Piercing Fishes: Porin Expansion and Adaptation to Hematophagy in the Vampire Snail *Cumia reticulata*

Marco Gerdol,[1] Manuela Cervelli,[2] Marco Oliverio,[3] and Maria Vittoria Modica*[,4,5]

[1]Department of Life Sciences, Trieste University, Italy

[2]Department of Biology, Roma Tre University, Italy

[3]Department of Biology and Biotechnologies "Charles Darwin", Sapienza University, Roma, Italy

[4]Department of Integrative Marine Ecology, Stazione Zoologica Anton Dohrn, Naples, Italy

[5]UMR5247, University of Montpellier, France

***Corresponding author:** E-mail: mariavittoria.modica@szn.it.

Associate Editor: Nicolas Vidal

Abstract

Cytolytic pore-forming proteins are widespread in living organisms, being mostly involved in both sides of the host–pathogen interaction, either contributing to the innate defense or promoting infection. In venomous organisms, such as spiders, insects, scorpions, and sea anemones, pore-forming proteins are often secreted as key components of the venom. Coluporins are pore-forming proteins recently discovered in the Mediterranean hematophagous snail *Cumia reticulata* (Colubrariidae), highly expressed in the salivary glands that discharge their secretion at close contact with the host. To understand their putative functional role, we investigated coluporins' molecular diversity and evolutionary patterns. Coluporins is a well-diversified family including at least 30 proteins, with an overall low sequence similarity but sharing a remarkably conserved actinoporin-like predicted structure. Tracking the evolutionary history of the molluscan porin genes revealed a scattered distribution of this family, which is present in some other lineages of predatory gastropods, including venomous conoidean snails. Comparative transcriptomic analyses highlighted the expansion of porin genes as a lineage-specific feature of colubrariids. Coluporins seem to have evolved from a single ancestral porin gene present in the latest common ancestor of all Caenogastropoda, undergoing massive expansion and diversification in this colubrariid lineage through repeated gene duplication events paired with widespread episodic positive selection. As for other parasites, these findings are congruent with a "one-sided arms race," equipping the parasite with multiple variants in order to broaden its host spectrum. Overall, our results pinpoint a crucial adaptive role for coluporins in the evolution of the peculiar trophic ecology of vampire snails.

Key words: **transcriptomics, Gastropoda, hematophagy, cytolysins.**

Article

Introduction

The presence of cytolytic pore-forming proteins has been documented from prokaryotes to eukaryotes, from plants to Metazoa. Despite their wide diversity, all these molecules are amphiphilic and contain hydrophobic regions directly involved in the formation of a pore in the target cell membrane (Iacovache et al. 2010). They are secreted as water-soluble proteins that do not have an endogenous target but are rather used for defense (against a predator or an invading pathogen) or attack (against a host or a prey) purposes (Andreeva-Kovalevskaya et al. 2008; Iacovache et al. 2008). Upon the establishment of an interaction with their biological target, due to either ingestion, injection, tissue colonization, or feeding, they are translocated into the cell membrane of the target organisms to form a pore (Parker and Feil 2005). Their toxic action is due to the disturbance of the ionic homeostasis and osmotic equilibrium of the cell and may eventually lead to cell lysis and, on a broader scale, tissue necrosis (Bischofberger et al. 2009).

Pore formers play important roles in host–pathogen interactions: many pathogenic microorganisms—bacteria or parasites—produce pore-forming proteins to promote infections (Parker and Feil 2005; Andreeva-Kovalevskaya et al. 2008; Bischofberger et al. 2009). On the other hand, pore-forming cytolysins are involved in host defense against enemies and are powerful antimicrobial compounds: as an example, perforins secreted by cytolytic lymphocytes (CTL and NK) of the mammalian immune system are able to kill both microorganisms and aberrant host cells such as cancer cells or cells damaged by virus (Voskoboinik and Trapani 2006; Hadders et al. 2007; Rosado et al. 2007).

Due to their cytotoxicity, pore forming toxins have been often recruited as components of the venom in many venomous taxa, including snakes, insects, scorpions, and sea anemones, where they contribute to prey capture and digestion (Parker and Feil 2005).

An intensive research effort has been dedicated to actinoporins from sea anemones, characterized by a single domain

of about 20 kDa lacking cysteine residues and with a high isoelectric point (>9.5) (Álvarez et al. 2009; Frazão et al. 2012; Valle et al. 2015; Podobnik and Anderluh 2017). Actinoporins belong to the alpha-pore forming toxins family (PFTs), since their structure cross the membrane by alpha-helical domains, containing a hydrophobic helical hairpin (Parker and Feil 2005; Iacovache et al. 2010; Monastyrnaya et al. 2010). They are unique as their ability to form pores evolved from the combination of an amphipathic alpha-helical structure similar to melittin, a beta-sandwich structures comprising several loops, and an interfacial binding site with great specificity for lipids (Rojko et al. 2016). Actinoporins perforate the cell membrane in a multistep process still not completely clarified, involving: 1) the binding of soluble monomers through recognition of sphingomyelin (or in some cases cholesterol; Parker and Feil 2005) on the cell membrane, via a specific aromatic cluster, phosphocholine binding sites, and basic regions (Hong et al. 2002); 2) the oligomerization and separation of the hydrophobic N-terminal, including the alpha helix 1, from the beta-sandwich region, to contact the lipid–water interface; and 3) the formation of the cation selective hydrophilic pores by translocation of the N terminal from several monomers in the lipid bilayer (Hong et al. 2002; Malovrh et al. 2003; Kristan et al. 2004; Mechaly et al. 2011; Rojko et al. 2016). Pore formation by actinoporins may cause hemolysis, while their pharmacological effects include cardiotoxicity, coronary vasospasm, and respiratory arrest (Anderluh and Macek 2002).

Their cytotoxicity, maintained at nanomolar ranges, made actinoporins, together with other pore forming proteins, appealing targets for biomedical development; their antitumoral activity was confirmed for several human cancers (reviewed in Mariottini and Pane 2013) although the underlying mechanisms in many cases have not been elucidated yet.

Actinoporin-like sequences are however not a unique prerogative of anthozoans within Metazoa. Indeed, several actinoporin-like genes have been evidenced by genome sequencing approaches in teleost fishes, although the functional role of the encoded proteins in these vertebrates is presently unclear, as they possess neither sphingomyelin-specificity, nor pore-forming properties (Gutiérrez-Aguirre et al. 2006). Actinoporin-like sequences are also found in bivalve molluscs. In the freshwater clam *Corbicula japonica* a single protein named clamlysin has been partly functionally characterized, revealing marked sphingomyelin-binding features, and implying a certain degree of functional homology with anthozoan toxins (Takara et al. 2011). Actinoporin-like genes, with uncharacterized function, are also present in the brachiopod *Lingula anatina* (Gerdol et al. 2018), whereas a few transcripts encoding members of this family have been identified in the transcriptomes of the venom gland of predatory bloodworms (Annelida, Polychaeta) and the venom fangs of the venomous remipede crustacean *Xibalbanus tulumensis* (Von Reumont et al. 2014).

In gastropods, toxins similar to actinoporins were firstly described in the predatory hairy triton *Monoplex echo* (Shiomi et al. 1994; Shiomi et al. 2002; Kawashima et al. 2003) and recently found in the salivary glands (SGs) of the closely related giant triton *Charonia tritonis* (Bose et al. 2017). These toxins, named echotoxins, exhibit both hemolytic and lethal activities in mice and are inhibited by gangliosides, suggesting a mechanism of action involving the gangliosides binding. With this respect, the behavior of echotoxins differ from that of sphingomyelin-binding actinoporins, rather resembling some bacterial hemolysins that are able to lyse erythrocytes (Ozawa et al. 1994; Kawashima et al. 2003).

Echotoxin-like sequences were subsequently reported in conoidean snails, both in the venom and in SGs, in some cases with multiple isoforms. Nine conoporin isoforms, with high sequence similarity but differing for their oxidative state, have been retrieved from *Conus geographus* venom gland with high expression levels, while the presence of seven isoforms in *Conus consors* injected venom was attributed to salivary secretion (Leonardi et al. 2012; Safavi-Hemami et al. 2014). A single tereporin sequence was also found in two species of auger snails (unpublished data). However, none of the actinoporin-like sequences identified in Conoidea has been so far functionally characterized, and these proteins have been simply assumed as being part of the toxin cocktail produced by predatory gastropods according to their expression profile.

More recently, we identified a high number of actinoporin-like sequences in the SG of the vampire snail *Cumia reticulata* (Neogastropoda, Colubrariidae) through a differential expression analysis on tissue-specific transcriptome data (Modica et al. 2015). Species of marine snails in the family Colubrariidae are specialized to feed on the blood on sleeping fishes (Johnson et al. 1995; Bouchet and Perrine 1996; Oliverio and Modica 2010). The snail contacts the fish with its long and extensible proboscis, gains access to the blood vessels, and feeds passively, taking advantage of the blood pressure of the fish. Hematophagous feeding is allowed by the active secretion of antihemostatics that act both externally, in the host, and internally in snail's digestive system, and facilitated by the reversible anesthetization of the fish until the end of the meal (Modica et al. 2015). Since the radula—the buccal teeth equipment used by most gastropods as a feeding tool—is vestigial in all Colubrariidae, and likely unable to pierce the host tegument alone, bioactive compounds that allow tissue degradation can play a central role in feeding.

Here, we aim at investigating the structural organization, molecular diversity, and evolution of the porins of *C. reticulata*, to shed light on their adaptive value in the establishment of the peculiar trophic habit of the colubrariid vampire snails.

Results

Coluporins Represent a Multigenic Family in *Cumia reticulata*

Our approach permitted to recover the full-length sequence of 30 nonredundant mRNA sequences encoding proteins characterized by an actinoporin-like domain in the transcriptome of *C. reticulata*, which will be hereafter named coluporins (GenBank accession numbers MH194204-194233). This result should be considered as a conservative estimate of the whole complement of coluporin sequences encoded by

the genome this species, as it does not take into account genes expressed at negligible levels in physiological conditions in the tissues available (i.e. salivary and midesophageal glands and whole body [WB]).

Overall, virtually translated coluporin sequences ranged from 192 (coluporin-11) to 265 (coluporin-29) amino acids in length and invariably displayed an N-terminal signal peptide for secretion. This region, comprising 20–28 residues, was highly conserved across all sequences, mirroring the situation often reported for other invertebrate toxins (Kaas et al. 2012; Pineda et al. 2014), and it can be summarized with the consensus sequence MXLQFPXLKTXLXIFLFVIGHXPXXVXX, based on an 80% threshold of residue conservation (fig. 3D). However, in spite of the remarkable conservation of the signal peptide, coluporins only showed a limited pairwise sequence similarity, ranging from a maximum of 85.26% identity in the highest scoring pair (coluporin-16 and -26) to values lower than 25%. Although genomic data concerning expected heterozygosity levels in *C. reticulata* are lacking, the low similarity levels between coluporins seem to exclude the possibility that the 30 sequences represent allelic variants of the same genomic locus, suggesting they might rather be the product of a relatively large multigenic family.

Only 23 out of 30 coluporins displayed a recognizable canonical "sea anemone cytotoxic protein domain" (PF06369), with coluporin-18 sequence showing the most significant match (e-value = $6.4e-^{12}$). Indeed, coluporins only displayed poor to nondetectable primary sequence homology with anthozoan actinoporins, in most cases showing the highest levels of similarity with the cytolysin-like protein form the sea anemone *Stomphia coccinea* (<25% sequence identity) (Macrander and Daly 2016). On the other hand, for all coluporins we found much more significant matches (30–40% sequence identity) with porin-like protein sequences previously described from other gastropods, i.e. cone snail coluporins (Violette et al. 2012; Li et al. 2017), *M. echo* echotoxins (Kawashima et al. 2003) and auger snail tereporins (unpublished), all produced by predatory snails. Although many coluporins, like actinoporins, were cationic (isoelectrical point > 8, in 15 out of 30 sequences), others displayed a negative net charge. Overall, the isoelectrical point of coluporins averaged 7.92, ranging from 4.58 (coluporin-14) to 9.89 (coluporin-23).

Coluporins are Highly Expressed and Tissue Specific

The analysis of gene expression levels revealed that two coluporin genes were transcribed at exceptionally high levels in the midesophageal gland (MO) (fig. 1). Namely, coluporin-16 and -26, reached TPM values of ~3,675 and ~370, respectively, ranking well within the top 50 and top 200 most expressed genes in MO. Both genes displayed marked tissue specificity, as their expression levels were ~100-fold lower in the SG and even lower in WB.

In stark contrast, all the other coluporins showed strong SG specificity, with a ratio between SG and MO expression exceeding, in most cases, 100-folds (fig. 1). Coluporin-1, the transcript produced at the highest levels, reached an average expression level of ~1,565 TPM in the three SG replicates, ranking within the 100 most expressed genes in this tissue. Two additional sequences, coluporin-1 and -3 exceeded 100 TPM, and can be therefore considered as transcripts expressed at high levels in SG. Other 19 coluporin genes were expressed at moderate levels (between 10 and 100 TPM), whereas the remaining six were expressed at poor, but not negligible levels (between 1 and 10 TPM).

Overall, the transcript levels of coluporins cumulatively accounted for 0.25–0.34% and 0.26–0.53% of the global transcriptional effort of SG and MO, respectively, although in the latter case this was extremely skewed towards the synthesis of messengers encoding coluporin -16 and -26. Comparatively, the overall contribution of coluporins to the global transcriptional activity of WB (with SG and MO removed) was lower than 0.003%.

Porin Genes have an Uneven Distribution in Molluscs

Actinoporin-like genes are seemingly absent in most molluscan classes (i.e. Polyplacophora, Caudofoveata, Solenogastres, Monoplacophora, Cephalopoda, and Scaphopoda). Curiously, a remarkable number of actinoporin-like genes was detected in some bivalve genomes, whereas they were completely absent in others. Most notably, no porin gene is encoded by the genomes of oysters (Ostreidae). On the other hand, 3 to 5 genes are found in clams (Veneridae), and scallops (Pectinidae) and up to 15 in different mussel species (Mytilidae) (fig. 2).

In gastropods, comparative genomics analyses permitted to exclude the presence of actinoporin-like sequences in the genomes and transcriptomes of some major lineages, including the primitive Patellogastropoda and Vetigastropoda (the latter with the single exception of the marine top-shell *Tegula atra*, where a single expressed sequence was detected) and the most advanced Heterobranchia (table 1).

In the most diverse gastropod lineage Caenogastropoda, porin genes could be detected in representative species from the orders Neogastropoda and Littorinimorpha, albeit with different numbers and with an apparently peppered taxonomical distribution. Among littorinimorphs, Ranellidae, Tateidae, and Crepidulidae possess porin genes, whereas it remains to be established whether their lack in Naticidae marks a gene loss event or lack of expression.

Within the large neogastropod superfamily Conoidea, including thousands of venomous predatory snails classified in three dozen families, only cone and auger snails are endowed with porin-like genes. Among the rachiglossate neogastropods, some species of Muricidae, Melongenidae, and Colubrariidae have actinoporin-like genes. Interestingly, while the finding in Colubrariidae of 30 coluporins most certainly underpins a gene family expansion, the detection of a single transcript in the somehow related family Melongenidae (*Volegalea cochlidium*) seems to restrict this expansion event to the family Colubrariidae within the superfamily Buccinoidea. However, only the inclusion of additional colubrariid genera (e.g. *Colubraria*, *Bartschia*, or *Metula*) in future

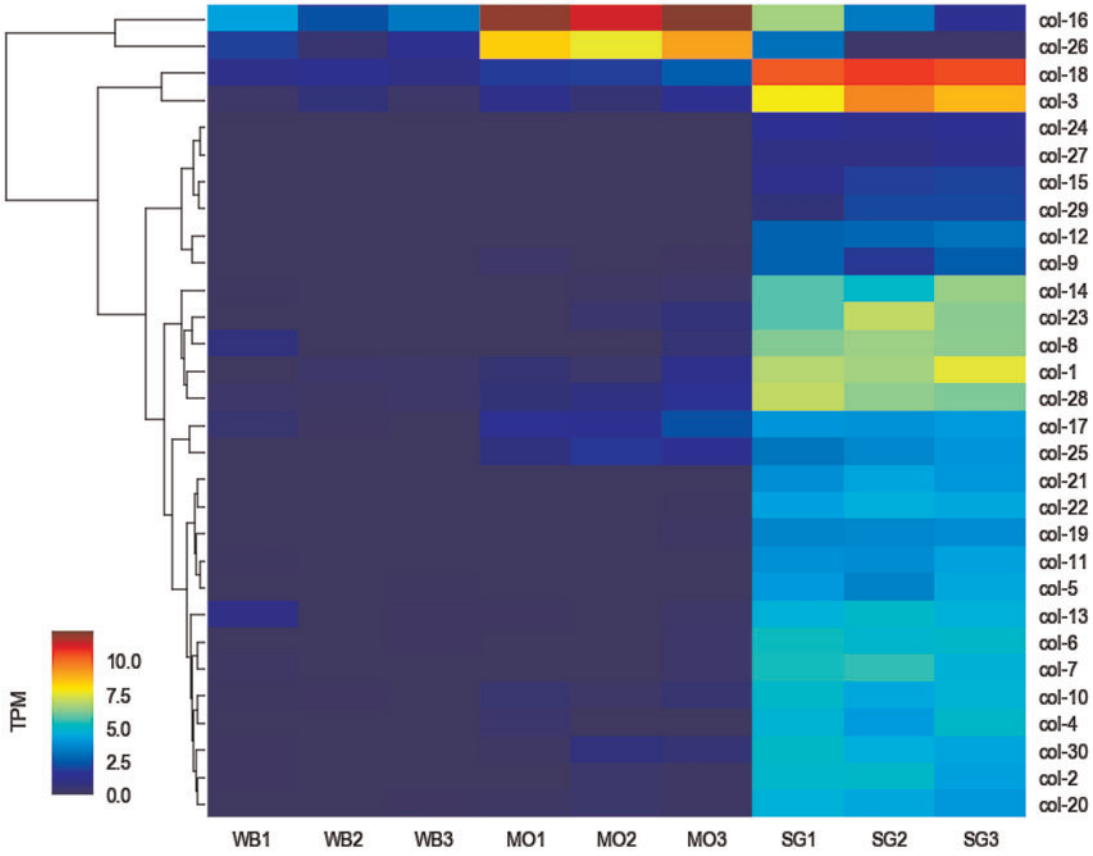

Fig. 1. Heat map summarizing expression levels of coluporins, shown as $\log_{10}$-transformed (TPM +1) values. Hierarchical clustering of coluporins was based on the calculation of Euclidean distance with average linkage.

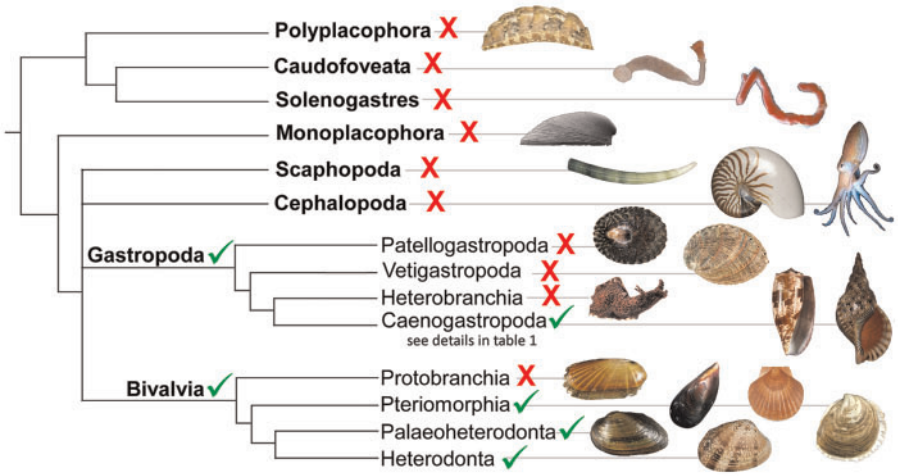

Fig. 2. Taxonomical distribution of actinoporin-like genes in Mollusca, inferred from the analysis of available genome and transcriptome resources (see Materials and Methods section for details). Red crosses identify lineages devoid of actinoporins-like genes, while green check marks indicate lineages possessing at least one actinoporins-like gene. Shell pictures courtesy of Guido and Philippe Poppe (www.conchology.be).

investigations will permit to test whether it is a general feature of this family.

Overall, the number of actinoporin sequences detected was always very low, never exceeding 3 (in *Ch. tritonis* and *Cinguloterebra anilis*), far smaller than the number detected in *C. reticulata* using the same approach. The only exception was represented by *Potamopyrgus antipodarum*, a grazer–scraper freshwater invasive snail pertaining to the family Tateidae.

Coluporins Display a Conserved Actinoporin-like Structure with Signatures of Diversifying Selection

In spite of a limited primary sequence homology and the presence of a poorly recognizable actinoporin-like domain,

Table 1. Number of nonredundant actinoporin-like sequences detected in gastropod genomes and transcriptomes.

Species	Subclass	Order	Superfamily	Family	Porins
Lottia gigantea	Patellogastropoda	—	Lottioideae	Lottiidae	0
Haliotis discus	Vetigastropoda	Lepetellida	Haliotoidea	Haliotidae	0
Aplysia californica	Heterobranchia	Aplysiida	Aplysioidea	Aplysiidae	0
Biomphalaria glabrata	Heterobranchia	Hygrophila	Planorboidea	Planorbidae	0
Crepidula atrasolea	Caenogastropoda	Littorinimorpha	Calyptraeoidea	Calyptraeidae	1
Crepidula navicella	Caenogastropoda	Littorinimorpha	Calyptraeoidea	Calyptraeidae	2
Euspira heros	Caenogastropoda	Littorinimorpha	Naticoidea	Naticidae	0
Potamopyrgus antipodarum	Caenogastropoda	Littorinimorpha	Truncatelloidea	Tateidae	13
Cipangopaludina cathayensis	Caenogastropoda	Architaenioglossa	Ampullarioidea	Viviparidae	0
Pomacea canaliculata	Caenogastropoda	Architaenioglossa	Ampullarioidea	Ampullariidae	0
Charonia tritonis	Caenogastropoda	Littorinimorpha	Tonnoidea	Ranellidae	3
Monoplex echo	Caenogastropoda	Littorinimorpha	Tonnoidea	Ranellidae	3
Volegalea cochlidium	Caenogastropoda	Neogastropoda	Buccinoidea	Melongidae	1
Cumia reticulata	Caenogastropoda	Neogastropoda	Buccinoidea	Colubrariidae	30
Dicathais orbita	Caenogastropoda	Neogastropoda	Muricoidea	Muricidae	0
Rapana venosa	Caenogastropoda	Neogastropoda	Muricoidea	Muricidae	1
Urosalpinx cinerea	Caenogastropoda	Neogastropoda	Muricoidea	Muricidae	0
Californiconus californicus	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1
Conus andremenezi	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1[a]
Conus consors	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1[a]
Conus ebraeus	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1[a]
Conus geographus	Caenogastropoda	Neogastropoda	Conoidea	Conidae	10[a]
Conus gloriamaris	Caenogastropoda	Neogastropoda	Conoidea	Conidae	0
Conus lividus	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1[a]
Conus miliaris	Caenogastropoda	Neogastropoda	Conoidea	Conidae	2
Conus monile	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1[a]
Conus virgo	Caenogastropoda	Neogastropoda	Conoidea	Conidae	1
Cinguloterebra anilis	Caenogastropoda	Neogastropoda	Conoidea	Terebridae	3
Terebra subulata	Caenogastropoda	Neogastropoda	Conoidea	Terebridae	2
Iotyrris cingulifera	Caenogastropoda	Neogastropoda	Conoidea	Turridae	0
Unedogemmula bisaya	Caenogastropoda	Neogastropoda	Conoidea	Turridae	0
Gemmula speciosa	Caenogastropoda	Neogastropoda	Conoidea	Turridae	0

NOTE—To avoid overestimates of porin numbers due to transcript fragmentation in the assembly process, only sequences displaying >50% of the expected full-length sequence (calculated as the mean length of coluporins) are reported in this table.
[a]The numbers reported are not representative of the full complement of actinoporins from the species as they represent single entries deposited in GenBank.

all coluporins displayed a remarkable predicted structural similarity with porins produced by anthozoans, as revealed by homology modelling and threading approaches by Phyre 2, HHpred, and I-TASSER approaches (fig. 3). The crystal structures of equinatoxin II (Athanasiadis et al. 2001) and sticholysin 2 (Álvarez et al. 2009) were in all cases identified as the best-fitting structural templates with 100% confidence by Phyre2. Potential templates with lower, but still significant homology (85–98% confidence) were identified in bacterial insecticidal toxins from *Bacillus thuringiensis* and *Alcaligensis faecalis* (Kelker et al. 2014; Yalpani et al. 2017) and in the fungal sphingomyelin-binding/toxic proteins nakanori (Makino et al. 2017) and pleurotolysin (Lukoyanova et al. 2015). HHpred confirmed these indications, revealing fragaceatoxin C (Morante et al. 2015) as the best fitting structure for all coluporins (with e-values ranging from $1E^{-40}$ to $1E^{-50}$). According to I-TASSER, the best fitting structures for coluporins (in most cases with TM-score > 0.7) were fragaceatoxin C, sticholysin II, and equinatoxin II, which resulted in a functional prediction (GOscore = 0.8) associated with "pore complex assembly" and "hemolysis," The predicted structure of most coluporins, well supported by several structural model evaluation methods and exemplified by coluporin-8 (fig. 3), comprises 10 to 12-sheets (9–12) arranged in a -sandwich structural scaffold, flanked by two alpha helical regions. Specifically for the coluporin-8 model obtained with PhyRe 2 we retrieved a very good LG-score value (3.881) and a fairly good MaxSub value (0.226) in ProQ, while the Ramachandran plot obtained with Procheck identified the majority of residues (85%) in the "most favored region". Additionally, the ProSA-web server calculated a z-score = −4.48 for the model (which fell well within the range of scores typically found for native proteins of similar size: Wiederstein and Sippl 2007) and the probability that the model was incorrect calculated with ModFOLD6 was = 5.349E−7, due to a global model score = 0.6176 (a score greater than 0.4 indicate models very similar to the native structure: Maghrabi and McGuffin 2017). The 3D model obtained with I-TASSER, highly similar to that obtained with Phyre2, obtained a TM-score = 0.70 ± 0.12, with Estimated RMSD = 5.7 ± 3.6 Å, further supporting the obtained structural model (Xu and Zhang 2010).

However, the presence of this conserved structural fold does not guarantee, by itself, that any pore-forming activity is maintained by coluporins (Gutiérrez-Aguirre et al. 2006). Functional studies have enabled the identification of a number of key residues that play a fundamental role in the pore-forming mechanism in anthozoan actinoporins, either in membrane binding or in pore assembly (García-Ortega

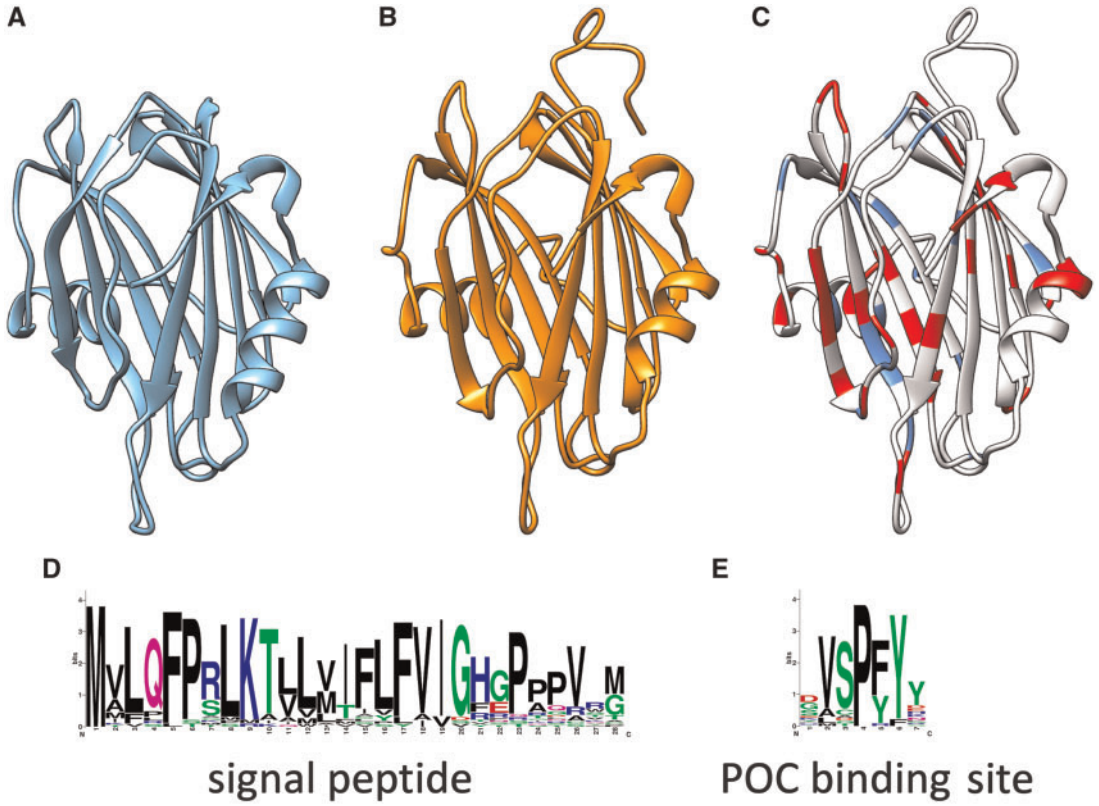

FIG. 3. (*A*) Ribbon diagram displaying the crystal structure of equinatoxin II (PDB ID: 1IAZ). (*B*) Predicted 3D structure of coluporin-8, produced by homology modeling with Phyre 2. Equinatoxin II was selected as the best structural template for coluporin-8, with confidence = 100% and model coverage = 93% (only 12 C-terminal residues could not be aligned with the template). (*C*) Coluporin-8 with sites predicted to be under positive/diversifying (red) or negative/purifying (blue) selection by at least one of the four methods used (MEME, FUBAR, SLAC, and FEL). (*D*) Consensus sequence of the signal peptide region of coluporins (residues 1–28), displayed as a conservation logo. (*E*) consensus of the residues putatively involved in the POC binding site in coluporins, homologous to positions 54, 89, 105, 107, 133, 137, and 138 of equinatoxin II, displayed as a conservation logo.

et al. 2011; Rojko et al. 2016), and their conservation in coluporins might provide useful insights for functional inference. For example, the highly conserved amphipathic alpha-helical N-terminal region of actinoporins, as demonstrated by site-directed mutagenesis, plays a pivotal role in the pore-forming mechanism, but not in membrane binding (Anderluh et al. 1997; Morante et al. 2015). This region shows high sequence diversity and extreme length polymorphisms in coluporins, as it is entirely missing in coluporin-19 and -23, but exceeds 40 amino acids in coluporin-22. Nevertheless, whenever present, the N-terminal region of coluporins shows a high prevalence of hydrophobic amino acids (36% on average), which might enable the adoption of an alpha helical secondary structure in hydrophobic environments, as suggested by the coluporin-8 model (fig. 3*B*).

The ability to recognize and bind target membranes is however provided by the POC binding site and by a partially overlapping exposed cluster of aromatic residues (García-Ortega et al. 2011). At the same time, a conserved Arg-Gly-Asp motif has been shown to play a key function in the maintenance of the oligomeric structure of the pore (García-Linares et al. 2014). However, these sites are not universally conserved, as they present a certain degree of variation in different families of sea anemones, in particular for what concerns the putative oligomerization site (Macrander and Daly 2016).

The structural comparison between coluporins and actinoporins permitted to identify a relatively good conservation of the exposed core of aromatic residues at the interface of interaction with the target membrane. As shown in figure 4, in coluporin-8 this cluster of aromatic residues consists of nine amino acids potentially involved in membrane interaction (F72, F74, Y129, F131, W137, W156, Y157, F161, and Y162).

The amino acid residues homologous to those involved in the POC binding site showed a very high level of conservation, although in several cases they diverged from those previously described in sea anemones (Macrander and Daly 2016). In detail, the first residue of the POC site (corresponding to S54 in equinatoxin II) was the one showing the highest variability and no clear pattern of conservation (fig. 3*E*), whereas the others were in large part coherent with the observations collected in anthozoan porins. Residue V89 was found to be either conserved or, in a few cases, replaced with other residues with hydrophobic side chains. S105 was conserved in the vast majority of sequences and P107 was always present, with the exception of coluporin-14 and -29, where a small deletion was detected in this position. Two out of the three conserved

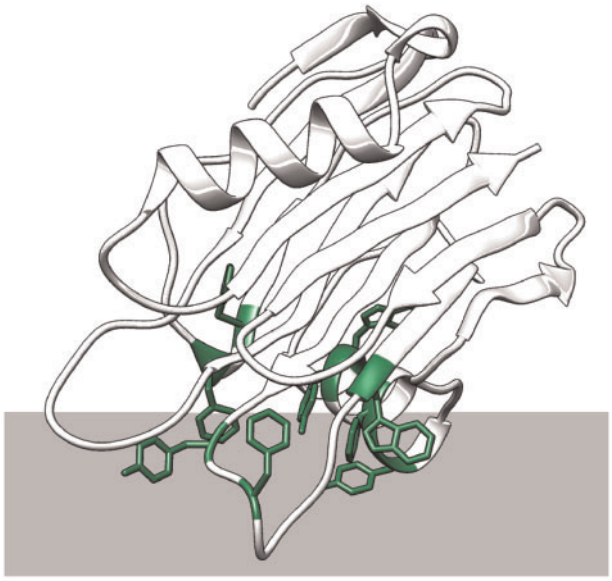

FIG. 4. Structure of coluporin-8 with the residues of the exposed aromatic cluster (F72, F74, Y129, F131 W156, Y157, F161, and Y162) highlighted in green. The region putatively involved in the interaction with the target membrane is indicated with a grey background (from Rojko et al. 2016).

tyrosine residues of the POC binding site of actinoporins (Y133 and Y 137) were found in most coluporins, or replaced with a phenylalanine residue. Finally, Y138 was conserved in about half of the coluporins identified.

In stark contrast, the residues present in a position homologous to those involved into actinoporin oligomerization showed high variability, suggesting that the aggregation of coluporins in molecular complexes might rely on other functionally important conserved residues yet to be identified.

We investigated whether the diversity of coluporins might be to some extent explained by evolutionary forces leading to diversifying selection. Our results pointed out that, while a non-negligible number of sites subject pervasive positive selection could be detected by multiple methods (7 with SLAC, 12 with FUBAR, and 17 with FEL), coluporins were massively subjected to episodic diversifying selection (i.e. only acting on a branch of the evolutionary tree). Indeed, a total of 40 sites under significant diversifying selection were detected by MEME (fig. 3C). The distribution of hypervariable sites along the sequence and even on the predicted molecular surface is rather uniform (fig. 3C).

In agreement with the structural constraints of the actinoporin fold, a number of residues were also found to be subject to strong purifying selection (25 by FEL, 18 by FUBAR, 17 by SLAC), most notably located within the highly conserved signal peptide region or within the central β-sandwich structural scaffold (fig. 3C). It is worth of a note that two tyrosine residues part of the putative POC binding site (homologous to equinatoxin II Y133 and Y137) were detected as sites under strong purifying selection by all the statistical methods used.

Phylogenetic Relationships between Coluporins and Other Molluscan Porins

From an evolutionary perspective, coluporins clearly represent a monophyletic, phylogenetically independent cluster compared to all the actinoporin-like proteins found in other molluscs (gastropods and bivalves, see above) (fig. 5). Coluporins were clustered by Bayesian inference in a highly supported clade (posterior probability = 0.99), showing a basal branch grouping the two MO-specific sequences coluporin-16 and -26, whose unique expression pattern is also matched by an ancestral origin deep within the *Cumia* lineage. Additionally, these two MO-specific coluporins share a level of pairwise similarity that is significantly higher than the level observed among their salivary counterparts.

All the SG-specific coluporin sequences are engaged in complex phylogenetic relationships with each other, often leading to poorly resolved nodes, reflecting the high average sequence divergence observed. However, the significant clustering of certain sequence pairs (e.g. coluporin-3 and -25, or coluporin-2 and -5) suggests that these sequences might be the product of relatively recent gene duplication events.

Overall, the molluscan sequences showing the closest relationship with coluporins were tereporins and conoporins. The latter, regardless of the tissue of origin (SG or venom duct) clustered with high confidence (posterior probability = 1) in a monophyletic clade, congruent with the current taxonomical classification of Colubrariidae and Conidae within Neogastropoda. However, while a marked expansion of porin genes took place in the former, porins appear to be only present with a very low number of genes per species in the latter. A clade comprising the echotoxins from Tonnoidea was placed as outgroup to the highly supported (posterior probability = 1) clade including coluporins, conoporins, and tereporins.

The porin-like sequences from bivalve molluscs and from other gastropods (*Crepidula* spp. and *P. antipodarum*, the latter being the only other gastropod where massive expansion of the porin repertoire seems to have taken place) were all grouped in a large heterogeneous clade with high statistical support (0.99).

Discussion

Colubrariidae are obligatory blood-feeding ectoparasites that rest buried in the sand or under stones in fish nests or shelters during the daytime when the host is active, taking periodic contacts with the host that are usually rather short (they can be assumed to be at maximum overnight). This behavior is remarkably similar to what has been reported in several hematophagous arthropods, in particular fleas, bedbugs, and soft ticks. Generally speaking, hematophagous parasites evolved multiple adaptations that involve morphology, behavior, and physiology for effectively feeding. The ability to impair hemostatic, immune and inflammatory responses of the host, while keeping it alive, is crucial in hematophagous feeding and is generally achieved through the evolution of a plethora of often redundant bioactive compounds that are injected into the host. Blood-feeding arthropods generally

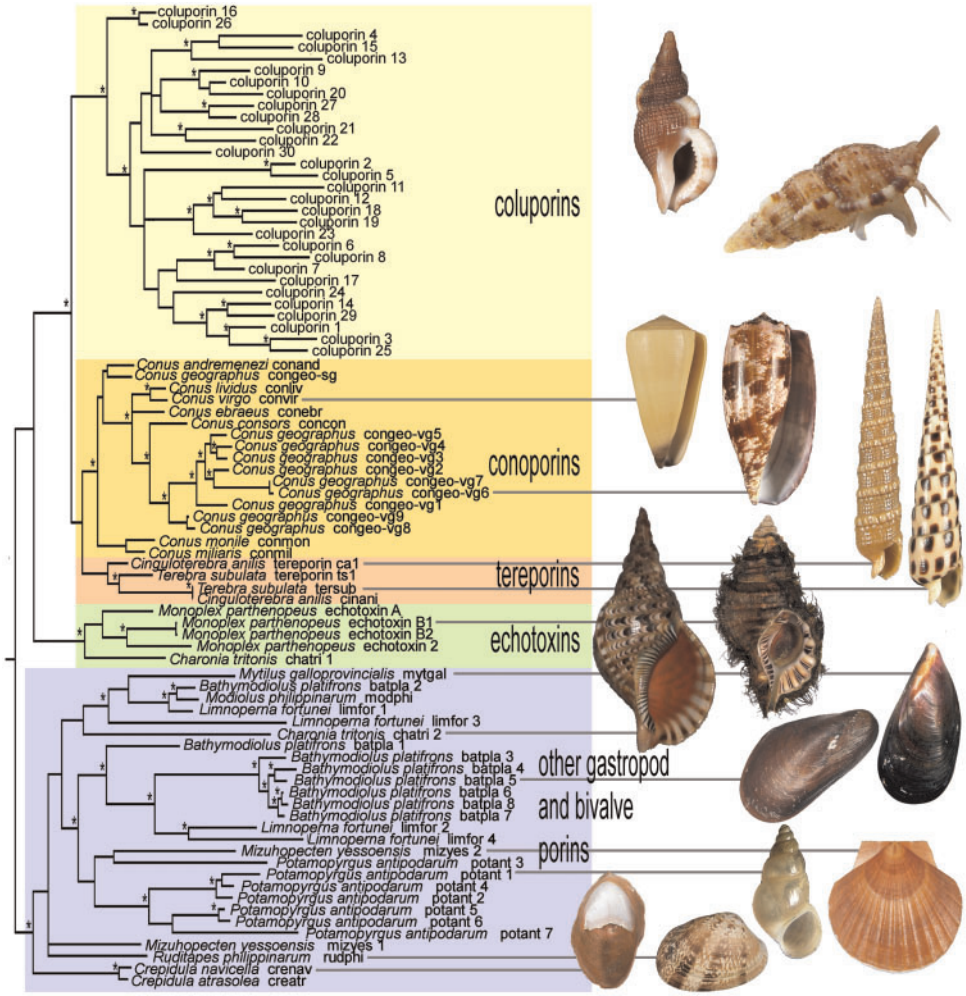

Fig. 5. Bayesian phylogenetic tree of molluscan porins. Posterior probability support values are shown for each node. Nodes supported by low posterior probability values (<0.5) were collapsed; nodes with a posterior probability value ≥0.98 are marked with an asterisk. The root of the tree was arbitrarily set at the node dividing the clade coluporins + conoporins + tereporins + echotoxins from other molluscan porins. Shell pictures courtesy of Guido and Philippe Poppe (www.conchology.be).

gain access to the blood vessels of the host mechanically, using piercing–sucking mouthparts (Krenn and Aspöck 2012) that penetrate the host's skin. The colubrariid morphological equivalents include an extensible proboscis, which is certainly able to reach the host even from a remarkable distance and probe the skin in search of a penetration point, but the hard parts are reduced to a vestigial radula that can barely protrude from the proboscis tip (Oliverio and Modica 2010). Such a delicate structure is unlikely to be responsible of the broad circular scars observed on the fish skin after detachment, invoking the contribution of secreted cytolytic and tissue degrading molecules to achieve penetration, especially given that colubrariids are able to target any part of the fish body surface, even if they preferentially seek access to the softer tegument of gills, periorbital, and perianal areas (Oliverio and Modica 2010, with references therein). Coluporins, given their putative function and their high expression level in the SGs that discharge their secretions at the very tip of the proboscis and thus at close contact with the host, are perfect candidates for such a key role in host penetration. The small vestigial radula is nevertheless likely to play a role in maintaining the proboscis in contact with the fish tegument, acting as an anchor, together with a weak suction from the proboscis, eliminating dilution issues that would reduce the activity of coluporins and other venom components.

Despite their limited sequence homology with actinoporins, coluporins share crucial structural characteristics with the coral toxins. These include the presence of a hydrophobic N-terminal region that plays a pivotal role in pore formation (Anderluh et al. 1997; Morante et al. 2015) and the conservation of the exposed core of aromatic residues at the interface of interaction with the target membrane (García-Ortega et al. 2011) (fig. 4). These findings suggest that a pore-forming activity might be maintained in coluporins. Such activity would be functional to the blood-feeding habit of Colubrariidae, allowing tissue degradation at the site of contact between the snail proboscis and the fish skin and thus facilitating the access to blood circulation.

The finding that coluporins are encoded by a relatively large multigenic family is supported by the high molecular diversity of the 30 sequences we identified in *C. reticulata*.

Moreover, the high level of conservation of the signal peptide is congruent with observations previously collected in many toxin families, including those produced by some marine snail species (Kaas et al. 2012; Pineda et al. 2014). The level of gene expansion observed in coluporins, together with variable and sequence-specific expression levels, is comparable to what has been reported in *Stichodactyla heliantus*, where 19 highly divergent actinoporin sequences have been detected, only two of which are usually produced in large amounts (Rivera-de-Torre et al. 2016). On the other hand, the 52 actinoporin isoforms detected in *Heteractis magnifica*, none of which is predominant in abundance, display a high level of sequence identity suggesting that they might, at least in part, constitute allelic variant of a same genetic locus (Wang et al. 2008). This observation implies that the levels of expansion and molecular diversification of the coluporin gene family are the highest so far observed in predation-related cytolytic lectins produced by Metazoa.

The evolutionary analysis of molluscan porins, and especially the detection of a single porin transcript in *V. cochlidium*, narrows down this expansion event to the family Colubrariidae after its split from the other members of the superfamily Buccinoidea. Such lineage specific duplications typically underlie the onset of key phenotypic traits that characterize lineages and provide adaptations to specific evolutionary niches (Fortna et al. 2004; Meyer and Van De Peer 2005; Hanada et al. 2008). In parasites, which are highly specialized groups having undergone long-term adaptation to their hosts, gene duplication has been identified as a major process driving evolution (Jackson 2015; Zarowiecki and Berriman 2015). In particular small-scale gene duplications, which are restricted to specific families, have been related to parasitism in trematodes (e.g. Wang et al. 2017) and tapeworms (e.g. Tsai et al. 2013), and have been also ubiquitously reported in the sialomes of hematophagous arthropods including ticks (Mans et al. 2002, 2017; Mans and Neitz 2004; Mans 2011; Mesquita et al. 2015), mosquitos (e.g. Calvo et al. 2006), and bedbugs (Henriques et al. 2017). These findings highlight the potentially central role of coluporins in the acquisition of hematophagous feeding. On the other hand, the remarkable number of porins reported in some bivalves (e.g. in Mytilidae) do not display any tissue specificity and may be regarded as an adaptation to defense from pathogens, a well-documented role for pore-forming molecules (Andreeva-Kovalevskaya et al. 2008; Iacovache et al. 2008). This function has been indeed demonstrated in the immune response of *Mytilus galloprovincialis* (Estévez-Calvar et al. 2011; Moreira et al. 2018). Similarly, in *P. antipodarum*, a well-studied host in trematode infections, the presence of multiple porin genes may be interpreted as an immune adaptation. It should also be noted that *P. antipodarum* recently underwent a whole genome duplication event (Logsdon et al. 2017) and that, in this species, triploidy is commonly used as a defense mechanism in populations exposed to parasites. Altogether, these factors might have led to an increased repertoire of porin genes (Osnas and Lively 2006).

In colubrariids, as in other hematophagous parasites, the expansion of protein families with a key role in feeding might have the function to produce antigenic variants that retain the same overall function. In ticks, such variants are secreted simultaneously at low levels during feeding to evade the immune system yet collectively achieving a concentration necessary for action, in the so-called multi-allelic diversifying selection model (Innan and Kondrashov 2010). This mechanism has been also invoked for the actinoporins, suggesting that it might be functional to broaden the prey spectrum (De Los Ríos et al. 2000; Wang et al. 2008; Monastyrnaya et al. 2010; Uechi et al. 2010; Valle et al. 2015). A positive relationship between dietary niche breadth and venom diversification is well documented in snakes (Li et al. 2005a, 2005b; Pahari et al. 2007) and has been recently described for *Conus* snails conotoxins, the functionally most important venom components for this snails group (Phuong et al. 2016). A broad dietary niche has indeed been anecdotally reported for several colubrariid species that are able to parasitize benthic fishes ranging from bony fishes to sharks (Oliverio and Modica 2010). Personal observations indicate that the main factor affecting host choice in colubrariids is their ability to reach it: in their natural habitat the snails tend to be associated to benthic, relatively sedentary fishes that spend the night in small cavities of the coral reef, mostly Scaridae (Bouchet 1989; Johnson et al. 1995); in aquarium, they would parasitize any benthic fish, and even attempt to reach highly motile pelagic fishes, including a juvenile *Mugil cephalus* (Oliverio and Modica, unpublished data). Indeed, the adaptation to a broad dietary spectrum would be fundamental for the survival of a slow-moving obligate parasite having no means to control the availability of potential hosts.

This scenario would also be supported by the massive occurrence of episodic diversifying selection detected in coluporins, which indicate that newly duplicated coluporin paralogous genes might have been targeted by positive selection. The functional significance of the hypervariable sites whose evolution has been shaped by episodic diversifying selection in coluporins is, however, unknown in absence of an experimentally validated structural model of the pore, and difficult to be interpreted due to their uniform distribution on the molecular surface.

Episodic selection has been frequently linked with coevolutionary dynamics, and is generally consistent with the evolutionary forces that may be imposed by shifts between different selective environments, including host shifts in immunogenic contexts (Rech et al. 2012; Cadavid et al. 2013; Murray et al. 2013; Grueber et al. 2014).

A Red Queen hypothesis, involving the evolution of new functions on both the host and the parasite side, making arsenals on both sides redundant, could hardly be invoked for the interaction between hematophagous snails and their hosts, given that the basic hemostatic functions of vertebrate hosts have been evolving since 400 mya, far before the appearance of the first known Colubrariidae (Upper Eocene, 36–37 mya: Maxwell 1966). A similar situation has been reported for ticks (Mans and Neitz 2004), and rather suggests a "one-sided arms race" (Murray et al. 2013) in which the selective pressure acts towards furnishing the predator with a number of different weapons allowing a broader host range.

This hypothesis is also complemented by recent findings on actinoporins, suggesting that different isoforms could take part in the formation of heteromeric functional pores (Rivera-de-Torre et al. 2016), a mechanism that would allow a finer modulation of the toxicity and/or a synergistic reinforcement of the activity in actinoporins.

The basal position of the two midesophageal coluporins in the phylogenetic tree compared to their salivary counterparts (fig. 5), coupled with the remarkable degree of sequence homology with conoporins (~50% identical amino acids) and short branch length, suggests that they maintained an higher similarity to the ancestral forms of colubrariid porins, likely due to a limited selective pressure. On the other hand, a duplication event deeply rooted into the colubrariid lineage might have given origin to a coluporin gene copy with specificity of expression to the SGs, which underwent multiple additional rounds of duplication coupled with diversifying selection, leading to the generation of a large and highly diversified multigenic family. This interpretation is also strongly supported by the results of the Tajima's relative rate test, highlighting a slower molecular rate of evolution for midesophageal coluporins when compared with salivary forms. These observation suggest a similar, secondary role of porins in colubrariid midesophagous and conoidean venom gland, which—not incidentally—are homologous (Ponder 1973; Kantor 2002), while the different evolutionary trajectories detectable in the salivary forms in the two snail groups highlight the functional importance of salivary coluporins.

Although data on porin numbers in other gastropod species reported in table 1 cannot be regarded as conclusive, it is interesting to note that the predatory snail species with the highest diversity of venom gland porins is *Co. geographus*. This piscivorous species has evolved a fishing-net feeding strategy, using its rostrum to engulf multiple fishes at one time, before envenomation, upon the release into the water of a venom containing insulin and sensory inhibitors to sedate and daze the preys (the so called "nirvana cabal") (Olivera et al. 2015). Despite their low level of molecular diversity, we might hypothesize that the multiple porin isoforms observed in *Co. geographus* might take part in the net feeding strategy, making fish tegument more permeable to the bioactive venom compounds released into the water: a putative functional convergence that is worth of further investigation.

While the ancestral porin gene, putatively present in the latest common ancestor of all Caenogastropoda, has been retained and maybe even neofunctionalized in some major gastropod families (e.g. Conidae and Terebridae), comparative genomics data suggest that it was lost in others (i.e. Turridae). Unfortunately, no NGS data are currently available for other species of Colubrariidae to assess whether the expansion event took specifically place in *C. reticulata* or before the split with other lineages. However, our results highlight a key role for coluporins in the adaptation to blood feeding on multiple hosts, a key characteristic for a successful feeding in such a peculiar hematophagous organism.

Material and Methods

Identification of Coluporin Transcripts in *Cumia reticulata*

Original sequencing data on *C. reticulata* (Colubrariidae; previously classified in the genus *Colubraria* but currently convincingly placed in the genus *Cumia*: MolluscaBase 2018) were derived after Modica et al (2015) (Bioproject: PRJEB9058). Paired-end raw trimmed sequencing data was trimmed *de novo* assembled using two different algorithms in parallel, i.e. Trinity v.2.5.1 (Grabherr et al. 2013) and the "*de novo* assembly" tool included in the CLC Genomics v.11 (Qiagen, Hilden, Germany). In the first case, default settings were used, allowing a minimum contig length of 200 nucleotides, whereas in the second case *word size* and *bubble size* parameters were automatically estimated. The transcriptomes assembled with the two methods (henceforth called CrTrinity and CrCLC, respectively) were virtually translated using TransDecoder v.5.01 (https://github.com/TransDecoder/), setting the minimum allowed ORF length to 100 codons. Proteomes were screened with Hmmer v.3.1b2 (Eddy 2011) to detect significant matches against the HMM profile denoting the Pfam "Sea anemone cytotoxic protein" domain (PF06369), based on an e-value threshold of 0.05.

Nucleotide sequence encoding positively matching contigs were recovered from the two assemblies and subjected to subsequent analyses. First, assembly consistency was evaluated by the pairwise comparison of contigs generated by the two algorithms and partial sequences were completed, whenever possible, by the use of Overlap-Layout-Consensus methods. Second, ambiguities were resolved by the backmapping of trimmed reads using highly stringent thresholds (length fraction = 0.75, similarity fraction = 0.98). Only transcripts displaying a uniform coverage of paired-end reads along the entire length of the CDS were considered as reliable, whereas the others were split and recursively reassembled, whenever needed. The reliability of the differentially spliced isoforms in CrTrinity was assessed using the same method.

The set of reliable full-length sequences obtained as described above were used to recursively recover additional sequences not meeting the initial search criteria, short contigs originated by fragmentation or misassembly using a tBLASTn approach. In detail, virtually translated full-length coluporin sequences were BLASTed against CrTrinity and CrCLC assemblies, and positive hits not matching with the contigs previously identified by Hmmer were selected and analyzed as described above.

In Silico Evaluation of Gene Expression Profiles

Expression levels of coluporin transcripts were calculated as Transcripts Per Million (TPM), as this normalized unit allows an efficient and reliable comparison of gene expression levels both within and between samples (Wagner et al. 2012). Briefly, trimmed reads from nine biological samples from Modica et al. (2015) (including three biological samples for WB, SG, and MO each) were separately mapped on CrTrinity, previously filtered to remove sequence redundancy by

considering only the longest transcript for each gene model. Coluporin-encoding contigs generated by Trinity were replaced by highly reliable and *in silico* validated sequences. Unique gene reads were used to compute TPM values as described by Wagner et al. (2012).

TPM expression levels +1 were transformed by $\log_{10}$ and used to generate a heat map graph. The hierarchical clustering of genes was based on the calculation of Euclidean distance with average linkage, using built-in Phyton modules.

Comparative Genomic Analyses

The NCBI nr protein sequence database was screened for the presence of sequences containing a porin-like domain as described above, limiting the search to the taxonomic entry "TaxID= 6447," i.e. the phylum Mollusca. Redundant sequences were removed with CD-hit v.4.6.8 (Li and Godzik 2006), based on a 99% similarity threshold.

In addition, RNA-sequencing data from selected species were downloaded from the SRA database, trimmed to remove adapter sequences, low quality nucleotides and short reads on a case-by-case basis, according to the specifications of each sequencing dataset. Namely, the following species and Bioprojects were considered: *Californiconus californicus*—PRJNA298293, *Conus virgo*—PRJNA298293, *Conus miliaris*—PRJNA257931, *Conus gloriamaris*—PRJNA385205 (Conoidea, Conidae), *Iotyrris cingulifera*, kindly provided by Dr Nicolas Puillandre, *Unedogemmula bisaya*—PRJNA260871 (Conoidea, Turridae), *Terebra subulata*—PRJNA286257, *Ci. anilis*—PRJNA286801 (Conoidea, Terebridae), *Rapana venosa*—PRJNA288836, *Dicathais orbita*—PRJEB12262, *Urosalpinx cinerea*—PRJNA253054 (Muricoidea, Muricidae), *Cipangopaludina cathayensis*—PRJNA264140 (Ampullarioidea, Viviparidae), *Ch. tritonis*—PRJNA383875 (Tonnoidea, Ranellidae), *Euspira heros*—PRJNA253054 (Naticoidea, Naticidae), *V. cochlidium*—PRJNA181275 (Buccinoidea, Melongidae). Details about the tissue of origin, sequencing depth, taxonomy and reference studied are reported in supplementary table S1, Supplementary Material online.

Trimmed reads for each species were used to *de novo* assembly reference transcriptomes as described above. In addition, the assembled transcriptomes of *P. antipodarum*—TSA: GFLZ00000000.1 (Truncatelloidea, Tateidae), *Crepidula atrasolea*—TSA: GFWJ00000000.1, *Crepidula navicella*—TSA: GELE00000000.1 (Calyptraeoidea, Calyptraeidae), and *Pomacea canaliculata*—TSA: GBZZ00000000.1 were downloaded from the NCBI TSA database. Assembled contigs, virtually translated with TransDecoder, were screened for positive actinoporin-like hits using the same methods applied to *C. reticulata*. To avoid overestimation in the number of unique porin-like sequences identified in each species, those were defined as the contigs (including full-length and fragments) that encoded protein fragments which length was >50% of the expected full-length sequence (calculated as the mean length of *C. reticulata* coluporins). Only full-length sequences were taken into account for phylogenetic analyses.

The presence of actinoporin-like sequences was further similarly evaluated in other molluscan classes with no reference genome available, exploiting publicly available sequence datasets. Namely, Polyplacophora (*Chiton olivaceus*, *Leptochiton rugatus*, and *Acanthochitona crinita*), Caudofoveata (*Scutopus ventrolineatus* and *Chaetoderma nitidulum*), Solenogastres (*Proneomenia sluiteri* and *Simrothiella margaritacea*) Monoplacophora (*Laevipilina hyalina*), Cephalopoda (*Octopus bimaculoides*), Bivalvia (*My. galloprovincialis*, *Limnoperna fortunei*, *Bathymodiolus platifrons*, *Modiolus philippinarum*, *Mizuhopecten yessoensis*, *Ruditapes philippinarum*, *Crassostrea gigas*, *Ennucula tenuis*, *Solemya velum*, *Uniomerus tetralasmus*, *Yoldia limatula*, and *Pyganodon grandis*), and Scaphopoda (*Antalis entalis*, *Entalina tetragona*, *Gadila tolmiei*, and *Graptacme eborea*) were considered.

Sequence Analysis and Structural Modelling

The obtained protein sequences for coluporins and the other actinoporin-like sequences for other molluscs were screened with Phobius (Käll et al. 2007) to detect the presence of N-terminal signal peptide targeting the protein precursors for secretion and discriminate these from N-terminal transmembrane regions. The presence of additional transmembrane alpha helical regions was assessed with TMHMM v.2 (Krogh et al. 2001). Secondary structures of coluporin sequences were predicted with the "predict secondary structure" tool included in the CLC Genomics Workbench, based on machine-learning methods from protein structures deposited in the PDB database (Rost 2001). The presence of accessory conserved protein domains was evaluated with InterProScan 5.27-66.0 (Finn et al. 2017).

The 3D structure of selected coluporin sequences, with particular regard with the porin domain, was modelled by homology using Phyre2 (Kelley et al. 2015) selecting the normal modelling mode. The N-terminal (after removing the signal peptide) and C-terminal extension regions were modelled *ab initio* selecting the intensive mode. Structural similarities with other proteins deposited in the PDB database were further inspected with HHpred (Söding et al. 2005) and with the threading approach implemented in I-TASSER (Zhang 2008). 3D structures were graphically represented with Chimaera1.11.2 (Pettersen et al. 2004). The quality of structural prediction was assessed using ProQ (Wallner and Elofsson 2003) and Procheck (Laskowski et al. 1993) methods implemented in PhyRe2, and the ProSA-web (Wiederstein and Sippl 2007) and ModFOLD6 (Maghrabi and McGuffin 2017) web servers.

Phylogenetic Analysis

Coluporins and the other molluscan actinoporin-like full-length sequences retrieved were trimmed to only include the region corresponding to the porin domain, based on the coordinates identified by the HMM profile. Sequences were aligned using MUSCLE (Edgar 2004) and the resulting multiple sequence analysis was subjected to a ModelTest-NG analysis to identify the best-fitting model of molecular evolution (Posada and Crandall 1998), which was determined as

the LG+G+F model (with Gamma distributed rates across sites and fixed empirical priors on state frequencies) based on the corrected Akaike Information Criteria (Hurvich and Tsai 1989).

This was used to carry out a phylogenetic Bayesian analysis with MrBayes v.3.2 (Huelsenbeck and Ronquist 2001), which included two independent runs. The analysis was stopped when the standard deviation of observed split frequencies dropped below 0.05 and the Effective Sample Size (ESS) for the estimated parameters reached a satisfying value (>200). Convergence was evaluated with Tracer v.1.6 (http://beast.bio.ed.ac.uk/Tracer).

Overall, the MCMC analysis was run for 300,000 generations. The 50% majority rule consensus tree was generated by discarding the trees sampled during the first part of the analysis (25% of the initial trees) with the burnin process. Posterior probabilities were computed for each node and those supported by low values (<50%) were collapsed.

Sequence Evolution Analysis

Codon-aligned coding sequences of *C. reticulata* coluporins were subjected to in depth analyses of molecular evolution with the Datamonkey Adaptive Evolution server (Pond and Frost 2005). Briefly, the multiple sequence alignment was tested for pervasive selection with a Fixed Effects Likelihood (FEL) method, inferring the nonsynonymous (dN) to synonymous (dS) rates of substitution on a per-site basis. Sites under positive selection were highlighted on the reference 3D structure of coluporin-8 to show their localization on the molecular surface.

We further inspected the conservation of key residues known to be involved in the function of anthozoan actinoporins in the interaction with the membrane and in the formation of the pore, obtaining a conservation sequence logo for *C. reticulata* putative POC binding site with WebLogo (Crooks et al. 2004).

The quality of evolutionary rates between MO-specific (ingroup 1) and SG-specific (ingroup 2) coluporins was tested with a Tajima's relative rate test (Tajima 1993), using the *Conus lividus* conoporin and tereporin-Ca1 as outgroups for the analysis. The null hypothesis was rejected for $P < 0.05$.

Supplementary Material

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

Acknowledgements

We thank Dr Lorenzo Graziani and Dr Marco Sollitto (University of Trieste, Italy) for technical assistance. We are grateful to Prof. Alberto Pallavicini (University of Trieste, Italy) for his helpful advice. Dr Nicolas Puillandre, Dr Paul Zaharias (Muséum National d'Histoire Naturelle, Paris, France), Dr Sébastien Dutertre (University of Montpellier, France), and Prof. Richard Lewis (University of Queensland, Australia) are gratefully acknowledged for sharing unpublished transcriptome data. Two anonymous reviewers provided useful suggestions and critical comments that helped improving the manuscript. Guido and Philippe Poppe (www.conchology.be) are acknowledged for kind permission to use their photographic database.

References

Álvarez C, Mancheño JM, Martínez D, Tejuca M, Pazos F, Lanio ME. 2009. Sticholysins, two pore-forming toxins produced by the Caribbean Sea anemone *Stichodactyla helianthus*: their interaction with membranes. *Toxicon* 54(8):1135–1147.

Anderluh G, Macek P. 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: actiniaria). *Toxicon* 40(2):111–124.

Anderluh G, Pungercar J, Krizaj I, Strukelj B, Gubensek F, Macek P. 1997. N-terminal truncation mutagenesis of equinatoxin II, a pore-forming protein from the sea anemone *Actinia equina*. *Protein Eng.* 10(7):751–755.

Andreeva-Kovalevskaya ZI, Solonin a. S, Sineva EV, Ternovsky VI. 2008. Pore-forming proteins and adaptation of living organisms to environmental conditions. *Biochemistry (Moscow)* 73(13):1473–1492.

Athanasiadis A, Anderluh G, Maček P, Turk D. 2001. Crystal structure of the soluble form of equinatoxin II, a pore-forming toxin from the sea anemone *Actinia equina*. *Structure* 9(4):341–346.

Bischofberger M, Gonzalez MR, van der Goot FG. 2009. Membrane injury by pore-forming proteins. *Curr Opin Cell Biol.* 21(4):589–595.

Bose U, Wang T, Zhao M, Motti CA, Hall MR, Cummins SF. 2017. Multiomics analysis of the giant triton snail salivary gland, a crown-of-thorns starfish predator. *Sci Rep.* 7(1):1–14.

Bouchet P. 1989. A marginellid gastropod parasitizes sleeping fishes. *Bull Mar Sci.* 45:76–84.

Bouchet P, Perrine D. 1996. More gastropods feeding at night on parrotfishes. *Bull Mar Sci.* 59:224–228.

Cadavid LF, Palacios C, Lugo JS. 2013. Bimodal evolution of the killer cell Ig-like receptor (KIR) family in New World primates. *Immunogenetics* 65(10):725–736.

Calvo E, Mans BJ, Andersen JF, Ribeiro JMC. 2006. Function and evolution of a mosquito salivary protein family. *J Biol Chem.* 281(4):1935–1942.

Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res.* 14(6):1188–1190.

De Los Ríos V, Oñaderra M, Martínez-Ruiz A, Lacadena J, Mancheño JM, Martínez Del Pozo Á, Gavilanes JG. 2000. Overproduction in *Escherichia coli* and purification of the hemolytic protein sticholysin II from the sea anemone *Stichodactyla helianthus*. *Protein Expr Purif.* 18(1):71–76.

Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol.* 7(10):e1002195.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.

Estévez-Calvar N, Romero A, Figueras A, Novoa B. 2011. Involvement of pore-forming molecules in immune defense and development of the Mediterranean mussel (*Mytilus galloprovincialis*). *Dev Comp Immunol.* 35(10):1017–1031.

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M, et al. 2017. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* 45(D1):D190–D199.

Fortna A, Kim Y, MacLaren E, Marshall K, Hahn G, Meltesen L, Brenton M, Hink R, Burgers S, Hernandez-Boussard T, et al. 2004. Lineage-specific gene duplication and loss in human and great ape evolution. *PLoS Biol.* 2(7):e207–e954.

Frazão B, Vasconcelos V, Antunes A. 2012. Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: an overview. *Mar Drugs* 10(8):1812–1851.

García-Linares S, Richmond R, García-Mayoral MF, Bustamante N, Bruix M, Gavilanes JG, Martínez-Del-Pozo Á. 2014. The sea anemone actinoporin (Arg-Gly-Asp) conserved motif is involved in maintaining the competent oligomerization state of these pore-forming toxins. *FEBS J.* 281(5):1465–1478.

García-Ortega L, Alegre-Cebollada J, García-Linares S, Bruix M, Martínez-Del-Pozo Á, Gavilanes JG. 2011. The behavior of sea anemone actinoporins at the water-membrane interface. *Biochim Biophys Acta Biomembr*. 1808(9):2275–2288.

Gerdol M, Luo YJ, Satoh N, Pallavicini A. 2018. Genetic and molecular basis of the immune system in the brachiopod *Lingula anatina*. *Dev Comp Immunol*. 82:7–30.

Grabherr MG, Haas BG, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol*. 29(7):644–652.

Grueber CE, Wallis GP, Jamieson IG. 2014. Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PLoS One* 9(3):e89632.

Gutiérrez-Aguirre I, Trontelj P, Maček P, Lakey JH, Anderluh G. 2006. Membrane binding of zebrafish actinoporin-like protein: AF domains, a novel superfamily of cell membrane binding domains. *Biochem J*. 398(3):381–392.

Hadders MA, Beringer DX, Gros P. 2007. Structure of C8-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* 317(5844):1552–1554.

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu S-H. 2008. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol* 148(2):993–1003.

Henriques BS, Gomes B, da Costa SG, Moraes C, da S, Mesquita RD, Dillon VM, Garcia E, de S, Azambuja P, Dillon RJ, Genta FA. 2017. Genome wide mapping of peptidases in *Rhodnius prolixus*: identification of protease gene duplications, horizontally transferred proteases and analysis of peptidase a1 structures, with considerations on their role in the evolution of hematophagy in triatomi. *Front Physiol*. 8:1–22.

Hong Y, Ohishi K, Inoue N, Kang JY, Shime H, Horiguchi Y, Van der Goot FG, Sugimoto N, Kinoshita T. 2002. Requirement of N-glycan on GPI-anchored proteins for efficient binding of aerolysin but not *Clostridium septicum* α-toxin. *EMBO J*. 21(19):5047–5056.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755.

Hurvich CM, Tsai C-L. 1989. Regression and time series model selection in small samples. *Biometrika* 76(2):297.

Iacovache I, Bischofberger M, van der Goot FG. 2010. Structure and assembly of pore-forming proteins. *Curr Opin Struct Biol*. 20(2):241–246.

Iacovache I, van der Goot FG, Pernot L. 2008. Pore formation: an ancient yet complex form of attack. *Biochim Biophys Acta – Biomembr*. 1778(7–8):1611–1623.

Innan H, Kondrashov F. 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet*. 11(2):97.

Jackson AP. 2015. Preface: the evolution of parasite genomes and the origins of parasitism. *Parasitology* 142(S1):S1–S5.

Johnson S, Johnson J, Jazwinski S. 1995. Parasitism of sleeping fish by gastropod mollusks in the Colubrariidae and Marginellidae at Kwajalein, Marshall Islands. *Festivus* 27:121–126.

Kaas Q, Yu R, Jin AH, Dutertre S, Craik DJ. 2012. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. *Nucleic Acids Res*. 40(D1):D325–D330.

Käll L, Krogh A, Sonnhammer ELL. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res*. 35(Web Server):W429–W432.

Kantor YI. 2002. Morphological prerequisites for understanding neogastropod phylogeny. *Boll Malacol*. 38:161–174.

Kawashima Y, Nagai H, Ishida M, Nagashima Y, Shiomi K. 2003. Primary structure of echotoxin 2, an actinoporin-like hemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*. *Toxicon* 42(5):491–497.

Kelker MS, Berry C, Evans SL, Pai R, McCaskill DG, Wang NX, Russell JC, Baker MD, Yang C, Pflugrath JW, et al. 2014. Structural and biophysical characterization of *Bacillus thuringiensis* insecticidal proteins Cry34Ab1 and Cry35Ab1. *PLoS One* 9(11):e112555.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 10(6):845.

Krenn HW, Aspöck H. 2012. Form, function and evolution of the mouthparts of blood-feeding Arthropoda. *Arthropod Struct Dev*. 41(2):101–118.

Kristan K, Podlesek Z, Hojnik V, Gutiérrez-Aguirre I, Gunčar G, Turk D, González-Mañas JM, Lakey JH, Maček P, Anderluh G. 2004. Pore formation by equinatoxin, a eukaryotic pore-forming toxin, requires a flexible N-terminal region and a stable β-sandwich. *J Biol Chem*. 279(45):46509–46517.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK—a program to check the stereochemical quality of protein structures. *J App Cryst*. 26(2):283–291.

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Mol Biol*. 305(3):567–580.

Leonardi A, Biass D, Kordiš D, Stöcklin R, Favreau P, Križaj I. 2012. *Conus consors* snail venom proteomics proposes functions, pathways, and novel families involved in its venomic system. *J Proteome Res*. 11(10):5046–5058.

Li M, Fry BG, Kini RM. 2005a. Eggs-only diet: its implications for the toxin profile changes and ecology of the marbled sea snake (*Aipysurus eydouxii*). *J Mol Evol*. 60(1):81–89.

Li M, Fry BG, Kini RM. 2005b. Putting the brakes on snake venom evolution: the unique molecular evolutionary patterns of *Aipysurus eydouxii* (marbled sea snake) phospholipase A 2 toxins. *Mol Biol Evol*. 22(4):934–941.

Li Q, Barghi N, Lu A, Fedosov AE, Bandyopadhyay PK, Lluisma AO, Concepcion GP, Yandell M, Olivera BM, Safavi-Hemami H. 2017. Divergence of the venom exogene repertoire in two sister species of *Turriconus*. *Genome Biol Evol*. 9(9):2211–2225.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659.

Logsdon J, Neiman M, Boore J, Sharbrough J, Bankers L, McElroy K, Jalinsky J, Fields P, Wilton P. 2017. A very recent whole genome duplication in *Potamopyrgus antipodarum* predates multiple origins of asexuality and associated polyploidy. *PeerJ* 5:e3046v1.

Lukoyanova N, Kondos SC, Farabella I, Law RHP, Reboul CF, Caradoc-Davies TT, Spicer BA, Kleifeld O, Traore DAK, Ekkel SM, et al. 2015. Conformational changes during pore formation by the perforin-related protein pleurotolysin. *PLoS Biol*. 13(2):e1002049–e1002015.

Macrander J, Daly M. 2016. Evolution of the cytolytic pore-forming proteins (Actinoporins) in sea anemones. *Toxins* 8(12):368–316.

Maghrabi AHA, McGuffin LJ. 2017. ModFOLD6: an accurate web server for the global and local quality estimation of 3D models of proteins. *Nucleic Acids Res*. 45(W1):W416–W421.

Makino A, Abe M, Ishitsuka E, Murate M, Kishimoto T, Sakai S, Hullin-Matsuda F, Shimada Y, Inaba A, Miyatake I, et al. 2017. A novel sphingomyelin/cholesterol domain-specific probe reveals the dynamics of the membrane domains during virus release and in Niemann-Pick type C. *FASEB J*. 31(4):1301–1322.

Malovrh P, Viero G, Dalla Serra M, Podlesek Z, Lakey JH, Maček P, Menestrina G, Anderluh G. 2003. A novel mechanism of pore formation: membrane penetration by the N-terminal amphipathic region of equinatoxin. *J Biol Chem*. 278(25):22678–22685.

Mans BJ. 2011. Evolution of vertebrate hemostatic and inflammatory control mechanisms in blood-feeding arthropods. *J Innate Immun*. 3(1):41–51.

Mans BJ, Featherston J, de Castro MH, Pienaar R. 2017. Gene duplication and protein evolution in tick-host interactions. *Front Cell Infect Microbiol*. 7:1–14.

Mans BJ, Louw AI, Neitz AWH. 2002. Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. *Mol Biol Evol*. 19(10):1695–1705.

Mans BJ, Neitz AWH. 2004. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem Mol Biol*. 34(1):1–17.
Mariottini GL, Pane L. 2013. Cytotoxic and cytolytic cnidarian venoms. A review on health implications and possible therapeutic applications. *Toxins* 6(1):108–151.
Maxwell PA. 1966. Some upper Eocene Mollusca from New Zealand. *N.Z. J Geol Geophys*. 9(4):439–457.
Mechaly AE, Bellomio A, Gil-Cartón D, Morante K, Valle M, González-Mañas JM, Guérin DMA. 2011. Structural insights into the oligomerization and architecture of eukaryotic membrane pore-forming toxins. *Structure* 19(2):181–191.
Mesquita RD, Vionette-Amaral RJ, Lowenberger C, Rivera-Pomar R, Monteiro FA, Minx P, Spieth J, Carvalho AB, Panzera F, Lawson D, et al. 2015. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc Natl Acad Sci U S A*. 112(48):14936–14941.
Meyer A, Van De Peer Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* 27(9):937–945.
Modica MV, Lombardo F, Franchini P, Oliverio M. 2015. The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). *BMC Genomics* 16:441.
Monastyrnaya M, Leychenko E, Isaeva M, Likhatskaya G, Zelepuga E, Kostina E, Trifonov E, Nurminski E, Kozlovskaya E. 2010. Actinoporins from the sea anemones, tropical *Radianthus macrodactylus* and northern *Oulactis orientalis*: comparative analysis of structure-function relationships. *Toxicon* 56(8):1299–1314.
Morante K, Caaveiro JMM, Tanaka K, González-Mañas JM, Tsumoto K. 2015. A pore-forming toxin requires a specific residue for its activity in membranes with particular physicochemical properties. *J Biol Chem*. 290(17):10850–10861.
Moreira R, Balseiro P, Forn-Cuní G, Milan M, Bargelloni L, Novoa B, Figueras A. 2018. Bivalve transcriptomics reveal pathogen sequences and a powerful immune response of the Mediterranean mussel (*Mytilus galloprovincialis*). *Mar Biol*. 165(4):61.
Murray GGR, Kosakovsky Pond SL, Obbard DJ. 2013. Suppressors of RNAi from plant viruses are subject to episodic positive selection. *Proc R Soc B Biol Sci*. 280(1765):20130965.
Olivera BM, Seger J, Horvath MP, Fedosov AE. 2015. Prey-capture strategies of fish-hunting cone snails: behavior, neurobiology and evolution. *Brain Behav Evol*. 86(1):58–74.
Oliverio M, Modica MV. 2010. Relationships of the haematophagous marine snail *Colubraria* (Rachiglossa: colubrariidae), within the neogastropod phylogenetic framework. *Zool J Linn Soc*. 158(4):779–800.
Osnas EE, Lively CM. 2006. Host ploidy, parasitism and immune defence in a coevolutionary snail-trematode system. *J Evol Biol*. 19(1):42–48.
Ozawa T, Kaneko J, Nariya H, Izaki K, Kamio Y. 1994. Inactivation of gamma-hemolysin H gamma II component by addition of monosialoganglioside GM1 to human erythrocyte. *Biosci Biotechnol Biochem*. 58(3):602–605.
Pahari S, Bickford D, Fry BG, Kini RM. 2007. Expression pattern of three-finger toxin and phospholipase A2 genes in the venom glands of two sea snakes, *Lapemis curtus* and *Acalyptophis peronii*: comparison of evolution of these toxins in land snakes, sea kraits and sea snakes. *BMC Evol Biol*. 7(1):175–179.
Parker MW, Feil SC. 2005. Pore-forming protein toxins: from structure to function. *Prog Biophys Mol Biol*. 88(1):91–142.
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 25(13):1605–1612.
Phuong MA, Mahardika GN, Alfaro ME. 2016. Dietary breadth is positively correlated with venom complexity in cone snails. *BMC Genomics* 17(1):1–15.
Pineda SS, Sollod BL, Wilson D, Darling A, Sunagar K, Undheim EAB, Kely L, Antunes A, Fry BG, King GF. 2014. Diversification of a single ancestral gene into a successful toxin superfamily in highly venomous Australian funnel-web spiders. *BMC Genomics* 15(1):177–116.
Podobnik M, Anderluh G. 2017. Pore-forming toxins in Cnidaria. *Semin Cell Dev Biol*. 72:133–141.
Pond SLK, Frost SDW. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10):2531–2533.
Ponder WF. 1973. The origin and evolution of the Neogastropoda. *Malacologia* 12(2):295–338.
Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14(9):817–818.
Rech GE, Vargas WA, Sukno SA, Thon MR. 2012. Identification of positive selection in disease response genes within members of the Poaceae. *Plant Signal Behav*. 7(12):1667.
Von Reumont BM, Campbell LI, Richter S, Hering L, Sykes D, Hetmank J, Jenner RA, Bleidorn C. 2014. A polychaete's powerful punch: venom gland transcriptomics of *Glycera* reveals a complex cocktail of toxin homologs. *Genome Biol Evol*. 6(9): 2406–2423.
Rivera-de-Torre E, García-Linares S, Alegre-Cebollada J, Lacadena J, Gavilanes JG, Martínez-del-Pozo Á. 2016. Synergistic action of actinoporin isoforms from the same sea anemone species assembled into functionally active heteropores. *J Biol Chem*. 291(27):14109–14119.
Rojko N, Dalla Serra M, MačEk P, Anderluh G. 2016. Pore formation by actinoporins, cytolysins from sea anemones. *Biochim Biophys Acta Biomembr*. 1858(3):446–456.
Rosado CJ, Buckle AM, Law RRP, Butcher RE, Kan W-T, Bird CH, Ung K, Browne KA, Baran K, Bashtannyk-Puhalovich TA, et al. 2007. A common fold mediates vertebrate defense and bacterial attack. *Science* 3:301–306.
Rost B. 2001. Protein secondary structure prediction continues to rise. *J Struct Biol*. 134(2-3):204–218.
Safavi-Hemami H, Hu H, Gorasia DG, Bandyopadhyay PK, Veith PD, Young ND, Reynolds EC, Yandell M, Olivera BM, Purcell AW. 2014. Combined proteomic and transcriptomic interrogation of the venom gland of *Conus geographus* uncovers novel components and functional compartmentalization. *Mol Cell Proteomics* 13(4):938–953.
Shiomi K, Kawashima Y, Mizukami M, Nagashima Y. 2002. Properties of proteinaceous toxins in the salivary gland of the marine gastropod (*Monoplex echo*). *Toxicon* 40(5):563–571.
Shiomi K, Mizukami M, Shimakura K, Nagashima Y. 1994. Toxins in the salivary gland of some marine carnivorous gastropods. *Comp Biochem Physiol*. 107(3):427–432.
Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*. 33(Web Server):W244–W248.
Takara T, Nakagawa T, Isobe M, Okino N, Ichinose S, Omori A, Ito M. 2011. Purification, molecular cloning, and application of a novel sphingomyelin-binding protein (clamlysin) from the brackishwater clam, *Corbicula japonica*. *Biochim Biophys Acta Mol Cell Biol Lipids* 1811(5):323–332.
Tsai IJ, Zarowiecki M, Holroyd N, Garciarrubio A, Sanchez-Flores A, Brooks KL, Tracey A, Bobes RJ, Fragoso G, Sciutto E, et al. 2013. The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* 496(7443):57–63.
Uechi GI, Toma H, Arakawa T, Sato Y. 2010. Molecular characterization on the genome structure of hemolysin toxin isoforms isolated from sea anemone *Actineria villosa* and *Phyllodiscus semoni*. *Toxicon* 56(8):1470–1476.
Valle A, Alvarado-Mesén J, Lanio ME, Álvarez C, Barbosa JARG, Pazos IF. 2015. The multigene families of actinoporins (Part I): Isoforms and genetic structure. *Toxicon* 103:176–187.
Violette A, Biass D, Dutertre S, Koua D, Piquemal D, Pierrat F, Stöcklin R, Favreau P. 2012. Large-scale discovery of conopeptides and conoproteins in the injectable venom of a fish-hunting cone snail using a combined proteomic and transcriptomic approach. *J. Proteomics* 75(17):5215–5225.
Voskoboinik I, Trapani JA. 2006. Addressing the mysteries of perforin function. *Immunol Cell Biol*. 84(1):66–71.
Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci*. 131(4):281–285.

Wallner B, Elofsson A. 2003. Can correct protein models be identified? *Protein Sci*. 12(5):1073–1086.

Wang S, Zhu X, Cai X. 2017. Gene duplication analysis reveals no ancient whole genome duplication but extensive small-scale duplications during genome evolution and adaptation of *Schistosoma mansoni*. *Front Cell Infect Microbiol*. 7:1–11.

Wang Y, Yap LL, Chua KL, Khoo HE. 2008. A multigene family of *Heteractis magnifica* lysins (HMgs). *Toxicon* 51(8):1374–1382.

Wiederstein M, Sippl MJ. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35(Web Server issue):W407–W410.

Xu J, Zhang Y. 2010. How significant is a protein structure similarity with TM.score=0.5? *Bionformatics* 26(7):889–895.

Yalpani N, Altier D, Barry J, Kassa A, Nowatzki TM, Sethi A, Zhao J-Z, Diehn S, Crane V, Sandahl G, et al. 2017. An *Alcaligenes* strain emulates *Bacillus thuringiensis* producing a binary protein that kills corn rootworm through a mechanism similar to Cry34Ab1/Cry35Ab1. *Sci Rep*. 7(1):10.

Zarowiecki M, Berriman M. 2015. What helminth genomes have taught us about parasite evolution. *Parasitology* 142(S1):S85–S97.

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40.