Trondhjem biologiske stasjon), Jacques Pelorce (côtes françaises) et Anders Warén (Swedish Museum of Natural History). [2] : dont 4079 Turridae, 342 Conus, 268 Terebridae et 39 gastéropodes non-Conoidea. [3] : ce nombre inclus les 450 séquences réalisées par Mandë Holford.

Taxonomie intégrative des Turridae : phylogénie, délimitation d'espèces et barcoding

Résumé

Les Turridae constituent un groupe de gastéropodes marins très diversifié, et dont une grande partie des espèces sont encore non décrites. Les approches taxonomiques traditionnelles se heurtent à des problèmes liés à la variabilité de la coquille, difficile à interpréter, aussi bien au niveau spécifique (α-taxonomie) qu'aux niveaux supérieurs (phylogénie). Afin d'accélérer le rythme de description des espèces, l'objectif de la thèse est de renouveler les pratiques taxonomiques au sein des Turridae en adaptant de nouvelles méthodes et de nouveaux outils. La taxonomie devient intégrative, et associe caractères moléculaires, morphologiques et écologiques mais également les méthodes issues de la phylogénie et de la génétique des populations. Le projet barcoding participe à cette nouvelle dynamique, en offrant une méthode d'analyse standardisée de la diversité génétique.

Dans une première partie, plusieurs gènes mitochondriaux et nucléaires ont été séquencés pour reconstruire la phylogénie des Conoidea, groupe incluant les Turridae mais également le genre *Conus* et les Terebridae. Les résultats obtenus mettent en évidence des relations phylogénétiques inédites, et permettent de proposer une nouvelle classification pour le groupe.

Dans une seconde partie, l'approche d'α-taxonomie intégrative est détaillée. Elle repose sur le cadre conceptuel suivant : les espèces sont des segments de lignées évolutives. De plus, les espèces acquièrent au cours de leur évolution différentes propriétés qui peuvent être utilisées pour proposer des hypothèses de délimitation d'espèces, notamment en utilisant des caractères moléculaires. Dans une première étape, le fragment barcode du gène COI a été séquencé. Il a été utilisé pour identifier des pontes de gastéropodes marins et pour tester les hypothèses de délimitation d'espèces déjà disponibles pour le genre *Eumunida* (Crustacea).

Le barcode COI peut également permettre de proposer de nouvelles hypothèses de délimitation d'espèces. Principalement deux méthodes exploratoires, dont l'une a été développée au cours de cette thèse, ont été utilisées pour délimiter des groupes de spécimens. Ces groupes ont ensuite été testés dans une seconde étape en analysant systématiquement un gène nucléaire (28S ou ITS2), mais également la variabilité morphologique. Cette approche de taxonomie intégrative a été appliquée à quatre groupes de Turridae, pour lesquels un grand nombre d'échantillons collectés dans le Pacifique étaient disponibles : les genres *Bathytoma*, *Benthomangelia* et *Xenuroturris*, et la sous-famille des Turrinae. Au total, 111 espèces ont été délimitées au sein de ces groupes, dont environ 70 seraient nouvelles pour la science.

Les caractères moléculaires ont donc été utiles pour éclaircir les relations phylogénétiques au sein de la famille des Turridae. Associés à l'analyse d'autres caractères, ils ont également permis d'analyser la diversité spécifique de ce groupe. La méthodologie mise au point au cours de la thèse est donc adaptée à un groupe hyperdiversifé comme les Turridae, et accélère le rythme de découverte et de description de nouvelles espèces.

Integrative taxonomy in the Turridae: phylogeny, species delimitation and barcoding

Abstract

The Turridae constitute a hyperdiverse group of marine gastropods within which most species are still undescribed. Traditional taxonomic approaches are challenged by the difficult interpretability of shell variability at both specific (α-taxonomy) and superior levels (phylogeny). In order to speed up the rate of species description, our goal was to replenish taxonomical practices within Turrinae by applying new tools and methods. Taxonomy becomes integrative, combining molecular, morphological and ecological characters, but also methods form phylogenetic and population genetics methods. The barcoding project is included in this new dynamics, providing a standardized method of genetic diversity analysis.

In a first part, several mitochondrial and nuclear genes were sequenced to reconstruct the phylogeny of Conoidea, a group including the Turridae but also the genus *Conus* and the Terebridae. Our results evidenced new phylogenetic relationships and allowed us to propose a new classification of the group.

In a second part, the integrative α-taxonomy approach was detailed. The latter lies on the following conceptual framework: species are segments of evolutionary lineages. Furthermore, species acquire several properties during their evolution that can be used to propose hypotheses of species delimitation, in particular using molecular characters. As a first step, the barcode fragment of the COI gene was sequenced. It was used to identify egg-capsules of marine gastropods and to test already available hypotheses of species delimitation in the genus *Eumunida* (Crustacea).

The COI barcode can also be used to propose new hypotheses of species delimitation. Mainly two exploratory methods, of which one was developed during this Ph.D., were used to delimit groups of specimens. These groups were then tested in a second step through analysis of a nuclear gene (28S or ITS2), but also morphological variability. This integrative taxonomy approach was applied to four groups of Turridae for which a great number of samples from the Pacific was available: the genera *Bathytoma*, *Benthomangelia* and *Xenuroturris*, and the subfamily Turrinae. Finally, 111 species were delimited, of which 70 are potentially new for science.

Molecular characters were thus useful to clarify phylogenetic relationships within the family Turridae. Combined with other characters, they also allowed the analysis of species diversity within the group. The methodology developed during this Ph.D. was thus adapted to a hyper-diversified group such as the Turridae, and accelerated the rate of species discovery and description.

Taxonomie intégrative des Turridae : phylogénie, délimitation d'espèces et barcoding

Résumé

Les Turridae constituent un groupe de gastéropodes marins très diversifié, et dont une grande partie des espèces sont encore non décrites. Les approches taxonomiques traditionnelles se heurtent à des problèmes liés à la variabilité de la coquille, difficile à interpréter, aussi bien au niveau spécifique (α-taxonomie) qu'aux niveaux supérieurs (phylogénie). Afin d'accélérer le rythme de description des espèces, l'objectif de la thèse est de renouveler les pratiques taxonomiques au sein des Turridae en adaptant de nouvelles méthodes et de nouveaux outils. La taxonomie devient intégrative, et associe caractères moléculaires, morphologiques et écologiques mais également les méthodes issues de la phylogénie et de la génétique des populations. Le projet barcoding participe à cette nouvelle dynamique, en offrant une méthode d'analyse standardisée de la diversité génétique.

Dans une première partie, plusieurs gènes mitochondriaux et nucléaires ont été séquencés pour reconstruire la phylogénie des Conoidea, groupe incluant les Turridae mais également le genre *Conus* et les Terebridae. Les résultats obtenus mettent en évidence des relations phylogénétiques inédites, et permettent de proposer une nouvelle classification pour le groupe.

Dans une seconde partie, l'approche d'α-taxonomie intégrative est détaillée. Elle repose sur le cadre conceptuel suivant : les espèces sont des segments de lignées évolutives. De plus, les espèces acquièrent au cours de leur évolution différentes propriétés qui peuvent être utilisées pour proposer des hypothèses de délimitation d'espèces, notamment en utilisant des caractères moléculaires. Dans une première étape, le fragment barcode du gène COI a été séquencé. Il a été utilisé pour identifier des pontes de gastéropodes marins et pour tester les hypothèses de délimitation d'espèces déjà disponibles pour le genre *Eumunida* (Crustacea).

Le barcode COI peut également permettre de proposer de nouvelles hypothèses de délimitation d'espèces. Principalement deux méthodes exploratoires, dont l'une a été développée au cours de cette thèse, ont été utilisées pour délimiter des groupes de spécimens. Ces groupes ont ensuite été testés dans une seconde étape en analysant systématiquement un gène nucléaire (28S ou ITS2), mais également la variabilité morphologique. Cette approche de taxonomie intégrative a été appliquée à quatre groupes de Turridae, pour lesquels un grand nombre d'échantillons collectés dans le Pacifique étaient disponibles : les genres *Bathytoma*, *Benthomangelia* et *Xenuroturris*, et la sous-famille des Turrinae. Au total, 111 espèces ont été délimitées au sein de ces groupes, dont environ 70 seraient nouvelles pour la science.

Les caractères moléculaires ont donc été utiles pour éclaircir les relations phylogénétiques au sein de la famille des Turridae. Associés à l'analyse d'autres caractères, ils ont également permis d'analyser la diversité spécifique de ce groupe. La méthodologie mise au point au cours de la thèse est donc adaptée à un groupe hyperdiversifé comme les Turridae, et accélère le rythme de découverte et de description de nouvelles espèces.

Integrative taxonomy in the Turridae: phylogeny, species delimitation and barcoding

Abstract

The Turridae constitute a hyperdiverse group of marine gastropods within which most species are still undescribed. Traditional taxonomic approaches are challenged by the difficult interpretability of shell variability at both specific (α-taxonomy) and superior levels (phylogeny). In order to speed up the rate of species description, our goal was to replenish taxonomical practices within Turrinae by applying new tools and methods. Taxonomy becomes integrative, combining molecular, morphological and ecological characters, but also methods form phylogenetic and population genetics methods. The barcoding project is included in this new dynamics, providing a standardized method of genetic diversity analysis.

In a first part, several mitochondrial and nuclear genes were sequenced to reconstruct the phylogeny of Conoidea, a group including the Turridae but also the genus *Conus* and the Terebridae. Our results evidenced new phylogenetic relationships and allowed us to propose a new classification of the group.

In a second part, the integrative α-taxonomy approach was detailed. The latter lies on the following conceptual framework: species are segments of evolutionary lineages. Furthermore, species acquire several properties during their evolution that can be used to propose hypotheses of species delimitation, in particular using molecular characters. As a first step, the barcode fragment of the COI gene was sequenced. It was used to identify egg-capsules of marine gastropods and to test already available hypotheses of species delimitation in the genus *Eumunida* (Crustacea).

The COI barcode can also be used to propose new hypotheses of species delimitation. Mainly two exploratory methods, of which one was developed during this Ph.D., were used to delimit groups of specimens. These groups were then tested in a second step through analysis of a nuclear gene (28S or ITS2), but also morphological variability. This integrative taxonomy approach was applied to four groups of Turridae for which a great number of samples from the Pacific was available: the genera *Bathytoma*, *Benthomangelia* and *Xenuroturris*, and the subfamily Turrinae. Finally, 111 species were delimited, of which 70 are potentially new for science.

Molecular characters were thus useful to clarify phylogenetic relationships within the family Turridae. Combined with other characters, they also allowed the analysis of species diversity within the group. The methodology developed during this Ph.D. was thus adapted to a hyper-diversified group such as the Turridae, and accelerated the rate of species discovery and description.