PHYLOGENOMICS AND COMPARATIVE TRANSCRIPTOMICS OF WEST AFRICAN CONE SNAILS

GOBIERNO DE ESPAÑA | MINISTERIO DE CIENCIA E INNOVACIÓN

CSIC Consejo Superior de Investigaciones Científicas

mncn museonacionaldecienciasnaturales

PHYLOGENOMICS AND COMPARATIVE TRANSCRIPTOMICS OF WEST AFRICAN CONE SNAILS · SAMUEL ABALDE LAGO 2019

PHYLOGENOMICS AND COMPARATIVE TRANSCRIPTOMICS OF WEST AFRICAN CONE SNAILS

SAMUEL ABALDE LAGO

TESIS DOCTORAL
Salamanca, 2019

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL
DEPARTAMENTO DE BIOLOGÍA ANIMAL, PARASITOLOGÍA, ECOLOGÍA, EDAFOLOGÍA Y QUÍMICA AGRÍCOLA

Esta tesis doctoral es un compendio de los siguientes artículos publicados

Capítulo I: Abalde S.[1], Tenorio, M.J.[2], Uribe, J.E.[1,3,4] and Zardoya, R.[1] (2019). Conidae phylogenomics and evolution. Zoologica Scripta, 48 (2): 194 - 214. https://doi.org/10.1111/zsc.12329

Capítulo II: Abalde S.[1], Tenorio, M.J.[2], Afonso, C.M.L.[5] and Zardoya R.[1] (2017). Mitogenomic phylogeny of cone snails endemic to Senegal. Molecular Phylogenetics and Evolution, 112: 79-87. https://doi.org/10.1016/j.ympev.2017.04.020

Capítulo III: Abalde, S.[1], Tenorio, M.J.[2], Afonso, C.M.L.[5] Uribe, J.E.[1], Echeverry, A.M.[1] and Zardoya, R.[1] (2017). Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitochondrial genomes. BMC Evolutionary Biology, 17: 231. https://dx.doi.org/10.1186%2Fs12862-017-1069-x

Capítulo IV: Abalde, S.[1], Tenorio, M.J.[2], Afonso, C.M.L.[5] and Zardoya, R.[1] (2018). Conotoxin diversity in *Chelyconus ermineus* (Born, 1778) and the convergent origin of piscivory in the Atlantic and Indo-Pacific cones. Genome Biology and Evolution, 10 (10): 2643-2662. https://doi.org/10.1093/gbe/evy150

[1] Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2] Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cadiz; 11510 Puerto Real, Cádiz, Spain

[3] Department of Invertebrate Zoology, Smithsonian Institution, National Museum of Natural History, Washington, DC, USA

[4] Grupo de Evolución, Sistemática y Ecología Molecular, Universidad del Magdalena, Santa Marta, Colombia

[5] Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

D. Rafael Zardoya San Sebastián, Profesor de Investigación en el Museo Nacional de Ciencias Naturales (MNCN-CSIC),

CERTIFICA:

Que la Tesis Doctoral titulada **"Phylogenomics and comparative transcriptomics of West African cone snails"** ha sido realizada bajo su dirección por **D. Samuel Abalde Lago**, con DNI 53189988-G, para optar al Título de Doctor en Biología por la Universidad de Salamanca, ha sido realizada bajo su dirección en el Departamento de Biodiversidad y Biología Evolutiva del Museo Nacional de Ciencias Naturales y reúne todos los requisitos científicos y formales necesarios para su defensa. Esta memoria está además adscrita a la Facultad de Biología de la Universidad de Salamanca, en el Departamento de Biología animal, Ecología, Parasitología, Edafología y Química Agrícola.

En Salamanca, a 2 de julio de 2019

Fdo.: Rafael Zardoya San Sebastián

Agradecimientos

Ésta es, sin duda, la sección que me parece más difícil de escribir. Después de estos cuatro años tan intensos, de momentos compartidos, estoy seguro de que me olvidaré de alguien que no se lo merece. Si eres tú, de verdad, lo siento.

Me gustaría agradecer en primer lugar a mi director, Rafa. Gracias por todo. Gracias por confiar en mí y darme esta beca a pesar de que no tenía experiencia. Gracias por las cervezas, las enseñanzas, los tirones de oreja y tu apoyo durante estos años. Y gracias por tirar de mí en esta recta final: cuando yo pensaba que no llegaba tú nunca has desfallecido.

Y gracias Valentín por tu sinceridad, tu ayuda y por estar siempre dispuesto a firmar lo que hiciese falta.

Gracias a mi familia: mis padres, mis hermanas, mis abuelos. Siempre me habéis ayudado, y siempre habéis tenido una palabra de ánimo cuando lo necesitaba. Me habéis apoyado sin condiciones, aún cuando yo no me he tomado el tiempo de ir más a veros, de explicaros qué estoy haciendo en Madrid. Si hoy estoy aquí es por vosotros.

Cuando me mudé a Madrid lo hice sólo, pero eso duró poco. En estos cuatro años he conocido más gente de la que sabría nombrar. Jorge, Miriam, Andrea, Paula, Ernesto, Guille, Violeta, Martí, Cristina, María, Mariana, Iker, Fer, Gabi, Lourdes, Joserra, Yi, FAFA y David (aunque no conviviéramos en el día a día) gracias por los cafés, los cotilleos, las cervezas. Esta tesis no habría sido lo mismo sin esos momentos. Paloma, gracias no sólo por todos los momentos, sino por tu apoyo, por acordarte de mí cuando sabías que necesitaba un empujón y, sobretodo, por una portada preciosa. Buckley, gracias por todo lo que me has enseñado. Hablar contigo es una lección constante. Iván, gracias por tu ayuda constante. Te desvives por ayudarnos a los demás aunque no tengas tiempo. Siempre se puede contar contigo. Y Karen, aunque no nos veamos muy a menudo, cuando coincidimos siempre eres capaz de hacerme sonreír. No importa que dan agobiado esté. Gracias a todos.

Entre todos, hay tres personas a las que me gustaría destacar. Michel: por enseñarme tanto, por apoyarme siempre, por ayudarme cuando lo he necesitado. Juanes: me has enseñado (casi) todo lo que sé hacer en el laboratorio. Me has instruido, me has apoyado y siempre has estado ahí. Gracias por ser un buen amigo. Y, cómo no, Silvi: la persona a la que visitar cuando me apetece dar un paseo, quien siempre tiene tiempo

para un café, para escuchar mis penas. Eres, como dirías tú, una de mis personas favoritas.

Cuando llegué al museo lo hice pensando que el de la ciencia era un mundo clasista: los estudiantes con los estudiantes. Querría agradecer a los investigadores del museo que me han demostrado que eso no es así: Marta, Íñigo, Mario, Borja, Nacho, Judith, Christoph, Pepe. Annie, gracias también por confiar en mí para participar en vuestras clases. Gracias a todos por lo mucho que me habéis enseñado.

En estos momentos me acuerdo especialmente de Piluchi e Isabel. Si en algún momento me gustó pasar por el laboratorio fue para coincidir con vosotras.

Esta tesis se habría quedado en mucho menos sin la inestimable ayuda de Manolo y Camané. Vaya capacidad para muestrear y, sobre todo, vaya conocimiento enciclopédico. Gracias por vuestra ayuda. Y gracias, Manolo, por estar siempre disponible cuando me encontraba alguna concha desconocida (casi siempre, vaya).

This small paragraph goes for the amazing people I met during my research stays. Sébastien, Vittoria, Julien, thanks for your patience and for teaching me every day. Tsuyoshi, Hiro, Takuya, Bastian, Franzis: thanks for an amazing time in the other side of the world. I miss it. Sara, gracias por enseñarme siempre algo nuevo y por tener buena palabras hacia mí. Y gracias por aceptar ser parte de mi tribunal, me hace mucha ilusión.

Gracias a todos mis compañeros del máster, siempre se puede contar con vosotros. Tete, gracias por tus incontables visitas y por las comidas de categoría. Margaux, me hace muy feliz que hayamos vuelto a vivir en la misma ciudad aunque sólo fuera un par de meses. David y Alicia (ya inseparables): puede que me haya olvidado de vosotros en el aeropuerto, pero eso no significa que no me apetezca veros en cada oportunidad que tengamos. Ojalá la próxima sea pronto (da igual cuándo leáis esto).

Para terminar en alto, Iris. Esta tesis no sólo ha sido dura para mí. Has aguantado pacientemente los largos días, los fines de semana de trabajo, los viajes. Los días en los que yo volvía a casa sólo con ganas de dormir. Has sido mi apoyo en los malos momentos, que los ha habido. Lo siento si a veces ha sido duro, pero me alegro de que hayas estado conmigo. Eres la razón por la que me apetecía volver a casa cada día.

Contents

0. - Summary

The family Conidae (Gastropoda: Caenogastropoda) is one of the most specious families of marine animals in the world. Cones are of interest for naturalists and biologists, who tried to define a systematic classification of the group. However, the astonishing diversity of species and their wide distribution range have hindered phylogenetic studies on the group, and it was not until recently (and not before the start of this project) that a robust evolutionary framework for the family was defined. Cones are active predators that use a sophisticated venom system to hunt their prey. This venom is highly complex, including hundreds of peptides named conotoxins. These secreted peptides target different ionic channels and neurotransmitter receptors in the body of the prey, with a degree of specificity that makes them a potential source of drug development. The advent of Next Generation Sequencing (NGS) technologies has brought the study of venom composition to a new level, allowing the identification of even the lowest expressed transcripts. Nonetheless, thus far, the variability of venom has been mainly studied from a proteomic perspective, and the transcriptomic approach has been limited mostly to the description of venom composition. Moreover, the study of the cones has been focused in those species inhabiting the Indo-Pacific region, whereas other regions that represent a hotspot of biodiversity for the group (like the Caribbean sea and West Africa) have been neglected.

In this regard, the present project was focused exclusively in the study of the cone snails from West Africa. We defined five main goals, which are worked in detail in the different chapters:

1) To establish a robust evolutionary framework for the family Conidae, thus studying the origin of the diet specializations
2) To study the phylogenetic relationships and evolution, and to explore the real diversity of the cones endemic to Senegal, Canary Islands and the archipelago of Cabo Verde
3) To study the evolution of the piscivorous diet in the family, by analysing through transcriptomics the venom composition in the Atlantic species *Chelyconus ermineus*

4) To characterize the venom composition of multiple species from two cone genera endemic to West Africa, and to apply a comparative study between them to study the evolution of their venoms

The mitochondrial genomes of four genera were sequenced, and the transcriptomic data of other 27 species were analysed to annotate the mitochondrial genes and 21 nuclear genes. All these markers and those already available in GenBank were used to infer the phylogenetic relationships of the family Conidae. The reconstructed phylogenetic trees recovered most of the nodes with high statistical support, although some of the deepest nodes remain unresolved. These results confirmed previous hypotheses regarding diet evolution, like the monophyly of the molluscivorous genera and of those that prey on amphinomids, but the double origin of piscivory of the genera from Indo-Pacific and Atlantic/ East Pacific (*Chelyconus*) regions.

The phylogenetic relationships among the endemic genera of cone snails (*Africonus*, *Lautoconus*, *Kalloconus*, and *Trovaoconus*) were inferred using more than 100 mitochondrial genomes, representative of most of the described species in this region. All the trees were recovered with maximum statistical support. The clade formed by the genera *Kalloconus* was the first offshoot. The sister group to *Africonus* is the species *Lautoconus ventricosus* from the Mediterranean Sea, and the clade comprising *Lautoconus* species from Senegal and the Canary Islands is the sister group to these two. Because of the paraphyly of *Lautoconus* due to *Africonus*, it was proposed the designation of a new generic name for the species of Senegal and the Canary Islands upon the corresponding taxonomic revision. Also, the genus *Trovaoconus* (endemic to Cabo Verde) should be synonymized to continental *Kalloconus*. The time of the cladogenetic events suggests that allopatry could play an important role in generating the diversity of these genera, since the eustatic sea level fluctuations and the direct development of the larvae could favour isolation. In all these genera, we found instances of species incorrectly described, whether because these species represented populational variability or because they had a convergent phenotype.

The genus *Chelyconus* represents the second adaptation to piscivory in Conidae, as mentioned above. In Cabo Verde inhabits the species *C. ermineus*, whose venom composition was studied using transcriptomic data. The comparison of the venom

composition among three individuals of this species confirmed the intraspecific variability of the venom, since only 20% of the identified conotoxins were present in the three specimens. This species presented the M, O1, O2 and T superfamilies as the most abundant, as it was also the case in previous studies of other cone species. The measure of the expression levels confirmed the importance of the A superfamily to prey on fish, as previously observed in the Indo-Pacific clade. Interestingly, this superfamily presented a cysteine pattern only described in this genus: an α4/4 spacing pattern of cysteines, in contrast with the α3/5 observed in the other clade. These results reinforce the idea that piscivory has evolved independently in both clades, and that the similarities in their hunting strategy are due to convergence.

Finally, the venom composition of 12 cone species endemic to West Africa (nine *Africonus* and four *Lautoconus*) was described. They all shared the same six expanded (with five or more conotoxin precursors) superfamilies , which could be explained by their presence in the common ancestor of these genera. The species of *Africonus* presented additional expanded superfamilies. This could be related to the higher levels of diversification of this genus and the ample opportunities of adaptation after the colonization of the archipelago of Cabo Verde. The venom composition of these genera show certain level of phylogenetic signal, so as that, in general, species more closely related have more sequences in common, as well as the same superfamilies, in the same proportion, and with similar levels of expression.

In summary, this work advances in the study of the West African cone snails, providing a robust evolutionary framework and information of the venom composition for half of the genera inhabiting there.

0. - Resumen

La familia Conidae (Gastropoda: Caenogastropoda) es una de las familias de animales marinos más diversas del mundo. Los conos son interesantes para naturalistas y biólogos, quienes han tratado de definir una clasificación del grupo. Sin embargo, la increíble diversidad de especies y su amplio rango de distribución han complicado los estudios filogenéticos, y no ha sido hasta hace poco (y no antes de comenzar este proyecto) que se definió un marco evolutivo robusto para la familia. Los conos son depredadores activos que usan un sofisticado sistema venenoso para cazar a sus presas. Este veneno es altamente complejo, incluyendo cientos de péptidos llamados conotoxinas. Estos péptidos secretados afectan a diferentes canales iónicos y receptores de neurotransmisores en el cuerpo de la presa, con un grado de especificidad que las hace una fuente potencial de nuevos medicamentos. El desarrollo de las técnicas de secuenciación masiva (NGS) ha llevado el estudio de la composición del veneno a un nuevo nivel, permitiendo la identificación de incluso los péptidos menos expresados. No obstante, la variabilidad del veneno ha sido estudiada principalmente desde una perspectiva proteómica, y la transcriptómica se ha limitado principalmente a estudios descriptivos. Además, el estudio de los conos se ha centrado en aquellas especies que habitan la región Indo-Pacífica, y otras regiones que representan puntos calientes de biodiversidad de conos (como el Caribe o África occidental) han sido ignoradas.

A este respecto, el presente proyecto está enfocado exclusivamente en el estudio de los conos de África occidental. Para ello, hemos definido cinco objetivos, trabajados en detalle en los siguientes capítulos:

1) Establecer un marco evolutivo robusto para la familia Conidae, estudiando así el origen de las dietas más especializadas.
2) Estudiar las relaciones filogenéticas y la evolución, así como explorar la diversidad real de los conos endémicos de Senegal, Islas Canarias y el archipiélago de Cabo Verde.
3) Estudiar la evolución de la dieta piscívora en la familia, analizando mediante transcriptómica la composición del veneno de la especie atlántica *Chelyconus ermineus*.
4) Caracterizar la composición del veneno de múltiples especies de conos de dos géneros endémicos de África occidental, y aplicar un estudio comparado entre ellas para estudiar la evolución de sus venenos.

Los genomas mitocondriales de cuatro géneros fueron secuenciados, y los transcriptomas de otras 27 especies fueron analizados, anotando sus genes mitocondriales y 21 genes nucleares. Todos esos marcadores y los ya publicados en GenBank fueron utilizados para inferir las relaciones filogenéticas de la familia Conidae. Los árboles filogenéticos reconstruidos recuperaron la mayoría de los nodos con alto apoyo estadístico, aunque algunos de los nodos más antiguos permanecen sin resolver. Estos resultados confirmaron hipótesis previas relacionadas con la evolución de la dieta, como la monofilia de los géneros moluscívoros y de aquellos que depredan anfinómidos, pero el doble origen de la piscivoría en los géneros de las regiones del Indo-Pacífico y del Atlántico/ Pacífico este (*Chelyconus*).

Las relaciones filogenéticas entre los géneros de conos endémicos (*Africonus*, *Lautoconus*, *Kalloconus* y *Trovaoconus*) fueron inferidas utilizando más de 100 genomas mitocondriales, representantes de la mayoría de las especies descritas en esta región. Todos los árboles se recuperaron con máximo apoyo estadístico. El clado formado por los géneros *Kalloconus* y *Trovaoconus* fue el primero en divergir. El grupo hermano de *Africonus* es *Lautoconus ventricosus* del mar Mediterráneo, y el clado que contiene las especies de *Lautoconus* de Senegal y las Islas Canarias es hermano de los dos. Debido a la parafilia de *Lautoconus* por la inclusión de *Africonus*, se ha propuesto la designación de un nuevo nombre genérico para las especies de Senegal y las Islas Canarias una vez realizada la correspondiente revisión taxonómica . Además, el género *Trovaoconus* (endémico de Cabo Verde) debería sinonimizarse con el continental *Kalloconus*. El tiempo de los eventos cladogenéticos sugiere que la alopatría podría estar jugando un importante papel en la generación de la diversidad de estos géneros, dado que las fluctuaciones eustáticas del nivel del mar y el desarrollo directo de las larvas podrían favorecer el aislamiento. En todos los géneros encontramos ejemplos de especies descritas incorrectamente, ya fuese porque esas especies representasen variabilidad poblacional o porque tenían fenotipos convergentes.

El género *Chelyconus* representa la segunda adaptación a la piscivoría, como se ha mencionado más arriba. En Cabo Verde podemos encontrar esta especie, cuyo veneno fue estudiado utilizando transcriptomas. La comparación de la composición del veneno de tres individuos ha confirmado la variabilidad intraespecífica, dado que sólo el 20% de las conotoxinas identificadas estaban presentes en los tres especímenes. Esta especie presenta las superfamilias M, O1, O2 y T como las más abundantes, tal y como

se ha reportado anteriormente para otra especies de conos. Los niveles de expresión han confirmado la importancia de la superfamilia A para depredar en peces, como se había observado en el clado Indo-Pacífico. Curiosamente, esta superfamilia presentaba un patrón de cisteinas sólo descrito en este género: un patrón de espaciado entre las cisteínas α4/4, en contraste con el α3/5 observado en el otro clado. Estos resultados refuerzan la idea de que la piscivoría ha evolucionado independientemente en ambos clados, y que las similitudes en la estrategia de caza son debidas a convergencia.

Finalmente, se describió la composición del veneno de 12 especies endémicas de África occidental (nueve *Africonus* y cuatro *Lautoconus*). Todas comparten las mismas seis superfamilias expandidas (con cinco o más precursores de conotoxinas), lo que podría explicarse por su presencia en el ancestro común de estos géneros. Las especies de *Africonus* presentaron superfamilias expandidas adicionales. Esto podría estar relacionado con los altos niveles de diversificación de este género y las amplias oportunidades de adaptación después de colonizar el archipiélago de Cabo Verde. La composición del veneno de estos géneros muestra cierto nivel de señal filogenética, de manera que, en general, las especies más cercanas evolutivamente tienen más secuencias en común, así como las mismas superfamilias en la misma proporción y con los mismo niveles de expresión.

En resumen, este trabajo avanza en el estudio de los conos de África occidental, aportando un marco evolutivo robusto e información sobre la composición del veneno para la mitad de los géneros habitando esta región.

1. - Introduction

1. - Introduction

Known since the time of Carl Linnaeus, whose *Systema Naturae* already included 30 species, cone snails have become an important group of study in numerous fields. For example, their astonishing diversity, with a whole variety of shell colours (Fig. 1), has drawn the attention of naturalists, collectors and artists (Tucker and Tenorio, 2013). Cone snails are also an important element for understanding the ecological dynamics of coral reefs and other intertidal environments, since they are key components of these marine communities (Kohn, 1959). Cone snails are marine predators that use for hunting their prey a highly sophisticated venom system, which contains hundreds of different toxins (Norton and Olivera, 2006). This venom is so powerful that even has the ability to kill human beings (Kohn, 2016). This is the reason why cone snails have also been a main subject in neuroscience and pharmacological research, since their venom can trigger very different physiological responses in the body (McIntosh and Jones, 2001; Miljanich, 2004; Garber, 2005; Anderson and Bokor, 2012; Romero et al., 2017). Moreover, cone snails have been used as model system in evolutionary biology for understanding the processes and mechanisms generating and maintaining the current diversity of species on Earth (Williams and Duda, 2008).

Fig. 1 – Small representation of the diversity of shapes and colours of cone snails/ Shells from the Malacology collection of the Museo Nacional de Ciencias Naturales (MNCN-CSIC)

1.1. – General features of cone snails

1.1.1. – Cone snail biology

Distributed in all tropical and subtropical seas around the world (see below for a more detailed description of their distribution), cone snails inhabit mainly intertidal environments associated to different substrata, such as rocky shores, sandy bottoms or coral reefs (Tucker and Tenorio, 2009). Conids present two types of larval development that will determine the range of distribution of the species. On one hand, there are species with planktotrophic development, whose larvae can move freely with the ocean currents and inhabit a wide geographical range. On the other hand, there are species that have lecitotrophic larvae and direct development, which means that the new individuals tend to move very little from where the eggs hatched. These species are confined to very narrow areas, sometimes even to particular bays (Kohn and Perron, 1994). Although the type of development has not been observed directly for many species, it can be inferred by the morphology of the protoconch, as having a paucispiral protoconch indicates that the larvae will spend little or no time feeding in the plankton (Jablonski and Lutz, 1983). Based on this, most of the species of cone snails present a planktotrophic stage, while the non-planktotrophic species are restricted to particular lineages. From an evolutionary perspective, the most likely hypothesis postulates that the ancestor of cone snails had planktotrophic larvae, and multiple adaptations to the lecitotrophic development occurred during the evolutionary history of the group (Duda and Palumbi, 1999a).

Cones are venomous voracious predators that have adapted to prey on very different groups of marine organisms. From a very wide point of view, they can be classified into vermivorous (preying on polychaete worms), molluscivorous (feeding on other gastropods) and piscivorous (preying on fish) (Kohn, 1959; Tucker and Tenorio, 2013). However, some observations might suggest that this classification may be too general, as some species could predate on more than one type of prey. For example, Aman et al. (2014) observed the vermivorous *Tesseliconus tessulatus* attacking a fish in an aquarium, and Songdahl (1973) demonstrated that the venom of the non-molluscivorous *Lindaconus spurius* was lethal for other gastropods. In any case, diet observations of cone snails are scarce, many times anecdotic, and being detailed studies restricted to very few species (Kohn, 1956, 1959; Duda et al., 2009a; Duda and Lee, 2009). These studies prove that cones tend to have highly specialized diets, so as that a

given cone species hunts only a particular family of preys. This specialization is likely a product of competition among species, as exemplified by *Miliariconus miliaris* from the Eastern Island (the most remote oceanic island in the world): it presents a wider diet than its conspecifics in other parts of the world, who need to compete for food with other cone species (Duda and Lee, 2009).

Having so different diets might suggest that cone snails evolved from a generalist diet and adapted over time to feed on each kind of prey. However, different phylogenetic analyses suggest a vermivorous origin and further diet shifts within the family. This is supported not only by the diversity of genera preying on polychaetes (the most common prey of cone snails) (Duda et al., 2001; Puillandre et al., 2014), but also by the relative position of *Californiconus californicus*, the only generalist species, placed among other vermivorous and not as the first off-shoot of the phylogenetic tree (Puillandre et al., 2014). These same studies suggest that the shift to prey on other molluscs happened only once, since these cone species from a monophyletic lineage. In contrast, the origin of piscivory could be more complex. While Duda and Palumbi (2004) suggested that there might be up to three or four independent shifts to piscivory, other studies (Duda et al., 2001; Puillandre et al., 2014) proposed that this diet had arisen only twice in the evolutionary history of the group, independently in the Indo-Pacific region versus the Atlantic and East Pacific regions. However, the relevant nodes in the trees lack the necessary statistical support, and the controversy requires further studies using molecular markers with additional phylogenetic information. Although the shift from preying on worms to fish could be hard to imagine, Aman et al. (2014) proposed that it may have been easier than expected. They studied the venom of *Tesseliconus tessulatus*, a vermivorous species, and found out that the δ-conotoxins of the O-superfamily (critical to prey on fish) of this species and of the piscivorous snails are rather similar, and they suggested that this species might have the ability to hunt fish, even if it does not normally have to.

In order to hunt their preys, cone snails present different attacking strategies, although the most studied ones are those related to fish-hunting cones. The first species reported to prey on fish were *Pionoconus striatus* and *Pionoconus catus* (Kohn, 1956). As observed in an aquarium, these species bury themselves in the sand, but keeping the proboscis out in the water. They track their prey chemically and, once the proboscis touches the fish, the radula is stung into the body of the prey, thus injecting paralytic

venom. Then, the snail comes out of the sandy bottom and engulfs the fish. This behaviour is called "taser-and-tether", since the effect on the prey resembles an electric shock (Olivera et al., 2015). A second strategy is the so-called "net-engulfment", and it is typical of the genus *Gastridium*. In this case, the prey is engulfed alive by the rostrum before the sting of the radula and the venom release (Olivera et al., 2015). The reason why the fish does not escape might be that the snail releases the venom straight to the water, thus sedating the fish (Safavi-Hemami et al., 2015). This strategy is normally used with schools of small fishes. Finally, the third strategy to hunt fish is known as "strike-and-release", which is similar to the "taser-and-tether" strategy, but in this case the fish is not immediately paralyzed, but the snail follows it until the venom makes effect and can eat it (Olivera et al., 2015). In contrast, molluscivorous species might inject the venom multiple times in the prey by stinging them repeatedly with the radula, as it was observed for *Cylinder textile* (Prator et al., 2014).

As predators, cones' fitness depends to a large extent on their ability to hunt their prey. While most gastropods use their radula to scrap on the surface of leaves and algae, in the superfamily Conoidea (including Conidae), the radular tooth has been modified, and it is specially designed to be stung into the prey. Each radular row contains independently functional radular teeth, that have the shape of a harpoon and are, in many cases, hollow to deliver the venom during the injection (Kohn et al., 1999). The radular teeth are in different stages of development in the radular sac, with those ready to use close to the pharynx and pointing to the opening of the sac (Endean and Duchemin, 1967). When hunting, the increasing hydrostatic pressure into the proboscis propels the tooth out of the snail (Kohn et al., 1999). Radular morphology is adapted to the particular prey of the snail, to the extent that it can be used to infer the diet of those species lacking any observation (Fig. 2). Actually, even if the different diet types have different morphologies, specific adaptations to particular preys can change the radular morphology (Duda et al., 2009b; Tucker and Tenorio, 2009). Radular characters have been widely used to classify conoidean snails into different families and genera (e.g. (Taylor et al., 1993; Kohn et al., 1999; Tucker and Tenorio, 2009, 2013). Although they can be used to discriminate quite well between genera and species (Duda et al., 2009b; Tucker and Tenorio, 2013), they are useless for inferring interrelationships among them (Kantor and Puillandre, 2012).

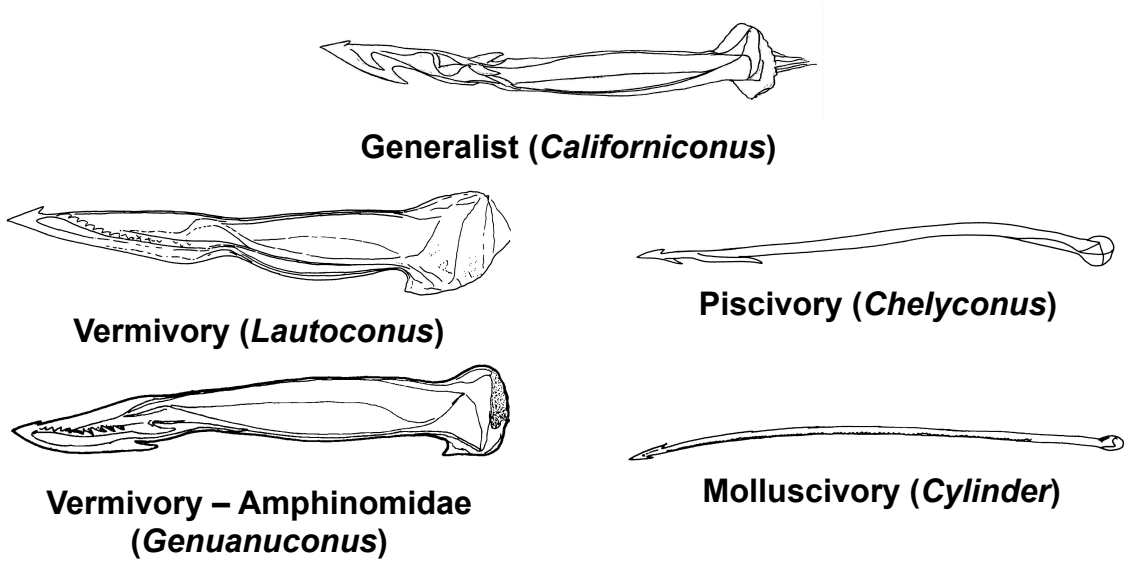


Fig. 2 - Representation of the radular tooth types for each kind of diet/ Images taken from Tucker and Tenorio (2009)

Together with radular evolution, conoidean snails have developed a sophisticated venom system to immobilize and capture preys. It is formed by: 1) the proboscis, used to locate the prey and sting the radula into it; 2) a radular sac, containing all the radular teeth; 3) the venom duct, that synthesizes the venom; and 4) the venom bulb, a muscular tissue in charge of propelling the venom out of the duct (Endean and Duchemin, 1967). This venom gland, which is present in all cone snails, has been lot several times in different families of the conoideans (Holford et al., 2009; Castelin et al., 2012).

1.1.2. - Venom

Cone snails present a venom system that could be, arguably, one of the most sophisticated in the animal kingdom. The venom of each species consists on hundreds of different peptides, named conopeptides or conotoxins, as well as some hormones and other proteins that have been proposed to enhance the effect of the venom (Terrat et al., 2012; Barghi et al., 2015; Safavi-Hemami et al., 2015). Conotoxins are rather short peptides that are synthesized as precursors, presenting a particular structure: a signal region, in the N-terminal end and highly conserved, that facilitates the pass through the endoplasmatic reticulum and the Golgi apparatus (Kaas et al., 2010); a propeptide region, more or less conserved, which participates in the folding of the final protein (Buczek et al., 2004) and/or favours posttranslational modifications (Bandyopadhyay et

al., 1997); a mature region, highly variable and often cysteine-rich, which represents the functional peptide (Kaas et al., 2010); and, occasionally, a final post-region (Fig. 3). All these regions are cleaved, and only the mature peptide is used during prey envenomation (Woodward et al., 1990).

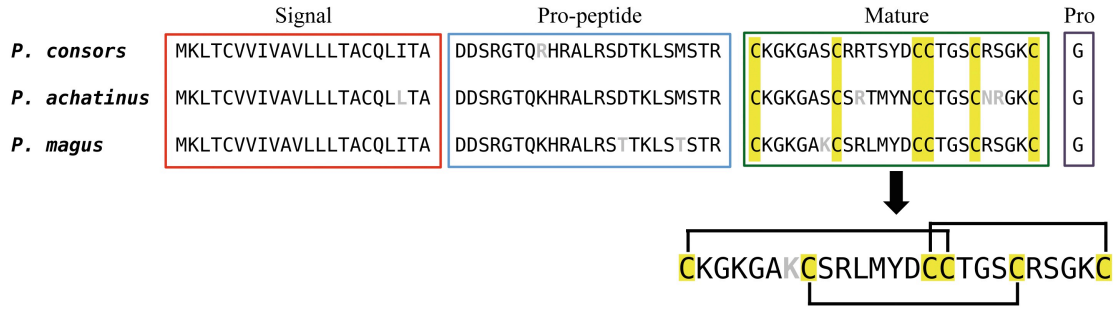


Fig. 3 - Typical conotoxin precursor structure: signal, pro-peptide, mature, and post region. The different amino acids are in grey and the cysteine patterns highlighted in yellow. The three precursors belong to the O1-superfamily, Framework VI/VII. The disulfide bonds of the precursor of *Pionoconus magus* are also depicted. The precursors come from the GenBank entries: CCI55497 (*Pionoconus consors*), P0C8V8 (*Pionoconus achatinus*) and P05484 (*Pionoconus magus*)

These regions have been used traditionally to classify conopeptides in different categories, and there are up to four different co-existing classifications. First, conotoxins have been broadly classified depending on the number of cysteines in their mature region into cysteine-rich or cysteine-poor peptides, which can or cannot form disulfide bonds, respectively (Terlau and Olivera, 2004; Olivera, 2006). Based on this classification, cysteine-poor peptides have been termed conopeptides (the same category as other proteins that might enhance venom activity), while the cysteine-rich ones were considered as "true" conotoxins (Lebbe and Tytgat, 2016). However, this classification makes no sense from an evolutionary point of view, since the cysteine-poor sequences do not form a cluster different from the so-called conotoxins (Puillandre et al., 2012). Moreover, cysteine-poor peptides also present their own biological function, as it is the case of the conantokin (B superfamily) (Robinson and Norton, 2014).

A second, widely used, classification is based on the sequence similarity of the signal regions. Conotoxins are grouped in "superfamilies", generally named with letters from the alphabet. Unlike the previous one, this classification does seem to reflect

evolutionary history (Robinson and Norton, 2014). However, every new studied species uncovers new superfamilies of conotoxins previously unknown (Terrat et al., 2012; Jin et al., 2013; Barghi et al., 2015), and thus, this classification might no be able to hold the real diversity of conotoxins. Third, the numbers of cysteine residues in the mature region together with the number of disulfide bonds formed during the peptide processing are used to classify mature peptides into "cysteine Frameworks". In some cases, these frameworks are congruent with the superfamily-based classification (e.g., the Framework I belongs to the superfamily A), but there are cases of different superfamilies sharing the shame framework of cysteines (e.g., Framework VI/VII is present in the superfamilies I3, M, O1, O2 and O3) (Kaas et al., 2010) and of superfamilies having more than one framework (e.g., the superfamily A presents the Frameworks I, II, IV, VI/VII, XIV, and XXII).

Finally, conotoxins can be classified based on their biological/pharmacologycal function and are named after Greek letters (Norton and Olivera, 2006). Conotoxins can target different ion channels, neurotransmitter transporters, and receptors in the nervous system of the prey (Terlau and Olivera, 2004; Norton and Olivera, 2006), triggering different physiological responses: from sedation to complete paralysis. Those conotoxins known as μ-conotoxins, for example, block sodium channels, whereas κ-conotoxins target potassium channels (Terlau and Olivera, 2004). More generally, it has been proposed that conotoxins might be acting together synergetically in the body of the prey, thus enhancing the venom activity, and forming the so-called "conotoxin cabals" (Norton and Olivera, 2006). There are currently described three types of cabals, each one of them characteristic for different prey-hunting strategies. First, the "nirvana cabal", proposed for those species that follow the "net engulfment" strategy, contains peptides that are released to the water and are meant for sedating the prey. They consist on hormones like insulin (Safavi-Hemami et al., 2015) and conantokins, known for producing a "sleeping" state (Olivera et al., 1985). Second, the "lightening-strike cabal", which produces an effect similar to an electric shock ("taser-and-tether" strategy), produces massive depolarization in the muscular axons, thus activating muscle activity that ends in tetanic paralysis. And third, the "motor cabal" (present in different hunting strategies) produces paralysis in the prey by inactivating the normal functioning of neurotransmitters (Norton and Olivera, 2006).

The incredible diversity of conopeptides has been long known (Olivera et al., 1990), and diet specificity is one of the explanations of this diversity. On one hand, different prey types require different conotoxins to target their receptors and neurotransmitters. On the other hand, venom composition can vary depending on diet specificity. Having a more generalist diet implies having peptides adapted to the physiological pathways of different preys, and requires producing a higher diversity of conopeptides (Remigio and Duda, 2008; Phuong et al., 2016). Chang and Duda (2016) tested this hypothesis by evaluating the change on conotoxins expression depending on the dietary breadth of *Virroconus ebraeus*, which varies depending on the age of the individual. Interestingly, they proved how the conotoxin diversity of this species could vary along the lifespan of the individuals, as their ecological necessities were also changing. From a molecular perspective, other mechanisms that can explain this diversity of conotoxins are: 1) as other secreted proteins, conopeptides evolve under strong selection pressures, mostly focussed on the mature region (Conticello et al., 2001; Espiritu et al., 2001; Duda et al., 2009a); 2) new conotoxin variants are generated via gene duplication (Duda and Palumbi, 1999b; Espiritu et al., 2001); 3) recombination events, with two gene copies generating a third variant (Espiritu et al., 2001); and 4) a reduced number of genes can produce high diversity of conotoxins by differential pre-processing of peptides, for example, via posttranslational modifications, cleavage site variability, etc. (Dutertre et al., 2013). Moreover, cone snails can adequate their venom depending on their final use, thus producing even higher conotoxin diversity. Dutertre et al. (2014) proposed that cone snails can produce a different cocktail of conotoxins for defending against predators *versus* hunting preys, and that this different production of venoms was regionalized, happening in different parts of the venom conduct. In fact, Prashanth et al. (2016) later confirmed differences in venom composition of predatory and hunting behaviours in the genus *Rhizoconus*. This genus uses almost exclusively D-superfamily conotoxins when they defend themselves against predators. Histological studies of the venom gland in the 1960s already proposed that different parts of the venom duct contained different venom bodies, that could actually have different functions (Endean and Duchemin, 1967).

However, fully understanding venom evolution and the role of particular conotoxin superfamilies in diet specificity cannot be achieved without: 1) a robust phylogeny of the group, providing the evolutionary framework for venom evolution; 2)

comparative analyses between species and genera, both with the same and different diets; 3) a good knowledge on the specific diet of the studied species. The recent advances in the study of venoms, which implement Next Generation Sequencing (NGS) techniques and high-throughput proteomic analyses, provide a better picture of the venom composition of particular specimens, thus facilitating these evolutionary analyses.

1.1.3. - Diversity and distribution

Despite being a (relatively) young clade originated in the Eocene (circa 55 mya; (Kohn, 1990), cone snails show a remarkably diversity, which includes more than 900 species described to date (WoRMS Editorial Board, 2019; the reference database for marine organisms) Therefore, this clade is one of the most specious groups of marine animals. Furthermore, there is probably much hidden diversity yet to be discovered within the group since, as shown in Fig. 4, cone snail species description is far from reaching stationarity. The advances on molecular techniques, and the effort to disentangle species complexes (Duda et al., 2008), can uncover cryptic diversity (Duda et al., 2009b), and even discover representatives of new genera (Monnier and Tenorio, 2015). Moreover, it is likely that many new species will be described in the future, as previously neglected habitats are sampled (e.g., deep ocean) and more attention is paid to new specimens (e.g. minute snails from sediments) little studied to date.

Nonetheless, this species diversity should be revisited under the light of molecular data. Although some authors did use characters from the internal anatomy to identify and describe species (Pin and Leung Tack, 1995), the general rule has been to employ shell shape and coloration patterns for these tasks. However, phenotypic characters have proven to be highly homoplasious and in many cases do not perform well in species diagnosis. For example, it is known the phenotypic plasticity of *Pionoconus magus* along its range of distribution. At the same time, there are cases of cryptic diversity within the genus (Duda et al., 2008; Duda et al., 2009b). Therefore, the total number of valid cone species needs to be taken cautiously.

This astonishing diversity, combined with the young age of the clade, has prompt some authors to propose that cone snails have a higher diversification rate than any other marine gastropod or bivalve (Stanley, 2008). There are several reasons that

could explain accelerated diversification rates. First, some authors propose that the dietary specialization of cone snails (explained above) reduces competition, thus allowing for a higher diversity in the same habitat (Taylor et al., 1980). In this regard, Stanley (2008) proposed that carnivory is a major promoter of diversity on one hand by controlling the expansion of species, and on the other hand by providing an opportunity for predators to specialize on new preys.

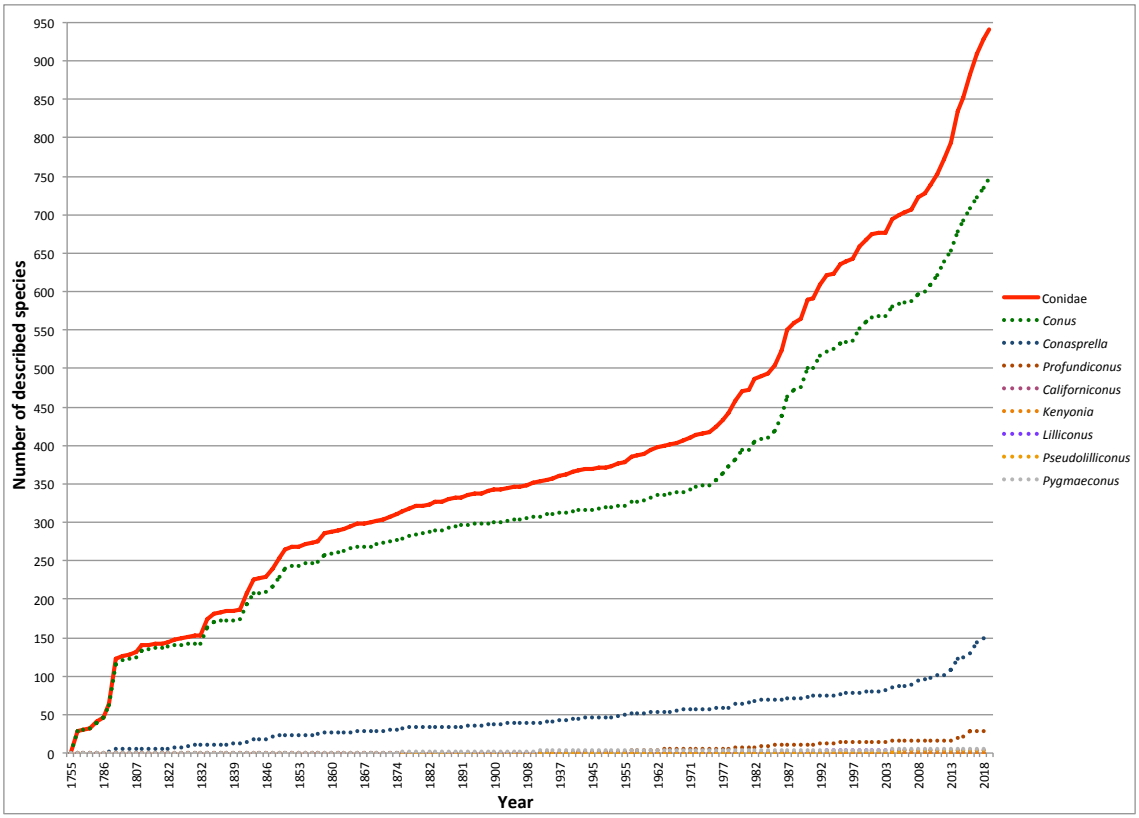


Fig. 4 - Cumulative number of described species of cone snails over years. The dotted lines represent the different genera, and the red line represents the total number of species for the family Conidae *sensu* Puillandre et al. (2014). These numbers were calculated based on the accepted species in the WoRMS database.

Second, venom production and, more specifically, venom evolution could have been a key innovation that led to a radiation of the Conoidea. Cone snails' venom has proven to be highly variable and specialized, displaying big differences even among individuals of the same species (Chang and Duda, 2016; Peng et al., 2016). This ability to produce and maintain high levels of diversity could have facilitated the adaptation to

hunt on new preys, which would enhance niche specialization and species diversification (Duda and Palumbi, 2004; Chang and Duda, 2012).

Third, larval development could play a major role on speciation. As explained by Stanley (2008), species with planktotrophic larvae have the ability to inhabit wider geographical habitats, while maintaining gene flow among populations. It is the case of *M. miliaris* and *V. ebraeus*, for example (Duda et al., 2009a; Duda and Lee, 2009). In contrast, local radiations tend to be related to genera and/or species presenting direct development, which limits the dispersal capacity of individuals. This is the pattern observed, for instance, in the genus *Africonus* from Cabo Verde (Duda and Rolán, 2005), the genus *Lautoconus* from Senegal (Pin and Leung Tack, 1995), and the genus *Jaspidiconus* from the Caribbean Sea (Berschauer, 2015).

And forth, different environmental and geological factors could affect speciation and extinction in species. Kohn (1990) documented several radiation and extinction events in the fossil record of cone snails that could be due to environmental factors, since the same trends were observed on other groups of marine gastropods. In the Anomura hermit crabs, for example, Davis et al. (2016) documented how the speciation rates increased in times of climate cooling, when the sea level drops isolating populations that diverge in allopatry. This pattern could be applied, potentially, to any other intertidal living invertebrate, since variations in sea level change the available surface, and connect and isolate cyclically populations. Tectonic activity also could affect the availability of suitable habitats for marine organisms, and it explains the high biodiversity present in the Indo-West Pacific (IWP) region, as proposed before ((Williams and Duda, 2008) and references therein). Tectonic events that occurred at the Oligocene-Miocene boundary, about 20-25mya, connected the IWP with the surrounding areas, thus promoting colonization from peripheral regions. Also, these events increased the amount of intertidal habitats and the length of the coastline, providing more suitable habitats for colonization and specialization. Williams and Duda (2008) studied the potential effect of these tectonic changes on the diversification rates of three groups of gastropods, including cone snails. They found an increased rate of diversification in the branch leading to the Major Clade (the current family Conidae), whose origin was proposed to have occurred coinciding with the increase of global tectonic activity in the Oligocene-Miocene boundary. In the three groups, most of the IWP clades have higher diversification rates, which could be due to the collision of the

Australia and New Guinea plate with South East Eurasia, increasing the heterogeneity and availability of habitats.

Molecular phylogenetics, together with new statistical models, could allow us to test among all these competing hypotheses. However, the results should be taken cautiously, since these analyses are strongly affected by the diversity included (or considered) for each clade. For instance, an underestimation of the diversity of any clade would bias the result by lowering the rate of speciation. Also, the different use of fossils might change the estimated age of the clades.

Cone diversity is currently distributed all over the world, in all the tropical and subtropical seas. The highest diversity is found in the Indo-Pacific province, with more than half of the total species live. However, local hotspots of biodiversity can be found in the Caribbean Sea and West Africa (20 and 14%, respectively) inhabiting a much narrower area (Tucker and Tenorio, 2013). One reason explaining this distribution might be the ecological suitability of the different areas. Cunha et al. (2014) found out that species richness in West Africa was correlated with sea surface mean temperature. Another factor is the availability of habitats, as assessed by Williams and Duda (2008) and explained above, with the tectonic activity of the IWP leading to a burst of speciation.

The current distribution of the species, together with the published phylogenies, might suggest that the geographical origin of this group is the Indo-Pacific. However, the fossil record contradicts this hypothesis. The oldest fossils assigned to cone snails are from the coasts of England and France (Kohn, 1990; Tracey et al., 2017), so this group could be of European origin and the current distribution is a consequence of its last radiations. This Atlantic European origin of cone snails, where they are not present today, is explained by the tropical conditions of this area in the Eocene ((Kohn, 1990) and references therein). Alternatively, the fossil record of earlier cones could be biased towards West Europe.

Cone snails occupy different habitats in terms of depth. Although most species live in the intertidal environment, there are species that live at hundreds of meters of depth, like *Dauciconus fenzani* (150m) or *Conasprelloides villepinii* (ca. 400m; (Tucker and Tenorio, 2013). In numbers, Peters et al. (2013) provide a good approximation: among the 632 evaluated species in the IUCN RedList, 53.6% lived between 1 and 5m, 27.7% above 50m, and less than 19% inhabited deeper waters.

Another striking comparison among geographical regions is related to the mode of larval development. While most of the species in the Indo-Pacific province have a planktotrophic larval development, the local hotspots of diversity (like the East and West Atlantic) are represented by species with direct development, showing these regions higher rates of endemicity (Tucker and Tenorio, 2013). This pattern correlates with the conservation status of the different species within the group. Peters et al. (2013) assessed the state of conservation of 632 species around the world, and evaluated three as Critically Endangered (CE), 11 as Endangered (EN) and 27 as vulnerable. The 14 species evaluated as CE or EN inhabit the Cabo Verde archipelago, whereas 18 of the 27 VU species occur in the East or West Atlantic. In summary, 25.5% of the eastern Atlantic species are threatened. Species with direct development are more prone to extinction, since catastrophic events can wipe out an entire species, while the species with wider distribution ranges could have populations at safe distance from these events (Stanley, 2008). In the case of Cabo Verde, CE species are restricted to specific bays, which have been recently radically transformed by human activity (Peters et al., 2016).

1.1.4. - Classification of cone snails

The incredible species richness and wide distribution of cone snails makes particularly difficult deciphering their phylogenetic relationships and establishing taxonomic levels. This section presents a brief review about the alternative historical classifications proposed for this group.

Despite general instability in the classification of the higher taxonomical levels of Gastropoda, the monophyly of the Conoidea has been widely accepted (Taylor et al., 1993; Puillandre et al., 2008) However, this superfamily includes thousands of species, which makes obtaining a robust phylogeny-based classification of the group cumbersome. Moreover, even the traditional name of the superfamily, Toxoglossa, has become inappropriate since many of the families included do not have the characteristic radular morphology that was considered as a synapomorphy of the group. In the last years, the inclusion of new species in the studies has made some families paraphyletic, setting the need for the creation of new ones. A recent reclassification of the three traditional families within the group, Conidae, Turridae and Terebridae, has led to the recognition of at least 16 families (Puillandre et al., 2008; Uribe et al., 2018; Abdelkrim et al., 2019).

This PhD thesis is focused in the family Conidae, the only one accepted from the very beginning of the taxonomic history of the group that is still considered monophyletic. The current taxonomic classification of the family is based on either the study by Tucker and Tenorio (2009) based on radular morphology or that carried out by Puillandre et al. (2015) based on a molecular phylogeny, which essentially coincide in the composition of the different groups and only differ in their taxonomic rank.

The family Conidae was for many years accepted as monogeneric, meaning that the more than 500 species described at the time were included into one single genus: *Conus* (Kohn, 1990). This classification remained unaltered, despite the numerous attempts to subdivide the genus *Conus* into different genera, until the exhaustive morphological revision of radular characters (Tucker and Tenorio, 2009). In this study, the authors divided the family into four different families and 69 genera, later raised to 90 (Tucker and Tenorio, 2013). Many of the new genera were formerly considered subgenera, so despite the drastic increase in number of genera, the change made by this proposal was not radical. However, while radular morphology proved to be a good diagnostic character of groups, it was rather useless in reconstructing their interrelationships.

The molecular study, based on the partial sequences of three mitochondrial genes, by Puillandre et al. (2014) set the basis for the currently accepted classification of cone snails (Puillandre et al., 2015). On this study, the authors recovered four main lineages later raised to the generic category: *Profundiconus*, *Conasprella*, *Californiconus* and *Conus*. Beyond the genus level, the phylogeny presented up to 71 monophyletic lineages that were considered subgenera: 11 within *Conasprella*, and 60 in *Conus*, the most specious genus of the four. Some of these clades had been also recovered in other molecular studies (e.g. (Duda and Kohn, 2005), although many less species were included. The main problem of this study was the little phylogenetic signal of the markers used. Cone snails have radiated several times during their evolutionary history (see next section) and as a consequence, most of the internal nodes in the molecular phylogeny were rather short, and their statistical support was low. A more recent molecular phylogeny based on mitochondrial genomes recovered with high statistical support the four genera and suggested the existence of additional main (cryptic) lineages within Conidae (Uribe et al., 2017).

In general, both the morphological and molecular classifications are highly congruent, as they agree in most of recognized groups, but differ in their taxonomic ranks. For instance, the family Conidae *sensu* Tucker and Tenorio (2009) corresponds to the genus *Conus sensu* Puillandre et al. (2015), and most of the genera described in the former are recovered as monophyletic and designed as subgenera in the latter (although there are some minor differences in composition eventually).

The reference classification of this work will be (Tucker and Tenorio, 2009) with the updates from (Tucker and Tenorio, 2013) as that of Puillandre et al 2015 with only four genera fails to describe fairly the high diversity of the group. The genera here included not only represent generally monophyletic assemblages in (Puillandre et al., 2014) classification, but also are defined by synapomorphies in the radular teeth, and are circumscribed to specific biogeographical regions, so we consider that the use of generic names is more informative and convenient for describing the diversity and evolutionary history of the group.

1.1.5. – The fossil record

The shell of the cone snails is made of calcium carbonate, which is preserve in good conditions when buried in sediments like limestones. Hence, cone snails present quite a rich fossil record, which merits here a brief review.

Despite the rich fossil record, many fossils are simply ascribed to the group because they present the characteristic conic shaped shell. In the worst cases, numerous fossils that had been initially identified as cone snails were later rejected as misidentifications. In other cases, identifying the genus of the fossil is challenging given the rampant convergence in the shape of the shell. Olsson (1967) observed that photographing the shells under ultraviolet light reveals the coloration patterns of the fossil shells, and Krueger (1974) proposed the utility of this technique for the study of neograstropodan shells. In fact, this technique has proven to be useful in the study of different mollusc fossil assemblages (e.g. (Kase et al., 2008; Caze et al., 2010; Caze et al., 2011). In recent years, this technique was applied to the study of Neogene cones (Hendricks, 2015, 2018) , allowing the assignment of these species to extant genera based on coloration patterns. However, this is still an ongoing line of research, and much work needs to be done to characterize the extensive cone snails' fossil record.

The study that has served as reference for understanding cone evolution and speciation was carried out by Kohn (1990), who did an exhaustive review of the 2500 cone fossil records known at that moment. He set the origin of the genus *Conus* (here considered family Conidae) in the lower Eocene, about 55 millions years ago (mya). Later, Tracey et al. (2017) proposed the origin of the family as a few million years earlier, based on new fossil evidence. The study by Kohn (1990) has another important output, which is the understanding of the radiation patterns of this group. By studying the abundance and diversity of fossils from each epoch, he could determine the pulses of speciation and extinction suffered along the evolutionary history of the group. Briefly, he could identify at least three radiation events: in the Middle Eocene (from five to 42 species), in the Miocene (from 19 to 158), and during the Pleistocene, leading to the astonishing diversity present today. These radiation events were accompanied by different extinction events: in the Oligocene (from 42 to 19) and in the Pliocene (from 158 to 53). As explained above, these speciation and extinction events are common to other groups of gastropods and marine invertebrates, so they could be explained by changes in different environmental factors or the tectonic activity.

1.2. - West African cones

This present work is focused on understanding cone snail diversity and evolution in West Africa and surrounding areas. Next, the geological history of West Africa and the main particularities of the cones endemic to this area are briefly summarized.

1.2.1. –West Africa geological history

West Africa and the surrounding areas form part of the Eastern Atlantic (EA) biogeographical region. Most of this region is located in the tropics, having waters with relatively high mean temperatures, although there is part of it in the subtropics (north from the Sahara coasts, including the Mediterranean Sea, and south from Angola coasts to Namibia), with waters having lower mean temperatures (Briggs and Bowen, 2012).

For the last tens of millions of years, the Mediterranean basin has suffered a complex tectonic and climatic activity (Carminati et al., 2012) that likely changed the living conditions for all organisms inhabiting the region. In particular, there are two main events that changed dramatically the geography of this region. First, during the

Eocene the collision of the African and Euroasian plates led to the closure of the Tethys Sea (Carminati et al., 2012). This event isolated the Mediterranean Sea (whose only remaining connection is the Atlantic Ocean) from the Indo-Pacific, preventing the dispersal and gene flow between both regions. Second, the Messinian Salinity Crisis (MSC) that separated temporarily the Atlantic Ocean and the Mediterranean Sea about 5 mya, between the Miocene and the Pliocene (Krijgsman et al., 1999). Probably due to a combination of tectonic and climatic factors, this event was catastrophic for the communities leaving in the Mediterranean Sea. Thus far, this event also affected greatly the communities outside the Mediterranean basin, by changing the conditions in the Atlantic Ocean (e.g., it affected the sea level and the ocean currents).

One of the hallmarks of the West Africa region is the presence of the Macaronesian archipelagos. These sets of oceanic islands present a great species diversity, with high rates of endemicity for multiple groups of flora and fauna, and are considered hotspots of biodiversity (Myers et al., 2000; Dimitrov et al., 2008; Carine et al., 2010; Briggs and Bowen, 2012; Tucker and Tenorio, 2013).

There are up to five Macaronesian archipelagos (from North to South): Azores, Madeira, Salvagens, Canary Islands and Cabo Verde. All these oceanic archipelagos have a volcanic origin, and their ages span from 25 mya to as little as 1 mya (Ancochea et al., 1990; Guillou et al., 1996; Holm et al., 2008; Dyhr and Holm, 2010). The volcanic activity of these regions is not over, and the islands have suffered different volcanic events along their history, thus expanding their area (Dyhr and Holm, 2010). These factors combined had a critical influence on the structure of the different communities living there, as this activity very likely had a strong influence on the speciation and extinction of different groups. For example, creating new islands provide more colonization opportunities, and the volcanic activity on existing islands might enlarge the coastline. Also, this activity can potentially wipe out the existing flora and fauna of the islands because of the extreme conditions. Moreover, volcanic islands are active habitats, in constant change because of the effect of the on-going volcanic activity, the impact and erosion of the oceanic waves, or the fluctuation of the sea level (Ramalho et al., 2013; Rijsdijk et al., 2014).

In summary, the East Atlantic has witnessed a complex geological history that has impacted drastically the biota of this region, connecting and disconnecting communities. At present, the region offers a whole variety of marine habitats (intertidal,

sandy and rocky shores, coral reefs, etc.) and differences in mean temperatures. Moreover, the proximity of the Macaronesian archipelagos to the coast (100 km between Fuerteventura and the African coast; 400 km between Cabo Verde and Senegal) facilitates colonization processes. All these factors combined offer a unique opportunity for diversification, which has led to the incredible biodiversity of flora and fauna present today.

1.2.2. – West African cone snails

West Africa (or the East Atlantic) represents a hotspot of diversity for the family Conidae, including more than 14% of the global species, distributed in a much narrower area than in other regions (as the Indo-West Pacific). In fact, among the ten most specious genera of cone snails three are endemic to these waters: *Africonus* from the Cabo Verde archipelago; *Varioconus* in Angola; and *Lautoconus* in Senegal, the Canary Islands and the Mediterranean Sea (Tucker and Tenorio, 2013). *Africonus* itself represents almost the 10% of the global diversity, with more than 90 accepted species (WoRMS Editorial Board, 2019).

In this region there are currently described eight genera: *Africonus*, *Chelyconus*, *Genuanoconus*, *Kalloconus*, *Lautoconus*, *Monteiroconus*, *Trovaoconus* and *Varioconus*. Although they represent a minority of the total genera diversity of the family, they show several adaptations that are worth mentioning (Tucker and Tenorio, 2013).

First, both modes of larval development are present in West African cones, and the different genera show very distinct ranges of distribution, accordingly. *Africonus* and *Trovaoconus* are only endemic to Cabo Verde archipelago, and *Varioconus* to Angola. *Lautoconus* is present in the Mediterranean Sea, Canary Islands, and along the West coast of Africa from Morocco to Senegal. *Genuanoconus* and *Monteiroconus* are present from Senegal to Angola, and in the archipelago of Cabo Verde and the Canary Islands. Finally, *Chelyconus* shows the largest distribution, being amphi-Atlantic. It is present in the East Atlantic coast, from Senegal to Angola, and in Cabo Verde, as well as in the West Atlantic coast, from the Caribbean Sea to the coasts of South America (Monteiro et al., 2004). In terms of species the distributions are also very different, from *Trovaoconus atlanticoselvagem*, present only in the João Valente seamont (between the islands of Maio and Boa Vista), to the amphi-Atlantic *Chelyconus ermineus*, including

Lautoconus ventricosus, present in all the Mediterranean Sea. Despite these distribution patterns, all genera are endemic to these waters except *C. ermineus*, which is present in other areas (Röckel et al., 1980; Monteiro et al., 2004; Tucker and Tenorio, 2013).

Second, several diet specializations are present as inferred from radular morphology together with field and aquarium observations. There is no West African molluscivorous cone snail. The species *C. ermineus* is the only one piscivorous. *Genuanoconus genuanus* is specialized on feeding on polychaetes of the family Amphinomidae (bristle worms), and it has been observed feeding on the species *Hermodice carunculata* (Tucker and Tenorio, 2009). Finally, the rest of the species present in West Africa are vermivorous, but a more detailed observation of their feeding patterns awaits (Tucker and Tenorio, 2009).

Because of this outstanding diversity, the cone snails from West Africa have drawn a lot of attention, although some areas more than others. The archipelago of Cabo Verde has been studied for decades, with several samplings, and morphological studies along the years (Röckel et al., 1980; Rolán, 1990; Tenorio et al., 2014). In fact, almost 100 species have been described, and more are being described every year as can be observed in WoRMS. In contrast, the coastlines of Senegal and Angola have been little studied despite presenting high levels of endemicity. In Senegal, after the extensive monograph based on different anatomical characters published by Pin and Leung Tack (1995), where they described 11 endemic and five non-endemic species, only five species more from deep waters were described (e.g., (Nolf and Verstraeten, 2008; Monnier and Limpalaër, 2010).

Although the cone snails from these waters are quite well known from a morphological point of view, these studies have faced several problems when trying to delimit species, such as the variation in colours of the species, the overlapping range of populations or the lack of differences in characters widely used to separate species (e.g., the radular teeth), among others (Röckel et al., 1980). In recent years, several molecular studies (Cunha et al., 2005; Duda and Rolán, 2005; Cunha et al., 2008) have been applied to the cones from West Africa, thus helping to understand the phylogenetic relationships and evolutionary history of these species. But, again, these studies have been focused on particular regions.

The cone snails that have received a wider attention are those endemic to the Cabo Verde archipelago. These studies, based on partial sequences of several genes,

confirmed the presence of two distinct clades within the islands (Cunha et al., 2005; Duda and Rolán, 2005; Cunha et al., 2008), later classified in two genera endemic to the archipelago: *Africonus* and *Trovaoconus* (Tucker and Tenorio, 2009). Both genera are very different morphologically, as the latter are significantly bigger than the former, reason why they were initially called "large shell" and "small shell", respectively. Based on these molecular studies, the two genera colonized the archipelago in two different stages: around 16.5 mya the genus *Africonus* and 4.6 mya the genus *Trovaoconus* (Cunha et al., 2014). However, the phylogenetic relationships of these genera remain unclear. While in (Duda and Rolán, 2005) the closest relative to *Africonus* was *Lautoconus ventricosus* from the Mediterranean Sea, in (Cunha et al., 2005) the sister group of *Africonus* were the *Lautoconus* species from Senegal. None of these studies recovered the closest relative to the genus *Trovaoconus*. Puillandre et al. (2014) did not recover clearly whether *L. ventricosus* or *Lautoconus* endemic to Senegal were the closest relative to *Africonus*, but it suggested that the closest relative to *Trovaoconus* was the continental genus *Kalloconus,* which was recovered as paraphyletic in their phylogenetic tree.

It is interesting to note that despite having quite similar volcanic origins and geological properties, the archipelago of the Canary Islands has very little species described compared to Cabo Verde, and none of them is endemic. Cunha et al. (2014) studied the possibility of this species being actually a species complex. However, they found no genetic structure between populations, neither between islands or respect to the continent. Based on abiotic data collected, they propose that the temperature of the sea surface is a key factor determining the cladogenetic differences in both archipelagos. The close distance of Canary Islands to the continent compare to Cabo Verde could also have allowed or restricted gene flow in the former and the latter, respectively.

The only other available molecular study of West African cone snails was based on the *cox1* marker and was focused on the only species of cone endemic to Saint Helena Island. The authors found that the species could be included in the genus *Varioconus* (*Varioconus jordani*), a genus that was shown to be monophyletic, including several species from Angola, sister to the endemic species to Senegal (Tenorio et al., 2016).

In summary, we still lack a robust evolutionary framework for the cone species inhabiting this region, which is of paramount importance for understanding the evolutionary patterns driving cone snail evolution, including accurate divergence times, and ecological adaptations, such as diet or venom evolution. Moreover, it is important to understand the real species diversity in this region as this has important consequences for taking conservation actions. Peters et al. (2013) found out that all the threatened species of cone snails were from West Africa. The main factors leading to the reduction of their populations, were tourism, cities development, pollution, among others (Peters et al., 2016). However, conservation status depends on having an accurate catalogue of species. If the taxonomic status of one species of restricted distribution, which is declared threatened, is revised and now the species is considered just a population of another with wider distribution, it is likely that the status of vulnerability will be reduced. In contrast, the identification of cryptic species will require updating conservation actions in the opposite direction.

1.2.3. – Why is this region so interesting?

All of the features above explained make the cones from the East Atlantic region an ideal model system for studying and understanding evolution. More specifically:

This region presents, in a (relative) narrow area, almost 15% of the global diversity of cone snails. Most of these species diversified in a relative short period of time through radiation events, and could be seen as independent "evolutionary experiments" in words of Ernst Mayr. Also, while some lineages have great species richness, such as *Africonus* in the Cabo Verde archipelago, others show little diversification such as *Lautoconus guanche* in Canary Islands or *L. ventricosus* in the Mediterranean region. Hence, the region offers a great opportunity for trying to disentangle the factors that promote diversification within the group, for example, through comparative studies (Cunha et al., 2014).

The different geographical ranges of the species also allow us to understand the processes that end in local adaptation. For example, *L. ventricosus* is found throughout the Mediterranean Sea, *Trovaoconus* species can inhabit different Cabo Verde islands, but *Africonus* species are usually restricted to single islands or even particular bays.

What are the mechanisms or factors promoting and restricting the species dispersal and colonization ability ultimately leading to speciation?

Most of these species are endemic and the genera are closely related within the larger phylogeny of the family Conidae. Therefore, West African cones represent a relatively closed model system. It is very likely that the same evolutionary processes are acting in *Africonus*, *Lautoconus*, *Trovaoconus* and *Varioconus*. These species share a recent common ancestor, which should facilitate, for example, understanding how the venom is evolving in this group.

Thus far, and assuming there were two independent shifts to piscivory in cone evolutionary history (Puillandre et al., 2014), West Africa provides an incomparable opportunity to understand how this particular adaptation evolved. In Cabo Verde inhabits the only Atlantic piscivorous species (*C. ermineus*). By comparing fish-hunting cones in both oceans, Indo-Pacific and Atlantic, we could infer the common adaptations to this diet, that would be convergent, and we could understand better what conotoxins are important for hunting on fish.

Finally, the geological history of the archipelagos of this region is well documented due to their volcanic origin, and thus we can have an accurate chronological perspective of cone snails evolution, which could help us disentangling the otherwise frequently misleading fossil record of cones.

1.3. – General methodology

The papers presented in the result section contain each a detailed explanation of the methodology followed in the different studies. However, these Materials and methods sections were focused on ensuring the reproducibility of the results presented there, and lacked any further explanation on why that methodology is particularly appropriate to tackle the proposed questions.

I believe it is interesting understanding how the different methods have evolved through time, what kind of information we gain with them, and how they can provide new insights on the biology and evolution of the cone snails (and their venom). Therefore, this section pretends to be a summary on the methodology that is used in the field. At the same time, I think that having such temporal perspective also allows us to

understand why the applied methodology is a sound choice for the studies here accomplished.

1.3.1. – Phylogenomics: cone snail evolution

As explained above, cone snails represent one of the most specious marine families. This species diversity makes the group interesting for evolutionary studies but at the same time, it makes challenging the reconstruction of a robust phylogenetic-based classification of the group. Yet, a robust phylogenetic framework is of paramount importance for understanding the evolutionary mechanisms and processes leading to this diversity of species.

Traditionally, and as many other animal groups, conoidean snails have been classified based on morphology (Taylor et al., 1993), using mainly shell and radular characters, and cone snails have not been an exception (Röckel et al., 1980; Da Motta, 1991; Tucker and Tenorio, 2009). Based on shell characters, Da Motta (1991) proposed six genera and 60 subgenera for the family Conidae *sensu* Puillandre et al. (2014). Later, Tucker and Tenorio (2009) used radular characters to create four families and 69 genera, raised to 90 in (Tucker and Tenorio, 2013). In fact, in WoRMS are up to 129 different synonymized genera, 98 of them accepted in a subgeneric form. As explained above, even if these morphological characters can distinguish between genera and species, they fail to resolve the evolutionary relationships among them.

In order to overcome these limitations, molecular phylogenetics became the standard tool for studying evolution. After the first molecular phylogeny published of the Conoidea in the 1990s (Harasewych et al., 1997), many others were inferred to answer different evolutionary questions (Duda and Palumbi, 1999a; Duda and Kohn, 2005; Duda and Rolán, 2005; Puillandre et al., 2014). However, all these phylogenies failed to recover a fully resolved and statistically supported topology. As suggested by the fossil record, cone snails have evolved through several radiation events, which produce short internal nodes in the trees (Philippe et al., 2011). Resolving these short nodes require abundant phylogenetic information, and using only few partial sequences of PCR-amplified genes, although meritorious at the time, was not enough in most cases. It was not until the advent of Next Generation Sequencing (NGS) techniques in

the last decade that obtaining phylogenetic markers at a genome scale has become feasible even for non-model organisms.

This new sequencing potential has certainly improved our capacity to reconstruct robust phylogenies even of complex radiations, but it also has brought new challenges: defining correctly groups of orthology, developing new algorithms to process the data, having enough computer power to analyse these datasets, among many others (Hedtke et al., 2006; Jeffroy et al., 2006). A new field, phylogenomics, has born to deal with such problems and is for the time being the best option in trying to resolve the most recalcitrant nodes of the tree of life (Kocot et al., 2011; Smith et al., 2011; Zapata et al., 2014), although it is on its infancy and there is yet a lot of work to be done.

One alternative to avoid some of the problems associated to the use of nuclear markers at a genome scale, is working with mitochondrial genomes (or mitogenomes). These markers present several advantages for phylogenetic reconstruction including higher evolutionary rates than nuclear genomes; enough sequence information/ phylogenetic signal but reasonably small size (about 16 Kb); single-copy genes, which make straightforward the orthology assignment; they can be obtained through direct NGS sequencing but also through long PCRs. Therefore, mitogenomes have been widely used in different animal groups to solve evolutionary questions at different taxonomic levels (Irisarri et al., 2012; Irisarri et al., 2014; Osca et al., 2014a; Osca et al., 2014b; San Mauro et al., 2014; Osca et al., 2015).

The age of the West African archipelagos, where inhabits the majority of the species, is younger than 25 my, and previous studies have suggested that the origin of the insular species is younger than 20 my (Cunha et al., 2005; Duda and Rolán, 2005). Therefore, using the mitochondrial genome as choice of marker to infer phylogenetic relationships of West African cones is an appropriate approach, as we would use it at its best performance level ensuring maximum resolution.

In order to resolve older phylogenetic relationships such as those among genera within the family Conidae, mitochondrial genomes are at their limit of resolution (Talavera and Vila, 2011; Stoger and Schrodl, 2013). In this case, it is possible to still achieve fair results by applying more complex evolutionary models that take into consideration rate heterogeneity and reduce the well-known long-branch attraction error (Irisarri et al., 2012). However, best results are achieved if mitochondrial genomes are

combined with tens or hundreds of nuclear genes, which are more suitable for phylogenetic inference at these deep levels (Lin and Danforth, 2004; Dool et al., 2016).

In the particular case of the family Conidae, the advent of NGS has facilitated the determination of the conotoxin precursors produced in the venom gland by transcriptome analyses (see below). These transcriptomes are an excellent source for nuclear genes that could be used in phylogenomics. Hence, we propose a phylogenomic pipeline as depicted in Fig. 5.

Another crucial and limiting factor in phylogenetic reconstruction is the taxon sampling. Cone snails diversity and distribution make difficult having representatives of all the genera in a single phylogenetic analysis, which ultimately could affect the resolution of the reconstructed tree. At the beginning of this work, there were available seven mitogenomes of cone snails, thus clearly underrepresenting the diversity of the family.

Table 1 - Catalogue of publicly available cone snail mitogenomes as of 2016.

Species	Acc. Number	Release date	Reference
Cylinder textile	NC_008797	25/01/2007	(Bandyopadhyay et al., 2008)
Africonus borgesi	NC_013243	07/07/2010	(Cunha et al., 2009)
Pionoconus consors	NC_023460	27/02/2014	(Brauer et al., 2012)
Kioconus tribblei	NC_027957	12/04/2016	(Barghi et al., 2016)
Cylinder gloriamaris	NC_030213	31/05/2016	Unpublished
Gastridium tulipa	NC_027518	17/05/2016	(Chen et al., 2016)
Rhizoconus capitaneus	NC_030354	17/06/2016	Unpublished

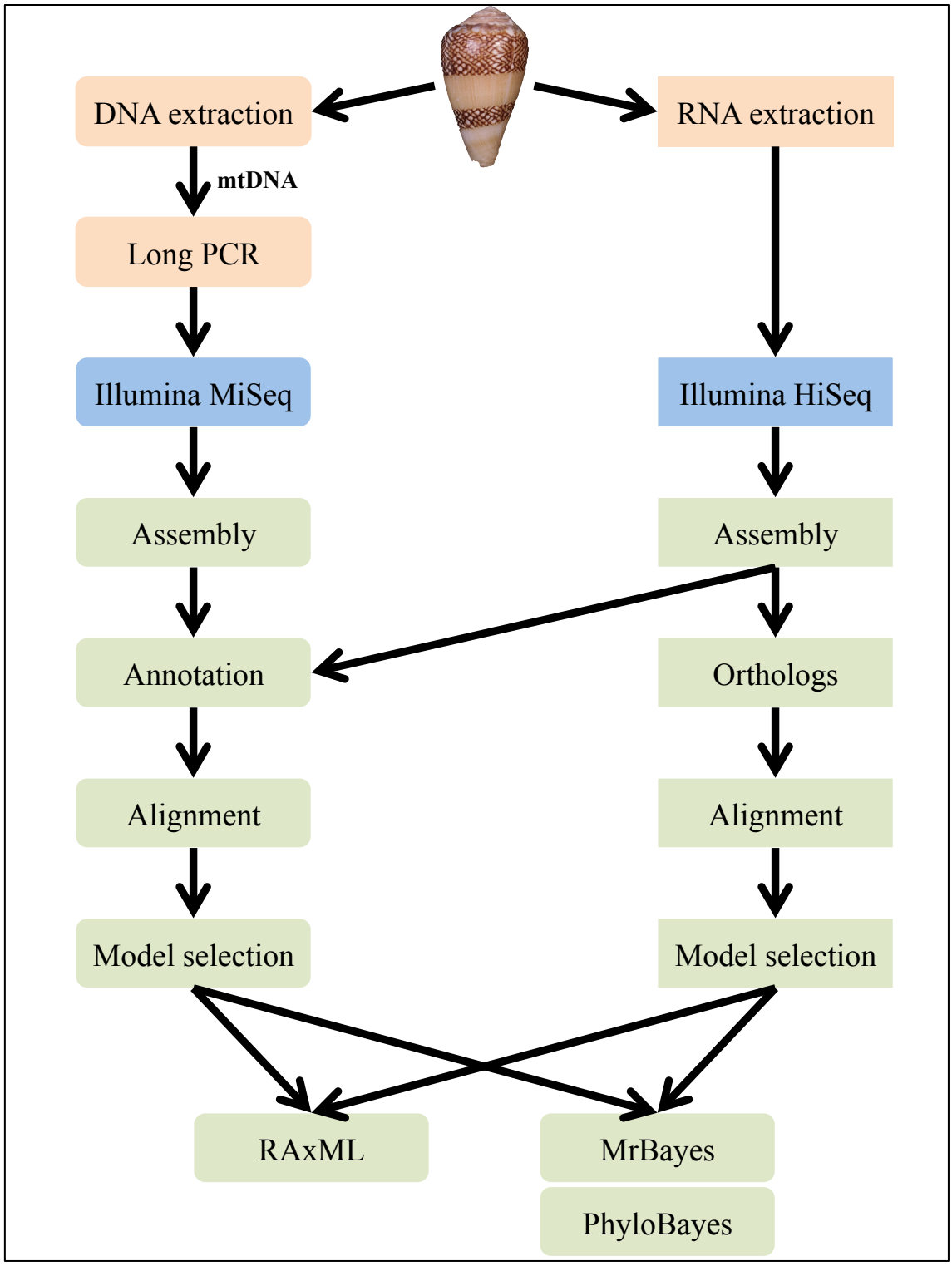


Fig. 5 – Main steps of the pipeline followed in the phylogenomic reconstruction. In orange, wet lab steps; in blue the outsourced sequencing; and in green, the bioinformatics analyses. The cone shell belongs to *Lautoconus mercator*. Picture taken by Manuel J. Tenorio

After five years and seven sampling trips to Cabo Verde and Senegal, we collected more than 1150 specimens, representing all the genera, most of the species and most of the populations of West African cone snails (not including the Mediterranean Sea). We PCR amplified and sequenced using Illumina MiSeq a total of 114 mitogenomes. In addition, since 2016, several mitogenomes of cone snails have been sequenced (Uribe et al., 2017; Uribe et al., 2018) and the transcriptomes of others have become available and could be use as source for assembling mitochondrial genes.

1.3.2. – Venom evolution through transcriptomics

Cone snails have been long known to be venomous, since they have caused several human fatalities (Kohn, 2016). As explained above, the venom of cone snails has important physiological properties, and it has drawn the attention of pharmacological researchers and companies. There are tens of thousands of different conotoxins that, potentially, could lead to the development of new drugs. Because of that (and its inherent biological interest) cone venom has been studied for decades.

At the beginning, many studies characterized histologically the venom gland, in trying to understand the biological mechanisms of conotoxin production (Kohn et al., 1960; Endean and Duchemin, 1967; Songdahl, 1973; Freeman et al., 1974). At the same time, biological studies started by injecting the extracted venom of different species of cone snails into model system organisms and observing the induced physiological responses (Kohn et al., 1960; Endean and Izatt, 1965; Endean and Rudkin, 1965; Cottrell and Twarog, 1972). This kind of studies continued for decades, providing invaluable information about the physiological effect of many different toxins (e.g. (Olivera et al., 1985), leading to the discovery of the MVIIA, the base of Ziconotide, the only conotoxin-based drug currently in use (Olivera et al., 1987).

More recently, the improvement of the proteomic techniques offered a new level of detail of the venom composition. On one hand, the proteome throughput techniques (MS-LC, MS/MS) are able to identify the whole set of proteins that are being secreted, thus providing an informative portrait of the actual venom used by the snail. This information allows making valuable inferences about the ecology of cone snails, and, for example, the interesting double-use of the venom (predation- or defence-evoked; (Dutertre et al., 2010; Prashanth et al., 2012; Dutertre et al., 2013; Dutertre et al., 2014;

Safavi-Hemami et al., 2014). On the other hand, thanks to the advances in peptide synthesis we can test more accurately the effect of particular toxins (Safavi-Hemami et al., 2015).

In parallel, the study of conotoxins by using the complementary DNA (cDNA) of the messenger RNA (mRNA) provided a new molecular perspective, shedding light over the mechanisms of conotoxin evolution (e.g. (Duda and Palumbi, 1999b; Duda and Palumbi, 2004; Santos et al., 2004; Luo et al., 2006; Pi et al., 2006a; Pi et al., 2006b; Remigio and Duda, 2008). This new methodology allowed sequencing tens of conotoxins from several species at the same time, and helped understanding, for example, the structure of the conotoxin precursors and the mechanisms generating the conotoxin variability.

Upon the advent of the NGS techniques a new molecular perspective was adopted. The transcriptome sequencing (first using 454 GS FLX titanium platforms, then Illumina HiSeq) outperformed completely the sequencing of cDNA, recovering as many as one order of magnitude more conotoxins (Hu et al., 2012; Terrat et al., 2012; Lavergne et al., 2013; Barghi et al., 2015). Because of the sensitivity of transcriptomics, we can identify even the less expressed transcripts, thus getting the full repertoire of conotoxins produced by one individual (even those not being secreted in the venom). Besides, transcriptomic data also reflect the level of expression of the different transcripts. Therefore, using enough biological replicates, we can infer the relative importance of the different conotoxins (or conotoxin superfamilies) in the final venom cocktail, and compare it between species, genera or diets.

Here, I am interested in uncovering the evolutionary patterns that gave rise to the astonishing diversity of conotoxins. As explained above, the most complete tool we have to infer venom composition, both in terms of cataloguing the different conotoxins and of determining their relative importance within the venom, is transcriptomics. We sequenced 25 transcriptomes from 13 species, nine from *Chelyconus ermineus* (three specimens, three samples per specimen), four from the genus *Lautoconus* from Senegal and the Canary Islands (two samples of *Lautoconus mercator*), and 10 from the genus *Africonus* (nine species, two specimens of *Africonus maioensis*).

1.4. – References

Aman, J.W., Imperial, J.S., Ueberheide, B., Zhang, M.-M., Aguilar, M., Taylor, D., Watkins, M., Yoshikami, D., Showers-Corneli, P., Safavi-Hemami, H., Biggs, J., Teichert, R.W., Olivera, B.M., 2014. Insights into the origins of fish hunting in venomous cone snails from studies of *Conus tessulatus*. Proceedings of the National Academy of Sciences 112, 5087-5092.

Ancochea, E., Fuster, J.M., Ibarrola, E., Cendrero, A., Coello, Juan, Hernan, F., Cantagrel, J.M., Jamond, C., 1990. Volcanic evolution of the island of Tenerige (Canary Islands) in the light of new K-Ar data. Journal of Volcanology and Geothermal Research 44, 231 - 249.

Anderson, P.D., Bokor, G., 2012. Conotoxins: Potential Weapons from the Sea. Journal of Bioterrorism & Biodefense 03, 1-4.

Bandyopadhyay, P.K., Colledge, C.J., Walker, C.S., Zhou, L.-M., Hillyard, D.R., Olivera, B.M., 1997. Conantokin-G precursor and its role in y.Carboxylation by a vitamin K-dependent carboxylase from a Conus snail. The Journal of Biological Chemistry 273, 5447-5450.

Bandyopadhyay, P.K., Stevenson, B.J., Ownby, J.P., Cady, M.T., Watkins, M., Olivera, B.M., 2008. The mitochondrial genome of *Conus textile*, coxI-coxII intergenic sequences and Conoidean evolution. Molecular Phylogenetics and Evolution 46, 215-223.

Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2015. High conopeptide diversity in *Conus tribblei* revealed through analysis of venom duct transcriptome using two High-Throughput sequencing platforms. Marine Biotechnology 17, 81-98.

Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2016. Characterization of the complete mitochondrial genome of *Conus tribblei* Walls, 1977. Mitochondrial DNA Part A 27(6), 4451-4452.

Berschauer, D.P., 2015. A comparison of adaptive radiation in Conidae and Conilithidae (Gastropoda: Conoidea) in the Eastern and Western Atlantic, together with an iconography of the conilithid genus *Jaspidiconus*. The Festivus 47, 99-113.

Brauer, A., Kurz, A., Stockwell, T., Baden-Tillson, H., Heidler, J., Wittig, I., Kauferstein, S., Mebs, D., Stocklin, R., Remm, M., 2012. The mitochondrial genome of the venomous cone snail *Conus consors*. PLoS One 7, e51528.

Briggs, J.C., Bowen, B.W., 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. Journal of Biogeography 39, 12-30.

Buczek, O., Olivera, B.M., Bulaj, G., 2004. Propeptide does not act as an intramolecular chaperone but facilitates protein disulfide isomerase-assisted folding of a conotoxin precursor. Biochemistry 43, 1093 - 1101.

Carine, M.A., Santos Guerra, A., Guma, R.R., Reyes-Betancourt, J.A., 2010. Endemisms and evolution of the Macaronesian flora. In: Williams, D.M., Knapp, S. (Eds.), Beyond cladistics: the branching of a paradigm. University of California Press, pp. 101 - 124.

Carminati, E., Lustrino, M., Doglioni, C., 2012. Geodynamic evolution of the central and western Mediterranean: Tectonics vs. igneous petrology constraints. Tectonophysics 579, 173-192.

Castelin, M., Puillandre, N., Kantor, Y.I., Modica, M.V., Terryn, Y., Cruaud, C., Bouchet, P., Holford, M., 2012. Macroevolution of venom apparatus innovations in auger snails (Gastropoda; Conoidea; Terebridae). Molecular Phylogenetics and Evolution 64, 21-44.

Caze, B., Merle, D., Meur, M.L., Pacaud, J.-M., Ledon, D., Martin, J.-P.S., 2011. Taxonomic Implications of the Residual Colour Patterns of Ampullinid Gastropods and Their Contribution to the Discrimination from Naticids. Acta Palaeontologica Polonica 56, 329-347.

Caze, B., Merle, D., Pacaud, J.-M., Saint Martin, J.-P., 2010. First Systematic Study using the Variability of the Residual Colour Patterns: The Case of the Paleogene Seraphsidae (Mollusca, Gastropoda, Stromboidea). Geodiversitas 32, 417-477.

Chang, D., Duda, T.F., Jr., 2012. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Molecular Biolology and Evolution 29, 2019-2029.

Chang, D., Duda, T.F., Jr., 2016. Age-related association of venom gene expression and diet of predatory gastropods. BMC Evolutionary Biology 16, 27.

Chen, P.-W., Hsiao, S.-T., Huang, C.-W., Chen, K.-S., Tseng, C.-T., Wu, W.-L., Hwang, D.-F., 2016. The complete mitochondrial genome of *Conus tulipa* (Neogastropoda: Conidae). Mitochondrial DNA Part A 27(4), 2738-2739.

Conticello, S.G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z., Fainzilber, M., 2001. Mechanisms for evolving hypervariability: The case of conopeptides. Molecular Biology and Evolution 18, 120- 131.

Cottrell, G.A., Twarog, B.M., 1972. Active factors in the venom duct of *Conus californicus*. British journal of pharmacology 44, 365P.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde islands. Systematic Biology 54, 634-650.

Cunha, R.L., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evolutionary Biology 9, 210.

Cunha, R.L., Lima, F.P., Tenorio, M.J., Ramos, A.A., Castilho, R., Williams, S.T., 2014. Evolution at a Different Pace: Distinctive Phylogenetic Patterns of Cone Snails from Two Ancient Oceanic Archipelagos. Systematic Biology 63, 971-987.

Cunha, R.L., Tenorio, M.J., Afonso, C., Castilho, R., Zardoya, R., 2008. Replaying the tape: recurring biogeographical patterns in Cape Verde *Conus* after 12 million years. Molecular Ecology 17, 885-901.

Da Motta, A.J., 1991. A systematic classification of the gastropod family Conidae at the generic level. La Conchiglia.

Davis, K.E., Hill, J., Astrop, T.I., Wills, M.A., 2016. Global cooling as a driver of diversification in a major marine clade. Nature Communications 7, 13003.

Dimitrov, D., Arnedo, M.A., Ribera, C., 2008. Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence for long-term occupancy yet rapid recent speciation. Molecular Phylogenetics and Evolution 48, 596-614.

Dool, S.E., Puechmaille, S.J., Foley, N.M., Allegrini, B., Bastian, A., Mutumi, G.L., Maluleke, T.G., Odendaal, L.J., Teeling, E.C., Jacobs, D.S., 2016. Nuclear introns outperform mitochondrial DNA in inter-specific phylogenetic reconstruction: Lessons from horseshoe bats (Rhinolophidae: Chiroptera). Molecular Phylogenetics and Evolution 97, 196-212.

Duda, T.F., Bolin, M.B., Meyer, C.P., Kohn, A.J., 2008. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. Molecular Phylogenetics and Evolution 49, 867-876.

Duda, T.F., Jr., Chang, D., Lewis, B.D., Lee, T., 2009a. Geographic variation in venom allelic composition and diets of the widespread predatory marine gastropod *Conus ebraeus*. PLoS One 4, e6245.

Duda, T.F., Jr., Palumbi, S.R., 2004. Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*. Proceedings of the Royal Society B: Biological Sciences 271, 1165-1174.

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., Kohn, A.J., Matheny, A.M., 2009b. Cryptic species differentiated in *Conus ebraeus*, a widespread tropical marine gastropod. Bioogical Bulletin 217, 292-305.

Duda, T.F., Kohn, A.J., Palumbi, S., R., 2001. Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. Biological Journal of the Linnean Society 73, 391-409.

Duda, T.F., Lee, T., 2009. Isolation and population divergence of a widespread Indo-West Pacific marine gastropod at Easter Island. Marine Biology 156, 1193-1202.

Duda, T.F., Palumbi, S.R., 1999a. Developmental shifts and species selection in gastropods. Proceedings of the National Academy of Sciences 96, 10272 - 10277.

Duda, T.F., Palumbi, S.R., 1999b. 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.

Duda, T.F., Rolán, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Molecular Ecology 14, 267-272.

Dutertre, S., Biass, D., Stocklin, R., Favreau, P., 2010. Dramatic intraspecimen variations within the injected venom of *Conus consors*: an unsuspected contribution to venom diversity. Toxicon 55, 1453-1462.

Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nature Communiactions 5, 3521.

Dutertre, S., Jin, A.H., Kaas, Q., Jones, A., Alewood, P.F., Lewis, R.J., 2013. Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. Molecular & Cell Proteomics 12, 312-329.

Dyhr, C.T., Holm, P.M., 2010. A volcanological and geochemical investigation of Boa Vista, Cape Verde Islands; 40Ar/39Ar geochronology and field constraints. Journal of Volcanology and Geothermal Research 189, 19-32.

Endean, R., Duchemin, C., 1967. The venom apparatus of *Conus magus*. Toxicon 4, 275-278.

Endean, R., Izatt, J., 1965. Pharmacological study of the venom of the gastropod *Conus magus*. Toxicon 3, 81-93.

Endean, R., Rudkin, C., 1965. Further studies of the venoms of Conidae. Toxicon 2, 225-249.

Espiritu, D.J.D., Watkins, M., Dia-Monje, V., Cartier, G.E., Cruz, L.J., Olivera, B.M., 2001. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon 39, 1899-1916.

Freeman, S.E., Turner, R.J., Silva, S.R., 1974. The venom and venom apparatus of the marine gastropod *Conus striatus* Linne. Toxicon 12, 587-592.

Garber, K., 2005. Peptide leads new class of chronic pain drugs. Nature Biotechnology 23, 399.

Guillou, H., Carracedo, J.C., TPérez Torrado, F., Rodríguez Badiola, E., 1996. K-Ar ages and magnetic stratigraphy of a hotspot-induced, fast grown oceanic island:

El hierro, Canary Islands. Journal of Volcanology and Geothermal Research 73, 141 - 155.

Harasewych, M.G., Adamkewych, S.L., Blake, J.A., Saudek, D., Spriggs, T., Bult, C.J., 1997. Neogastropod phylogeny: a molecular perspective. Journal of Molluscan Studies 63, 327-351.

Hedtke, S.M., Townsend, T.M., Hillis, D.M., 2006. Resolution of Phylogenetic Conflict in Large Data Sets by Increased Taxon Sampling. Systematic Biology 55, 522 - 529.

Hendricks, J.R., 2015. Glowing seashells: Diversity of fossilized coloration patterns on coral reef-associated cone snail (Gastropoda: Conidae) shells from the Neogene of the Dominican Republic. PLoS One 10, e0120924.

Hendricks, J.R., 2018. Diversity and preserved shell coloration patterns of Miocene Conidae (Neogastropoda) from an exposure of the Gatun Formation, Colón Province, Panama. Journal of Paleontology 92, 804 - 837.

Holford, M., Puillandre, N., Terryn, Y., Cruaud, C., Olivera, B., Bouchet, P., 2009. Evolution of the Toxoglossa venom apparatus as inferred by molecular phylogeny of the Terebridae. Molecular Biology and Evolution 26, 15-25.

Holm, P.M., Grandvuinet, T., Friis, J., Wilson, J.R., Barker, A.K., Plesner, S., 2008. An40Ar-39Ar study of the Cape Verde hot spot: Temporal evolution in a semistationary plate environment. Journal of Geophysical Research 113.

Hu, H., Bandyopadhyay, P.K., Olivera, B.M., Yandell, M., 2012. Elucidation of the molecular envenomation strategy of the cone snail *Conus geographus* through transcriptome sequencing of its venom duct. BMC Genomics 13, 284.

Irisarri, I., Eernisse, D.J., Zardoya, R., 2014. Molecular phylogeny of Acanthochitonina (Mollusca: Polyplacophora: Chitonida): three new mitochondrial genomes, rearranged gene orders and systematics. Journal of Natural History 48, 2825-2853.

Irisarri, I., San Mauro, D., Abascal, F., Ohler, A., Vences, M., Zardoya, R., 2012. The origin of modern frogs (Neobatrachia) was accompanied by acceleration in mitochondrial and nuclear substitution rates. BMC Genomics 13, 626.

Jablonski, D., Lutz, R.A., 1983. Larval ecology of marine benthic invertebrates: paleobiological implications. Biological Reviews 58, 21 - 89.

Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends in Genetics 22, 225-231.

Jin, A.H., Dutertre, S., Kaas, Q., Lavergne, V., Lewis, R.J., Alewood, P.F., 2013. Transcriptomic messiness in the venom duct of Conus miles contributes to conotoxin diversity. Molecular & Cellular Proteomics 12, 3824-3833.

Kaas, Q., Westermann, J.C., Craik, D.J., 2010. Conopeptide characterization and classifications: an analysis using ConoServer. Toxicon 55, 1491-1509.

Kantor, Y.I., Puillandre, N., 2012. Evolution of the radular apparatus in Conoidea (Gastropoda: Neogastropoda) as inferred from a molecular phylogeny. Malacologia 55, 55-90.

Kase, T., Kitao, F., Aguilar, Y.M., Kurihara, Y., Pandita, H., 2008. Reconstruction of color markings in Vicarya, a Miocene potamidid gastropod (Mollusca) from SE Asia and Japan. Paleontological Research 12, 345-353.

Kocot, K.M., Cannon, J.T., Todt, C., Citarella, M.R., Kohn, A.B., Meyer, A., Santos, S.R., Schander, C., Moroz, L.L., Lieb, B., Halanych, K.M., 2011. Phylogenomics reveals deep molluscan relationships. Nature 477, 452-456.

Kohn, A.J., 1956. Piscivorous gastropods of the genus *Conus*. Proceedings of the National Academy of Sciences 42, 168-171.

Kohn, A.J., 1959. The Ecology of *Conus* in Hawaii. Ecological Monographs 29, 47-90.

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

Kohn, A.J., 2016. Human injuries and fatalities due to venomous marine snails of the family Conidae. International journal of clinical pharmacology and therapeutics 54, 524-538.

Kohn, A.J., Nishi, M., Pernet, B., 1999. Snail spears and scimitars: a character analysis of *Conus* radular teeth. Journal of Molluscan Studies 65, 461-481.

Kohn, A.J., Perron, F.E., 1994. Life history and biogeography: patterns in *Conus*. Oxford University Press on Demand.

Kohn, A.J., Saunders, P.R., Wiener, S., 1960. Preliminary studies on the venom of the marine snail *Conus*. Annals of the New York Academy of Sciences 90, 706-725.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652.

Krueger, K.K., 1974. The use of ultraviolet light in the study of fossil shells. Curator: The Museum Journal 17, 36 - 49.

Lavergne, V., Dutertre, S., Jin, A.-H., Lewis, R.J., Taft, R.J., Alewood, P.F., 2013. Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. BMC Genomics 14, 708.

Lebbe, E.K., Tytgat, J., 2016. In the picture: disulfide-poor conopeptides, a class of pharmacologically interesting compounds. Journal of Venomous Animals and Toxins including Tropical Diseases 22, 30.

Lin, C.-P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Molecular Phylogenetics and Evolution 30, 686-702.

Luo, S., Zhangsun, D., Wu, Y., Zhu, X., Xie, L., Hu, Y., Zhang, J., Zhao, X., 2006. Identification and molecular diversity of T-superfamily conotoxins from *Conus lividus* and *Conus litteratus*. Chemical Biology & Drug Design 68, 97-106.

McIntosh, J.M., Jones, R.M., 2001. Cone venom-from accidental stings to deliberate injection. Toxicon 39, 1447 - 1451.

Miljanich, G.P., 2004. Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. Current medicinal chemistry 11, 3029-3040.

Monnier, E., Limpalaër, L., 2010. *Conus dorotheae* (Gastropoda, Conidae) a new species of cone from the Cape Verde Peninsula in Senegal. Visaya 3.

Monnier, E., Tenorio, M.J., 2015. *Malagasyconus* (Gastropoda: Conidae), a new genus from Southern Madagascar. Xenophora Taxonomy 6, 47 - 51.

Monteiro, A., Tenorio, M.J., Poppe, G.T., 2004. The family Conidae, the west African and Mediterranean species of *Conus*. ConchBooks, Hackenheim, Germany.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots or conservation priorities. Nature 403, 853 - 858.

Nolf, F., Verstraeten, J., 2008. *Conus trencarti* (Mollusca: Gastropoda: Conidae): a new cone from Senegal. Neptunea 7, 1 - 13.

Norton, R., S., Olivera, B.M., 2006. Conotoxins down under. Toxicon 48, 780-798.

Olivera, B.M., 2006. *Conus* peptides: biodiversity-based discovery and exogenomics.The Journal of Biological Chemistry 281, 31173-31177.

Olivera, B.M., Cruz, L.J., de Santos, V., LeCheminant, G.W., Griffin, D., Zeikus, R., McIntosh, J.M., Galyean, R., Varga, J., Gray, W.R., Rivier, J., 1987. Neuronal calcium channel antagonists. Discrimination between calcium channel subtypes using w-Conotoxin from *Conus magus* venom. Biochemistry 26, 2086 - 2090.

Olivera, B.M., McIntosh, J.M., Clark, C., Middlemas, D., Gray, W.R., Cruz, L.J., 1985. A sleep-inducing peptide from *Conus geographus* venom. Toxicon 23, 277 - 282.

Olivera, B.M., Rivier, J., Clark, C., Ramilo, C.A., Corpuz, G.P., Abogadie, F.C., Mena, E.E., Woodward, S.R., Hillyard, D.R., Cruz, L.J., 1990. Diversity of *Conus* neuropeptides. Science 249, 257 - 263.

Olivera, B.M., Seger, J., Horvath, M.P., Fedosov, A.E., 2015. Prey-Capture Strategies of Fish-Hunting Cone Snails: Behavior, Neurobiology and Evolution. Brain, Behaviour and Evolution 86, 58-74.

Olsson, A.A., 1967. Some Tertiary mollusks from south Florida and the Caribbean. Paleontological Research Institution.

Osca, D., Irisarri, I., Todt, C., Grande, C., Zardoya, R., 2014a. The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca: Chaetodermomorpha) supports the Aculifera hypothesis. BMC Genomics 14, 197.

Osca, D., Templado, J., Zardoya, R., 2014b. The mitochondrial genome of *Ifremeria nautilei* and the phylogenetic position of the enigmatic deep-sea Abyssochrysoidea (Mollusca: Gastropoda). Gene 547, 257-266.

Osca, D., Templado, J., Zardoya, R., 2015. Caenogastropoda mitogenomics. Molecular Phylogenetics and Evolution 93, 118-128.

Peng, C., Yao, G., Gao, B.-M., Fan, C.-X., Bian, C., Wang, J., Cao, y., Wen, B., zhu, Y., Ruan, Z., Zhao, X., You, X., Bai, J., Li, J., Lin, Z., Zou, S., Zhang, X., Qiu, Y., Chen, J., Coon, S.L., Yang, J., Chen, J.-S., Shi, Q., 2016. High-throughput identification of novel conotoxins from the Chinese tubular cone snail (*Conus betulinus*) by multitranscriptome sequencing. GigaScience 5, 17.

Peters, H., O'Leary, B.C., Hawkins, J.P., Carpenter, K.E., Roberts, C.M., 2013. *Conus*: first comprehensive conservation red list assessment of a marine gastropod mollusc genus. PLoS One 8, e83353.

Peters, H., O'Leary, B.C., Hawkins, J.P., Roberts, C.M., 2016. The cone snails of Cape Verde: Marine endemism at a terrestrial scale. Global Ecology and Conservation 7, 201-213.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: Why more sequences are not enough. PLoS Biology 9, e1000602.

Phuong, M.A., Mahardika, G.N., Alfaro, M.E., 2016. Dietary breadth is positively correlated with venom complexity in cone snails. BMC Genomics 17, 401.

Pi, C., Liu, J., Peng, C., Liu, Y., Jiang, X., Zhao, Y., Tang, S., Wang, L., Dong, M., Chen, S., Xu, A., 2006a. Diversity and evolution of conotoxins based on gene expression profiling of *Conus litteratus*. Genomics 88, 809-819.

Pi, C., Liu, Y., Peng, C., Jiang, X., Liu, J., Xu, B., Yu, X., Yu, Y., Jiang, X., Wang, L., Dong, M., Chen, S., Xu, A.L., 2006b. Analysis of expressed sequence tags from the venom ducts of *Conus striatus*: focusing on the expression profile of conotoxins. Biochimie 88, 131-140.

Pin, M., Leung Tack, K.D., 1995. The Cones of Senegal. Conchiglia.

Prashanth, J.R., Dutertre, S., Jin, A.H., Lavergne, V., Hamilton, B., Cardoso, F.C., Griffin, J., Venter, D.J., Alewood, P.F., Lewis, R.J., 2016. The role of defensive ecological interactions in the evolution of conotoxins. Molecular Ecology 25, 598-615.

Prashanth, J.R., Lewis, R.J., Dutertre, S., 2012. Towards an integrated venomics approach for accelerated conopeptide discovery. Toxicon 60, 470-477.

Prator, C.A., Murayama, K.M., Schulz, J., R., 2014. Venom variation during prey capture by the cone snail, *Conus textile*. PLoS One 9, e98991.

Puillandre, N., Bouchet, P., Duda, T.F., Jr., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Molecular Phylogenetics and Evolution 78, 290-303.

Puillandre, N., Duda, T.F., Meyer, C., Olivera, B.M., Bouchet, P., 2015. One, four or 100 genera? A new classification of the cone snails. Journal of Molluscan Studies 81, 1-23.

Puillandre, N., Koua, D., Favreau, P., Olivera, B.M., Stocklin, R., 2012. Molecular phylogeny, classification and evolution of conopeptides. Journalof Molecular Evolution 74, 297-309.

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

Ramalho, R.S., Quartau, R., Trenhaile, A.S., Mitchell, N.C., Woodrofe, C.D., 2013. Coastal evolution on volcanic oceanic islands: a complex interplay between volcanism, erosion, sedimentation, sea-level change and biogenic production. Earth-Science Reviews 127, 140-170.

Remigio, E.A., Duda, T.F., Jr., 2008. Evolution of ecological specialization and venom of a predatory marine gastropod. Molecular Ecology 17, 1156-1162.

Rijsdijk, K.F., Hengl, T., Norder, S.J., Otto, R., Emerson, B.C., Ávila, S.P., López, H., van Loon, E.E., Tjørve, E., Fernández-Palacios, J.M., Heaney, L., 2014. Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. Journal of Biogeography 41, 1242-1254.

Robinson, S.D., Norton, R.S., 2014. Conotoxin gene superfamilies. Marine Drugs 12, 6058-6101.

Röckel, D., Rolan, E., Monteiro, A., 1980. Cone shells from Cape Verde islands. A difficult puzzle. I. Feito, Vigo, Spain.

Rolán, E., 1990. Descripción de nuevas especies y subespecies del género *Conus* (Mollusca, Neogastropoda) para el archipiélago de Cabo Verde. Iberus Supplement 2, 5-70.

Romero, H.K., Christensen, S.B., Mannelli, R.D.C., Gajewiak, J., Ramachandra, r., Elmslie, K.S., Vetter, D.E., Ghelardini, C., Iadonato, S.P., Mercado, J.K., Olivera, B.M., McIntosh, J.M., 2017. Inhibition of a9a10 nicotinic acetylcholine receptors prevents chemotherapy-induced neuropathic pain. Proceedings of the National Academy of Sciences114, E1825–E1832.

Safavi-Hemami, H., Gajewiak, J., Karanth, S., Robinson, S.D., Ueberheide, B., Douglass, A.D., Schlegel, A., Imperial, J.S., Watkins, M., Bandyopadhyay, P.K., Yandell, M., Li, Q., Purcell, A.W., Norton, R.S., Ellgaard, L., Olivera, B.M., 2015. Specialized insulin is used for chemical warfare by fish-hunting cone snails. Proceedings of the National Academy of Sciences 112, 1743 - 1748.

Safavi-Hemami, H., Hu, H., Gorasia, D.G., Bandyopadhyay, P.K., Veith, P.D., Young, N.D., Reynolds, E.C., Yandell, M., Olivera, B.M., Purcell, A.W., 2014. Combined proteomic and transcriptomic interrogation of the venom gland of *Conus geograhus* uncovers novel components and functional compartmentalization. Molecular & Cell Proteomics 13, 938-953.

San Mauro, D., Gower, D.J., Muller, H., Loader, S.P., Zardoya, R., Nussbaum, R.A., Wilkinson, M., 2014. Life-history evolution and mitogenomic phylogeny of caecilian amphibians. Molecular Phylogenetics and Evolution 73, 177-189.

Santos, A.D., McIntosh, J.M., Hillyard, D.R., Cruz, L.J., Olivera, B.M., 2004. The A-superfamily of conotoxins: structural and functional divergence. Journal of Biological Chemistry 279, 17596-17606.

Smith, S.A., Wilson, N.G., Goetz, F.E., Feehery, C., Andrade, S.C., Rouse, G.W., Giribet, G., Dunn, C.W., 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480, 364-367.

Songdahl, J.H., 1973. The venom and venom apparatus of the Atlantic cone, *Conus spurius atlanticus* (Clench). Bulletin of Marine Sciences 23, 600 - 612.

Stanley, S.M., 2008. Predation defeats competition on the seafloor. Paleobiology 34, 1-21.

Stoger, I., Schrodl, M., 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). Molecular Phylogenetics and Evolution 69, 376-392.

Talavera, G., Vila, R., 2011. What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. BMC Evolutionary Biology 11, 315.

Taylor, J.D., Kantor, Y.I., Sysoev, A.V., 1993. Foregut anatomy, feeding mechanisms, relationships and classification of the Conoidea (= Toxoglossa) (Gastropoda). Bulletin of the Natural History Museum, London (Zoology) 59, 125 - 170.

Taylor, J.D., Morris, N.J., Taylor, C.N., 1980. Food specialization and the evolution of predatory prosobranch gastropods. Palaeontology 23, 375 - 490.

Tenorio, M.J., Afonso, C.M.L., Cunha, R.L., Rolán, E., 2014. New species of *Africonus* (Gastropoda, Conidae) from Boa Vista in the Cape Verde Archipelago: Molecular and Morphological Characterization. Xenophora Taxonomy 2, 5-21.

Tenorio, M.J., Lorenz, F., Dominguez, M., 2016. New insights into *Conus jourdani* da Motta, 1984 (Gastropoda, Conidae), an endemic species from Saint Helena Island. Xenophora Taxonomy 11, 32-42.

Terlau, H., Olivera, B.M., 2004. Conus venoms: A rich source of novel ion channel-targeted peptides. Physiological Reviews 84, 41-68.

Terrat, Y., Biass, D., Dutertre, S., Favreau, P., Remm, M., Stöcklin, R., Piquemal, D., Ducancel, F., 2012. High-resolution picture of a venom gland transcriptome: Case study with the marine snail *Conus consors*. Toxicon 59, 34-46.

Tracey, S., Craig, B., Belliard, L., Gain, O., 2017. One, four or forty species? - early Conidae (Mollusca, Gastropoda) that led to a radiation and biodiversity peak in the late Lutetian Eocene of the Cotentin, NW France. Carnets de Voyages Paléontologiques dans le Bassin Anglo-Parisien 3, 1-38.

Tucker, J.K., Tenorio, M.J., 2009. Systematic Classification of Recent and Fossil Conoidean Gastropods, with keys to the genera of cone shells. ConchBooks.

Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MdM Publishing.

Uribe, J.E., Puillandre, N., Zardoya, R., 2017. Beyond *Conus*: Phylogenetic relationships of Conidae based on complete mitochondrial genomes. Molecular Phylogenetics and Evolution 107, 142-151.

Uribe, J.E., Zardoya, R., Puillandre, N., 2018. Phylogenetic relationships of the conoidean snails (Gastropoda: Caenogastropoda) based on mitochondrial genomes. Molecular Phylogenetics and Evolution 127, 898-906.

Williams, S.T., Duda, T.F., Jr., 2008. Did tectonic activity stimulate oligo-miocene speciation in the Indo-West Pacific? Evolution 62, 1618-1634.

Woodward, S.R., Cruz, L.J., Olivera, B.M., Hillyard, D.R., 1990. Constant and hypervariable regions in conotoxin propeptides. The EMBO Journal 9, 105-1020.

WoRMS Editorial Board, 2019. World Register of Marine Species. http://www.marinespecies.org at VLIZ, VLIZ.

Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C., Jorger, K.M., Schrodl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proceedings of the Royal Society B: Biological Sciences 281, 20141739.

2. - Aims

2. – Aims

As explained in the introduction, cone snails (Gastropoda: Conidae) are a hyperdiverse group of marine gastropods, presenting a wide distribution, different larval developmental modes and two specializations that could be key in the success of the group: the venom gland and the harpoon-shaped radular tooth. In particular, West African cone snails represent a sound model system to study cone evolution as they evolved through radiations events, they include representatives of the particular adaptations of the family, and they present high levels of endemicity. However, we currently lack a well-supported phylogeny for both the family and the West African genera, hindering the study of the mechanisms leading to cone snail diversification and adaptation.

The main goal of this work was studying West African cone snail evolutionary and adaptation processes, trying to uncover the main factors that have led to and maintain the current diversity of species of this successful group. This goal could be further divided in two main objectives:

First (**phylogenomics**), **Chapters I**, **II** and **III** of results reconstruct phylogenies and set the evolutionary framework. The goal of **Chapter I** is reconstructing a robust phylogeny of the family Conidae, thus inferring the relative position and sister relationships of the West African genera. The available mitochondrial genome repertoire at the beginning of this work was scarce, so more mitogenomes were sequenced, thus increasing the generic representation. Transcriptomic data from venomics studies was also used to assembly the corresponding mt genes. **Chapters II** and **III** are specifically focused on West African species. The evolutionary history of the genus *Lautoconus* species endemic to Senegal, plus the closely related *Lautoconus guanche*, (**Chapter II**) and the endemic genera to Cabo Verde (**Chapter III**) were studied using mitochondrial genomes. All these trees were time calibrated and the evolution of different characters (diet, radula, etc.) were studied by mapping them onto the phylogeny. The mitochondrial genomes of the **Chapters II** and **III** were also used to test the reliability of the morphological characters used to describe species in this region.

Second (**venomics**), **Chapters IV** and **V** were focused on venom evolution. The main goal of **Chapter IV** was digging on the particular adaptation of cone snails to prey on fish. The species *Chelyconus ermineus*, the only Atlantic piscivorous species and representative of the second shift to piscivory is present in Cabo Verde. Its venom was sequenced, analysed using transcriptomics, and compared to the venom of other piscivorous cone species from the Indo-Pacific clade. The **Chapter V** is focused on the venoms of the two genera endemic to West Africa. The genera *Africonus* and *Lautoconus* are vermivorous, closely related phylogenetically and present an allopatric distribution (*Lautoconus* is present in the African coast and the Canary Islands, while *Africonus* is restricted to Cabo Verde). This chapter studies the variability of the venom composition along a phylogenetic gradient, from a biogeographic framework, and uses the data from **Chapter IV** to dig into the differences between piscivory and vermivory.

In summary, the specific goals of this work were:

1. To increase the mitogenomic catalogue of cone snails, maximizing the representation of different genera.
2. To develop a bioinformatics pipeline to annotate mitochondrial genes and ortholog nuclear genes from venom gland transcriptomes to be used in phylogenetic inference.
3. To collect the maximum number of species possible from West Africa, and amplify their mitogenomes.
4. To infer, using probabilistic methods, the evolutionary history of the family and the African genera.
5. To date major cladogenetic events in the different phylogenies.
6. To map character evolution onto these phylogenies.
7. To characterize the venom composition (both in terms of number of conotoxins and levels of expression) of the Atlantic piscivorous species *Chelyconus ermineus*.
8. To study the compartmentalization of the venom gland.
9. To determine the mechanisms leading to the shift to piscivory in cone snails.
10. To characterize the venom composition of 14 species of cone snails endemic to West Africa.
11. To compare the differences in venom composition along a phylogenetic gradient.
12. To analyse venom composition from a biogeographical perspective.

3. - Results

3.1. - Chapter I: “Conidae Phylogenomics and evolution”

3.1. - Capítulo I: “Filogenómica y evolución de Conidae”

Zoologica Scripta, 48 (2): 194 – 214

La comprensión del papel relativo de los diferentes procesos evolutivos que conducen a la extraordinaria diversidad morfológica, ecológica y de especies de los conos requiere una filogenia robusta, que hasta ahora ha sido difícil de obtener. Aquí, reunimos un conjunto de datos de genomas mitocondriales (mt), que incluye cuatro genomas mt recién secuenciados, 25 genomas mt disponibles públicamente, así como los genes codificantes de proteínas y de RNAs ribosomales ensamblados a partir de transcriptomas de glándulas venenosas. En total, analizamos 42 diferentes especies que representan 27 géneros de conos, es decir, alrededor de un tercio de la diversidad genérica del grupo. Además, utilizamos las lecturas de ARN-Seq para ensamblar 21 genes nucleares, que se concatenaron en un conjunto de datos nucleares. Finalmente, también contruímos una matriz combinada que incluía genes mitocondriales y nucleares. Las tres matrices de datos se analizaron con métodos probabilísticos, modelos homogéneos de sitio y heterogéneos de sitio y con los genes codificantes de proteínas tanto a nivel de aminoácidos como de nucleótidos. La especialización en dietas, la morfología radular y el tipo de protoconcha (paucispiral o multispiral, que indica larvas lecitotróficas o planctónicas, respectivamente), así como la diversidad de conotoxinas se mapearon en la filogenia mt reconstruida. También se dataron los principales eventos cladogenéticos dentro del grupo.

Received: 19 July 2018 | Revised: 19 October 2018 | Accepted: 23 October 2018

DOI: 10.1111/zsc.12329

ORIGINAL ARTICLE

WILEY Zoologica Scripta

Conidae phylogenomics and evolution

Samuel Abalde[1] | **Manuel J. Tenorio**[2] | **Juan E. Uribe**[1,3,4] | **Rafael Zardoya**[1]

[1]Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, Spain

[2]Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cádiz, Puerto Real, Spain

[3]Department of Invertebrate Zoology, Smithsonian Institution, National Museum of Natural History, Washington, District of Columbia, USA

[4]Grupo de Evolución, Sistemática y Ecología Molecular, Universidad del Magdalena, Santa Marta, Colombia

Correspondence
Rafael Zardoya, Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, Spain.
Email: rafaz@mncn.csic.es

Funding information
Spanish Ministry of Science, Industry, and Innovation, Grant/Award Number: CGL2016-75255-C2-1-P (AEI/FEDER, UE) , BES-2014-069575 and BES-2011-051469

Abstract

Understanding the relative role of different evolutionary processes leading to the extraordinary morphological, ecological and species diversity of cone snails requires a robust phylogeny, which thus far has been elusive. Here, we constructed a mitochondrial (mt) genome data set, which included four newly sequenced mt genomes, 25 publicly available mt genomes and 24 data sets with all mt protein-coding and rRNA genes assembled from venom gland transcriptomes. In total, we analysed 42 different species representing 27 genera of cone snails, that is, about one third of the generic diversity of the group. In addition, we used the RNA-Seq reads to assemble 21 nuclear genes, which were concatenated in a nuclear data set. Finally, a combined data set including mt and nuclear genes was also constructed. The three data matrices were analysed with probabilistic methods, site-homogeneous and site-heterogeneous models, and with protein-coding genes both at the amino acid and nucleotide levels. Diet specialization, radular morphology and the type of protoconch (paucispiral or multispiral indicating lecithotrophic or planktonic larvae, respectively) as well as conotoxin diversity were mapped onto the reconstructed mt phylogeny, and a chronogram dating mayor cladogenetic events within the group was also reconstructed.

KEYWORDS

Conidae, conotoxins, mitogenome, phylogenomics, transcriptomics

1 | INTRODUCTION

In few decades of existence, the field of molecular phylogenetics has come of age and undoubtedly represents a quantum leap for inferring the evolutionary history of organisms. For many years, one of the main limitations of molecular phylogenetics derived from the use of single or few genes, which could often render unresolved trees (Bleidorn et al., 2009; Rokas, Williams, King, & Carroll, 2003). This problem has been particularly severe in the case of evolutionary radiations (i.e., major diversification events occurring in a relatively short period of time) since phylogenetic relationships among taxa are difficult to disentangle due to the limited associated phylogenetic signal, which is visualized as short internal nodes in the tree (Philippe et al., 2011). Upon the advent of the next-generation sequencing (NGS) techniques, the capacity for generating genomewide markers for virtually any taxon was thought to overcome this limitation and open the possibility of reconstructing fully resolved trees (Rokas et al., 2003). Although NGS results are encouraging, it has become apparent that just adding many genes is not enough, as the taxon sampling (including outgroup selection), fit of the evolutionary models and marker congruence, among others, also have capital effects on the results (Hedtke, Townsend, & Hillis, 2006; Jeffroy, Brinkmann, Delsuc, & Philippe, 2006; Philippe et al., 2011). A new field, phylogenomics, has emerged to uncover drawbacks associated with large sequence data sets and develop new analytical methods, which are effectively helping in the resolution of most recalcitrant nodes in the "Trees of Life" (Vargas & Zardoya, 2014) of diverse groups such as molluscs (Smith et al., 2011), annelids (Weigert et al., 2014), vertebrates (Irisarri et al., 2017), arthropods (Espeland et

al., 2018; Sharma et al.., 2018) or land plants (Wickett et al., 2014), among many others rapidly accumulating.

Cone snails constitute a paradigm of a natural group, which has diversified and evolved through radiation events. Their rich fossil record shows that successive radiations along cone evolutionary history have contributed to their current great species diversity (Kohn, 1990), including most recent diversifications in the Cabo Verde archipelago (Cunha, Castilho, Rüber, & Zardoya, 2005; Duda & Rolán, 2005), the Senegalese coast (Pin & Leung Tack, 1995), Madagascar (Monnier, Tenorio, Bouchet, & Puillandre, 2018) or the Caribbean Sea (Kohn, 2014). Altogether, old and recent radiations have led to astonishing species diversity with a worldwide distribution. Currently, there are more than 900 species of cone snails described and this number grows steadily every year (MolluscaBase, 2018), being present in a broad range of depths in all tropical and subtropical seas (Tucker & Tenorio, 2013).

Cones are a key ecological component in intertidal and subtidal habitats, where most species feed mainly on worms, but also some on molluscs and others on fishes (Kohn, 1959). Cone predatory capacity relies on a sophisticated venom system, formed by hollow harpoon-like radular teeth, which inject a cocktail of hundreds of different peptides named conotoxins (Li et al., 2017; Norton & Olivera, 2006; Peng et al., 2016). The great specificity and biological potential of conotoxins have attracted the interest of pharmacological research (Miljanich, 2004; Yang et al., 2017), and venom gland transcriptomics are currently the main tool for cataloguing the cocktail composition in the different species (Abalde, Tenorio, Afonso, & Zardoya, 2018; Barghi, Concepcion, Olivera, & Lluisma, 2015; Dutertre et al., 2014; Hu, Bandyopadhyay, Olivera, & Yandell, 2012; Li et al., 2017; Peng et al., 2016). In this regard, understanding the evolutionary processes involved in conotoxin diversification and adaptation to different preys requires a robust phylogeny of cones.

The extraordinary species diversity of cones has made their systematics particularly challenging. For many years, the consensus was to classify cone snails into the single genus *Conus*, but recent morphological (Tucker & Tenorio, 2009) and molecular (Duda & Kohn, 2005; Puillandre et al., 2014; Uribe, Puillandre, & Zardoya, 2017) phylogenetic studies discovered enough divergence among main lineages inside the group to ultimately propose the splitting of genus *Conus* into several genera. A first molecular phylogenetic study (Puillandre et al., 2014) recognized four main genera: *Profundiconus*, *Californiconus*, *Conasprella* and *Conus*; the latter holding most of the species diversity with up to 60 monophyletic groups classified as subgenera. A subsequent molecular phylogenetic study (Uribe, Puillandre, et al., 2017) added two more genera: *Lilliconus* and *Pseudolilliconus* (later redefined *Pygmaeconus*; Puillandre & Tenorio, 2017). Alternatively, an exhaustive review of morphological characters classified cones into 84 extant genera (Tucker & Tenorio, 2009), 63 of them included within the family Conidae (equivalent to the genus *Conus* in Puillandre et al., 2014). These numbers were raised in Tucker and Tenorio (2013) to 114 extant genera, and 90 within the family Conidae. Therefore, both the morphological and the molecular proposals agreed on the need of going beyond the single genus classification, were generally congruent in the definition of the groups and only differed in the taxonomic rank that should be used. Furthermore, these groups are monophyletic, share shell and radula synapomorphies and often correspond to biogeographical assemblies, suggesting that the use of the corresponding generic names may be more suitable and convenient. Therefore, we will follow here the classification of Tucker and Tenorio (2009) with the updates of Tucker and Tenorio (2013) and focus on the phylogenetic relationships among genera within Conidae sensu these authors.

Therefore, the challenge is going beyond phylogenies of cones based on few genes (Aman et al., 2015; Puillandre et al., 2014) and reconstruct robust phylogenetic relationships among cone genera sensu Tucker and Tenorio (2009) based on multilocus sequences data sets such as the recent one based on concatenating hundreds to thousands of exon sequences (Phuong & Mahardika, 2018). An additional source of molecular markers for phylogenetic inference is mitochondrial (mt) genomes, which have been widely used in gastropods and proven to be particularly useful for resolving phylogenies at the family level (Osca, Templado, & Zardoya, 2015; Uribe, Williams, Templado, Abalde, & Zardoya, 2017) as well as for disentangling recent radiation events (Abalde, Tenorio, Afonso, & Zardoya, 2017; Abalde, Tenorio, Afonso, Uribe, et al., 2017). Although there are currently about 150 mt genomes of the family Conidae available in GenBank at NCBI (https://www.ncbi.nlm.nih.gov/), the total diversity of the group is clearly underrepresented as these mitogenomes belong to only 11 out of the currently 89 described genera (the genus *Trovaoconus* of Tucker and Tenorio (2009) was recently synonymized with *Kalloconus* by Abalde, Tenorio, Afonso, Uribe, et al., 2017). Alternatively, the many recent Illumina RNA-Seq studies of cone venom glands have deposited millions of raw reads on the Sequence Read Archive (SRA) at NCBI, which belong to 19 different genera, and could be used for cone phylogenomics. Here, we used all these publicly available sequence data and four newly sequenced mt genomes aiming to (a) reconstruct a robust phylogeny for the family Conidae, based on mitochondrial and nuclear data; (b) infer the evolution of diet specialization, radular morphology and conotoxin diversity; and (c) date major cladogenetic events within the family.

TABLE 1 Mitochondrial genomes analysed in this study

Species	Diet	Data source	Transcriptome		mt genome	
			SRA No.	Reference	GenBank Acc. No.	Reference
Africonus borgesi (Trovão, 1979)	Vermivorous	GenBank	—	—	NC_013243	Cunha, Grande, and Zardoya (2009)
Africonus infinitus (Rolán, 1990)	Vermivorous	GenBank	—	—	KY864967	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Africonus miruchae (Röckel, Rolán & Monteiro, 1980)	Vermivorous	GenBank	—	—	KY864971	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Calamiconus quercinus (Lightfoot, 1791)	Vermivorous	GenBank	—	—	KY609509	Gao, Peng, Chen, Zhang, and Shi (2018)
Calamiconus quercinus (Lightfoot, 1791)	Vermivorous	SRA	SRR2609537	Phuong, Mahardika, and Alfaro (2016)	—	This study
Chelyconus ermineus (Born, 1778)	Piscivorous	GenBank	—	—	KY864977	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Chelyconus ermineus (Born, 1778)	Piscivorous	SRA	SRR6983166, 68, 69	Abalde et al. (2018)	—	This study
Conus marmoreus Linnaeus, 1758	Molluscivorous	SRA	SRR2609532	Phuong et al. (2016)	—	This study
Cylinder gloriamaris (Chemnitz, 1777)	Molluscivorous	GenBank	—	—	NC_030213	Chen, Hsiao, Huang, et al. (2016) unpublished
Cylinder gloriamaris (Chemnitz, 1777)	Molluscivorous	SRA	SRR5499408	Robinson et al. (2017)	—	This study
Cylinder textile (Linnaeus, 1758)	Molluscivorous	GenBank	—	—	NC_008797	Bandyopadhyay et al. (2008)
Cylinder victoriae (Reeve, 1843)	Molluscivorous	SRA	SRR833564	Robinson *et al*. (2014)	—	This study
Darioconus episcopatus (da Motta, 1982)	Molluscivorous	SRA	DRR034332	Lavergne *et al*. (2015)	—	This study
Dendroconus betulinus (Linnaeus, 1758)	Vermivorous	SRA	SRR2124881	Peng et al. (2016)	—	This study
Eugeniconus nobilis (Linnaeus, 1758)	Molluscivorous	GenBank	—	—	KX263253	Uribe, Puillandre, et al. (2017)
Fulgiconus goudeyi (Monnier & Limpalaër, 2012)	Vermivorous	PCR	—	—	KY864975	This study
Gastridium geographus (Linnaeus, 1758)	Piscivorous	SRA	SRR503413-16	Hu et al. (2012)	—	This study
Gastridium tulipa (Linnaeus, 1758)	Piscivorous	GenBank	—	—	NC_027518	Chen, Hsiao, Huang, et al. (2016)
Genuanoconus genuanus (Linnaeus, 1758)	Vermivorous (Amphinomidae)	PCR	—	—	KY864974	This study

(Continues)

TABLE 1 (Continued)

Species	Diet	Data source	Transcriptome		mt genome	
			SRA No.	Reference	GenBank Acc. No.	Reference
Harmoniconus sponsalis (Hwass in Bruguière, 1792)	Vermivorous	SRA	SRR2609541	Phuong *et al*. (2016)	—	This study
Kalloconus ateralbus (Kiener, 1850)	Vermivorous	GenBank	—	—	KY864970	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Kalloconus pulcher ([Lightfood]), 1786)	Vermivorous	GenBank	—	—	KY864972	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Kalloconus pulcher ([Lightfood]), 1786)	Vermivorous	GenBank	—	—	KY864973	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Kalloconus trochulus (Reeve, 1844)	Vermivorous	GenBank	—	—	KY864969	Abalde, Tenorio, Afonso, Uribe, et al. (2017)
Kalloconus venulatus (Hwass in Bruguière, 1792)	Vermivorous	GenBank	—	—	KX263250	Uribe, Puillandre, et al. (2017)
Kioconus lenavati (da Motta & Röckel, 1982)	Vermivorous	SRA	SRR1803942	Barghi et al. (2015)	—	This study
Kioconus tribblei (Walls, 1977)	Vermivorous	GenBank	—	—	NC_027957	Barghi, Concepcion, Olivera, and Lluisma (2016)
Kioconus tribblei (Walls, 1977)	Vermivorous	SRA	SRR1802610	Barghi et al. (2015)	—	This study
Kioconus tribblei (Walls, 1977)	Vermivorous	SRA	SRR1803938	Barghi et al. (2015)	—	This study
Lautoconus guanche (Lauer, 1993)	Vermivorous	GenBank	—	—	KY801847	Abalde, Tenorio, Afonso, and Zardoya (2017)
Lautoconus hybridus (Kiener, 1847)	Vermivorous	GenBank	—	—	KX263252	Uribe, Puillandre, et al. (2017)
Lautoconus ventricosus (Gmelin, 1791)	Vermivorous	GenBank	—	—	KX263251	Uribe, Puillandre, et al. (2017)
Lindaconus spurius (Gmelin, 1791)	Vermivorous	PCR	—	—	KY864976	This study
Lividoconus lividus (Hwass in Bruguière, 1792)	Vermivorous	SRA	SRR2609539	Phuong et al. (2016)	—	This study
Miliariconus coronatus (Gmelin, 1791)	Vermivorous	SRA	SRR2609545	Phuong et al. (2016)	—	This study
Miliariconus miliaris (Hwass in Bruguière, 1792)	Vermivorous	SRA	SRR1548185	Weese and Duda (2015)	—	This study
Monteirococonus tabidus (Reeve, 1844)	Vermivorous	PCR	—	—	KY864968	This study
Pionoconus consors (G. B. Sowerby I, 1833)	Piscivorous	GenBank	—	—	NC_023460	Brauer et al. (2012)
Pionoconus consors (G. B. Sowerby I, 1833)	Piscivorous	SRA	SRR1955039	Leonardi et al. (2012)	—	This study
Pionoconus striatus (Linnaeus, 1758)	Vermivorous	GenBank	—	—	NC_030536	Chen et al. (2016b)

(Continues)

TABLE 1 (Continued)

Species	Diet	Data source	Transcriptome		mt genome	
			SRA No.	Reference	GenBank Acc. No.	Reference
Puncticulis arenatus (Hwass in Bruguière, 1792)	Vermivorous	SRA	SRR2609544	Phuong et al. (2016)	—	This study
Rhizoconus capitaneus (Linnaeus, 1758)	Vermivorous	GenBank	—	—	NC_030354	Chen et al. (2016a)
Rhizoconus vexillum (Gmelin, 1791)	Vermivorous	SRA	SRR2890189	Prashanth et al. (2016)	—	This study
Rhombiconus imperialis (Linnaeus, 1758)	Vermivorous	SRA	SRR2609542	Phuong et al. (2016)	—	This study
Rolaniconus varius (Linnaeus, 1758)	Vermivorous	SRA	SRR2609543	Phuong et al. (2016)	—	This study
Splinoconus biliosus (Röding, 1798)	Vermivorous	SRA	SRR1956759	Unpublished	—	This study
Virgiconus virgo (Linnaeus, 1758)	Vermivorous	SRA	SRR2608262	Phuong et al. (2016)	—	This study
Virroconus ebraeus (Linnaeus, 1758)	Vermivorous	SRA	SRR2609538	Phuong et al. (2016)	—	This study
Conasprella wakayamaensis (Kuroda, 1956)	Vermivorous	GenBank	—	—	KX263254	Uribe, Puillandre, et al. (2017)
Californiconus californicus (Reeve, 1844)	All	GenBank	—	—	KX263249	Uribe, Puillandre, et al. (2017)
Californiconus californicus (Reeve, 1844)	All	SRA	SRR2609536	Phuong et al. (2016)	—	This study
Profundiconus teramachii (Kuroda, 1956)	Vermivorous	GenBank	—	—	KX263256	Uribe, Puillandre, et al. (2017)
Tomopleura sp.	—	GenBank	—	—	KX263259	Uribe, Puillandre, et al. (2017)

Note. Genbank: published mtDNA; PCR: amplified mtDNA; SRA: RNA sequences.

TABLE 2 Nuclear genes analysed in this study

Gene	Length (bp)	Length (aa)
Translocon-associated protein subunit alpha	921	307
Cathepsin Z	942	314
Eukaryotic translation elongation factor 1 beta 2-like	660	220
Transmembrane protein 59-like	1,023	341
RWD domain-containing protein 1-like	750	250
Eukaryotic translation initiation factor 3 subunit G-like	825	275
Alcohol dehydrogenase class-3-like protein	1,143	381
Eukaryotic translation initiation factor 3 subunit F-like	879	293
Sodium-/potassium-transporting ATPase subunit beta-like	894	298
Ferritin	522	174
Syntenin-1	909	303
Bax inhibitor-1	600	200
CD63 antigen	720	240
Translocation protein SEC62-like	1,185	395
Cyclin-I	990	330
Nucleoside diphosphate kinase B	507	169
Mitochondrial import receptor subunit TOM20	459	153
Ragulator complex protein LAMTOR3-A	378	126
Coatomer subunit epsilon-like	903	301
Nascent polypeptide-associated complex subunit alpha-like protein	630	210
Elongation factor 1-gamma	1,284	428
Total	17,124	5,708

2 | MATERIALS AND METHODS

2.1 | Taxon sampling

We studied 48 specimens belonging to 41 different species and 27 genera within the family Conidae (Table 1). As outgroup, we selected five specimens of another four highly divergent genera (*Californiconus*, *Conasprella* and *Profundiconus* of cone snails and *Tomopleura* of the family Borsoniidae). During November 2017, a total of 25 mt genome and 24 RNA-Seq entries were downloaded from GenBank and SRA databases at NCBI, respectively (Table 1). We also sequenced the nearly complete (only missing the control region and flanking regions) mt genomes of four species: *Fulgiconus goudeyi* (voucher MNCN/ADN 95093; from New Caledonia), *Genuanoconus genuanus* (voucher MNCN/ADN 95096; from Cabo Verde), *Lindaconus spurius* (voucher MNCN/ADN 95097; from Aruba) and *Monteiroconus tabidus* (voucher MNCN/ADN 95098; from Cabo Verde).

2.2 | Mitochondrial DNA extraction, PCR amplification, sequencing and annotation

Total genomic DNA was isolated from up to 1–3 mg of foot tissue of one individual of *F. goudeyi*, *G. genuanus, L. spurius* and *M. tabidus* following a standard phenol-chloroform extraction. The mtDNA was PCR-amplified in two or three overlapping fragments following Uribe, Puillandre, et al. (2017). Long PCR products from the same mitogenome were pooled together in equimolar concentrations, and an indexed library was constructed using the NEXTERA XT DNA library prep kit (Illumina, San Diego, CA, USA). The four libraries were sequenced together with others from different projects in a single run of an Illumina MiSeq platform (2 × 150 paired-end reads) at Sistemas Genómicos (Valencia, Spain).

Raw reads were uploaded to the TRUFA webserver (Kornobis et al., 2015), where the 150 bp reads were trimmed and filtered out if the PHRED quality was below 20 using PRINSEQ v.0.20.3 (Schmieder & Edwards, 2011) and de novo assembled using Trinity r2012-06-08 (Grabherr et al., 2011) with default parameters. The resulting contigs with a minimum length of 5 kb were selected and used as reference to map all remaining reads using Geneious® 8.1.8. The mapping requirements were a minimum overlap of 60 bp with a 100% of identity and one mismatch allowed.

Newly sequenced mitogenomes were annotated using already published mtDNAs of other cone snails as a reference with Geneious. The open reading frames (ORFs) of the 13 mt protein-coding genes were manually checked for potential misannotations in the start and stop codons. All mt transfer RNA (tRNA) genes were identified using tRNAscan-SE 1.21 (Lowe & Chan, 2016), which infers cloverleaf secondary structures (with a few exceptions that were determined manually). The mt ribosomal RNA (rRNA) genes were identified by comparison with orthologous genes in other Conidae mt genomes and assumed to extend to the boundaries of adjacent genes (Boore, Macey, & Medina, 2005).

2.3 | Transcriptome assembly and annotation

The raw reads from the different cone venom gland transcriptome projects were downloaded from the SRA repository at NCBI and trimmed, filtered and de novo assembled in the TRUFA webserver, using default parameters in PRINSEQ v.0.20.3 (Schmieder & Edwards, 2011) and Trinity r2012-06-08 (Grabherr et al., 2011). All contigs with more than 200 bp were kept for annotation.

The different mt genes were identified by BLASTn searches against a custom database, formed by the already published mt genomes of cone snails. The sequences of the assembled mt genes are presented in Supporting Information Table S1. In addition, nuclear ORFs were identified using TransDecoder (Haas, 2016). Those genes which were common to all studied species were selected using the AGALMA pipeline (Dunn, Howison, & Zapata, 2013). Briefly, all homologous sequences were identified through an all-by-all blast approach. The resulting gene clusters were aligned and cleaned, and a phylogenetic tree for each gene cluster was built using RAxML v.8.1.16 (Stamatakis, 2014). Orthologs were identified by looking for repeated species trees in the gene tree (indicative of gene duplications). Only those orthologs present in at least 21 of the 24 analysed individuals were kept. For each of the selected orthologs, sequences were aligned using TranslatorX (Abascal, Zardoya, & Telford, 2010) and manually curated and filtered, as suggested by Philippe et al. (2011) in order to detect translational frameshifts and local sequencing errors as well as unusually divergent sequences indicating potential contaminations. Finally, the 21 most complete and divergent nuclear gene sequences were chosen (Table 2) in order to obtain similar amounts of nuclear and mt sequence data in the final combined data set.

As a proof of concept, in those cases for which transcriptomic and mitogenomic sequence data were available independently for the same species (*Calamiconus quercinus*, *Californiconus californicus*, *Chelyconus ermineus*, *Cylinder gloriamaris*, *Kioconus tribblei* and *Pionoconus consors*), the number of different positions and percentages of similarity were examined.

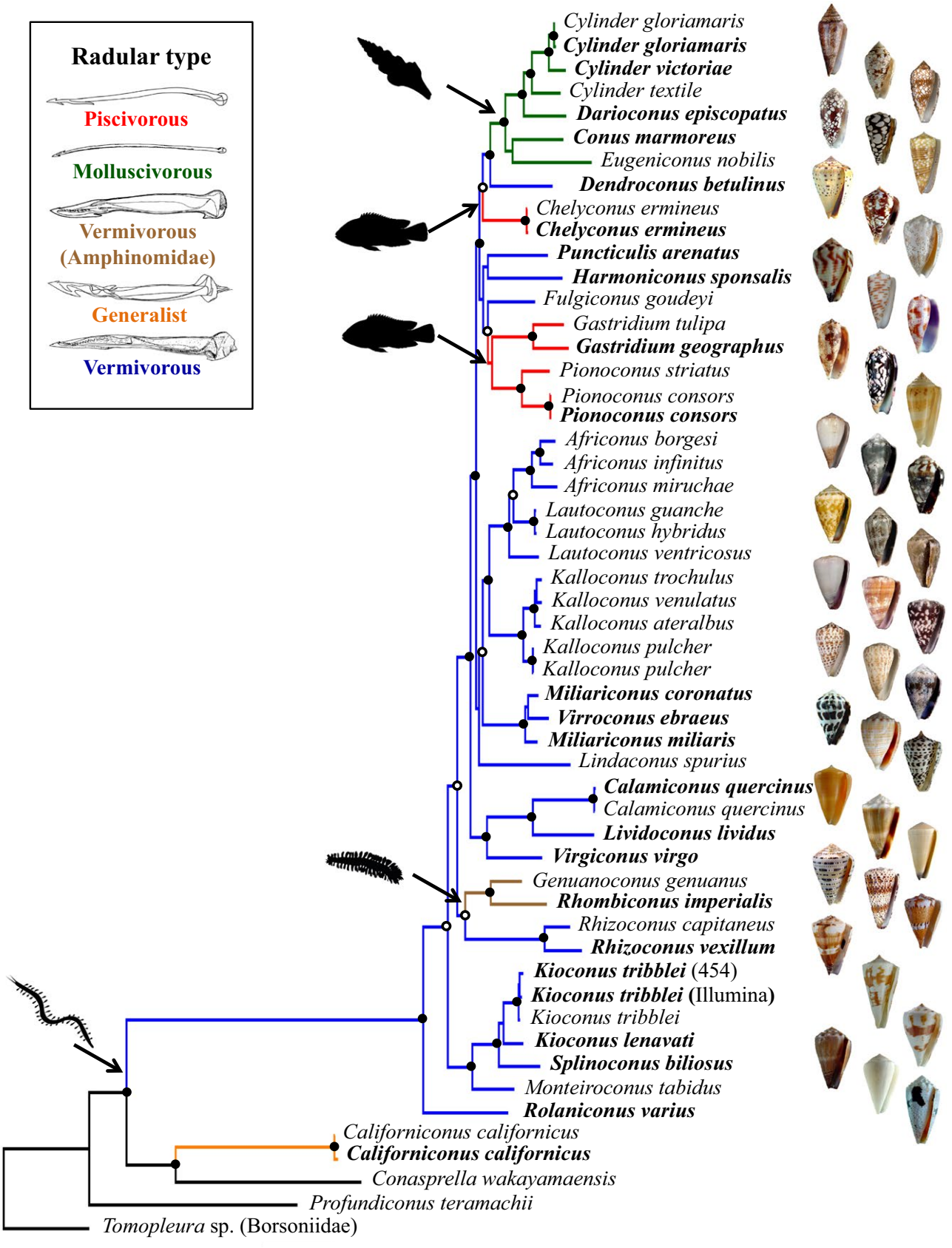
2.4 | Sequencing alignment and phylogenetic analyses

Three data sets were compiled: (a) mt genome, which included 13 mt protein-coding and two rRNA genes; (b) nuclear, which included 21 nuclear genes; and (c) combined, which included mt and nuclear genes. Protein-coding genes were analysed both at the nucleotide and the amino acid levels. They were individually aligned using TranslatorX (Abascal et al., 2010), whereas the nucleotide sequences of the rRNA genes were aligned using MAFFT v7 (Katoh & Standley, 2013). All ambiguously aligned positions were removed using GBlocks v.0.9.1b (Castresana, 2000) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. Sequences were format converted for further analyses using the ALTER webserver (Glez-Pena, Gomez-Blanco, Reboiro-Jato, Fdez-Riverola, & Posada, 2010). Finally, the different single alignments were concatenated using Geneious. Alignments can be accessed at TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S22475).

The best-fit partition schemes and models of substitution for each data set were identified using PartitionFinder (Lanfear, Calcott, Ho, & Guindon, 2012) with the Bayesian information criterion (BIC; Schwarz, 1978). The following partitions were tested: all genes together, all mt genes together versus all nuclear genes together, all genes arranged in subunits (*atp*, *cob*, *cox*, *nad, rrn* and *nuclear*) and all genes separated (except *atp6-atp8* and *nad4-nad4L)*. In addition, for those data sets in which protein-coding genes were analysed at the nucleotide level, we also tested separately the three codon positions.

Phylogenetic relationships were inferred using maximum likelihood (ML; Felsenstein, 1981) and Bayesian inference (BI; Huelsenbeck & Ronquist, 2001). For ML, we used RAxML v.8.1.16 (Stamatakis, 2014) with the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates. BI analyses were conducted using (a) MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003), performing two independent runs (to increase the chance of adequate mixing of the Markov chains and of convergence) with four simultaneous Markov chains for 10 million of generations, sampling every 1,000 generations, and discarding the first 25% generations as burn-in (as judged by plots of ML scores and low *SD* of split frequencies) to prevent sampling before reaching stationarity; and (b) PhyloBayes MPI v1.5 (Lartillot, Rodrigue, Stubbs, & Richer, 2013), running two independent chains under a site-heterogeneous CAT-GTR model and based only on

FIGURE 1 Phylogeny of the family Conidae. The reconstructed BI tree using best-fit partitions and site-homogeneous models based on the mt genome data set (concatenated 13 protein-coding genes analysed as amino acid sequences plus two rRNA genes at nucleotide level) is shown. Bayesian posterior probabilities (BPP) supporting nodes are shown as black (BPP = 1) and white (BPP = 0.95–0.99) dots. Bold names indicate those species whose genes were assembled from transcriptomic data. Branch colours and silhouettes (downloaded from PhyloPic) represent ancestral character state reconstruction under unordered parsimony of diet specializations: blue, vermivory; brown, vermivory on amphinomids; green, molluscivory; red, piscivory. In the inset, radular teeth from (Tucker & Tenorio, 2009) corresponding to the different feeding modes are illustrated. Shell pictures are from the authors or from Alexander Medvedev (www.coneshells-am.ruwileyonlinelibrary.com]). Scale bar indicates substitutions/site. [Colour figure can be viewed at wileyonlinelibrary.com]


Radular type
Piscivorous
Molluscivorous
Vermivorous (Amphinomidae)
Generalist
Vermivorous
Cylinder gloriamaris
Cylinder gloriamaris
Cylinder victoriae
Cylinder textile
Darioconus episcopatus
Conus marmoreus
Eugeniconus nobilis
Dendroconus betulinus
Chelyconus ermineus
Chelyconus ermineus
Puncticulis arenatus
Harmoniconus sponsalis
Fulgiconus goudeyi
Gastridium tulipa
Gastridium geographus
Pionoconus striatus
Pionoconus consors
Pionoconus consors
Africonus borgesi
Africonus infinitus
Africonus miruchae
Lautoconus guanche
Lautoconus hybridus
Lautoconus ventricosus
Kalloconus trochulus
Kalloconus venulatus
Kalloconus ateralbus
Kalloconus pulcher
Kalloconus pulcher
Miliariconus coronatus
Virroconus ebraeus
Miliariconus miliaris
Lindaconus spurius
Calamiconus quercinus
Calamiconus quercinus
Lividoconus lividus
Virgiconus virgo
Genuanoconus genuanus
Rhombiconus imperialis
Rhizoconus capitaneus
Rhizoconus vexillum
Kioconus tribblei (454)
Kioconus tribblei (Illumina)
Kioconus tribblei
Kioconus lenavati
Splinoconus biliosus
Monteiroconus tabidus
Rolaniconus varius
Californiconus californicus
Californiconus californicus
Conasprella wakayamaensis
Profundiconus teramachii
Tomopleura sp. (Borsoniidae)
0.09

the protein-coding genes at the amino acid or nucleotide levels. Convergence between chains was assessed a posteriori using the *bpcomp* and *tracecomp* tools implemented in PhyloBayes.

The outgroups used with the mt genome data set were *Conasprella wakayamaensis*, two specimens of *C. californicus*, *Profundiconus teramachii* and *Tomopleura sp*. (Family Borsoniidae). The outgroup used with the nuclear and combined data sets was one specimen of *C. californicus*.

Ancestral character state reconstructions of diet specialization, radular morphology and the type of protoconch (paucispiral or multispiral indicating lecitotrophic or planktonic larvae, respectively) as described in Tucker and Tenorio (2009) were performed using unordered maximum parsimony with Mesquite (Maddison & Maddison, 2018) and mapped onto our best working hypothesis for the phylogeny of Conidae (see Results), that is, the one recovered by the BI analysis based on the mt genome data set with protein-coding genes analysed at the amino acid level.

2.5 | Estimation of divergence times

Divergence times were estimated following a Bayesian approach using the software BEAST v.1.7.5 (Drummond & Rambaut, 2007). An uncorrelated relaxed molecular clock was used to infer branch lengths and nodal ages. The tree topology was fixed using our best working hypothesis for the phylogeny of Conidae but with only one tip per species and fixing only those nodes with high statistical support (BPP = 0.95–1) . For the clock model, the lognormal relaxed-clock model was selected, which allows rates to vary among branches without any a priori assumption of autocorrelation between adjacent branches. For the tree prior, a Yule process of speciation was employed. Concatenated mt protein-coding genes (amino acids) plus rRNA (nucleotides) genes were analysed. We used the mtREV (+I+G) substitution model for amino acids (the closest model available in the programme to the one selected by PartitionFinder) and the GTR (+I+G) substitution model for the rRNA genes (see Results). The final Markov chain was run twice for 26 million generations,

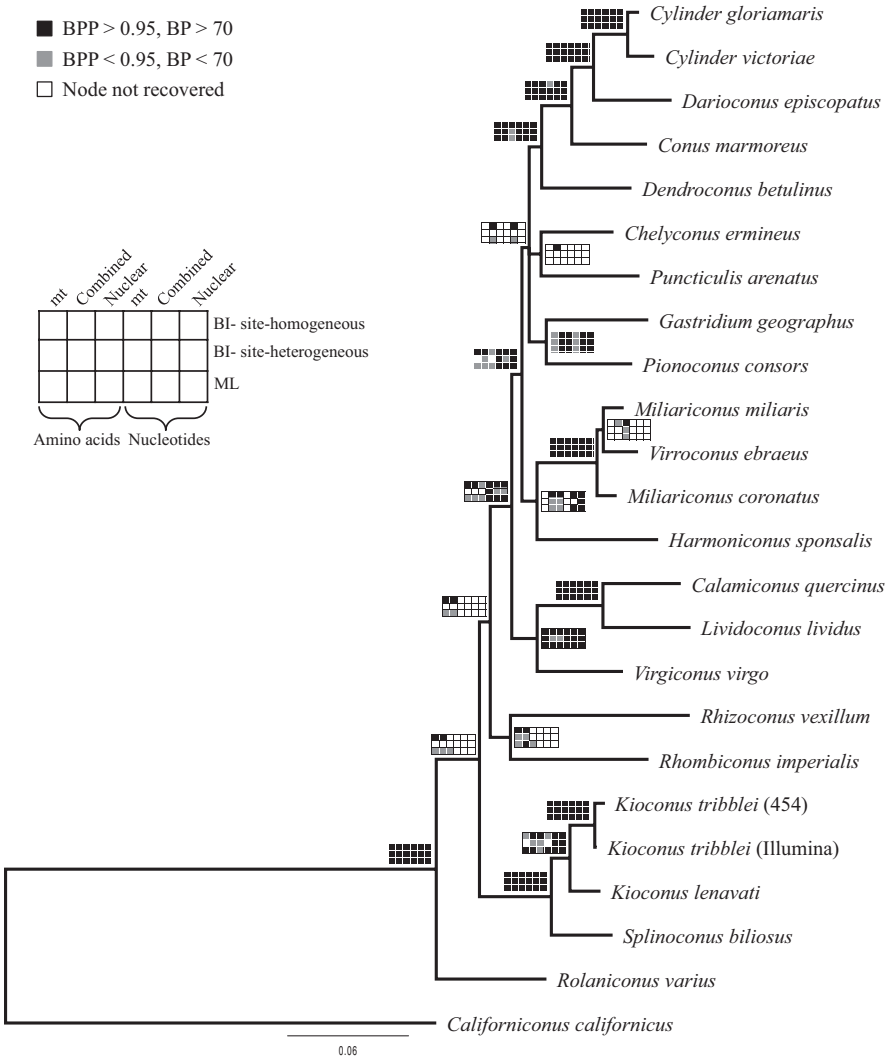


FIGURE 2 The reconstructed BI tree using best-fit partitions and site-homogeneous models based on the combined (mt + nuclear) data set with protein-coding genes analysed at the amino acid level. For each node, statistical support based on the 18 different phylogenetic analyses is shown: A black square represents BPP > 0.95 or BP > 70% in BI and ML, respectively; a gray square represents BPP between 0.90 and 0.95 or BP between 50% and 70%; a white square indicates no recovery of the node. Scale bar indicates substitutions/site. The genus *Californiconus* was used as outgroup

sampling every 10,000 generations, and the first 1,000 trees were discarded as part of the burn-in process, according to the convergence of chains checked with Tracer v.1.6. (Rambaut & Drummond, 2007). The ESS of all parameters was above 200 except for seven of 42 tmrca (time to most recent common ancestor) statistics, which were above 150.

Although there are many reported fossils of cone snails, their identification is not always straightforward, since shell morphology is prone to homoplasy (Abalde, Tenorio, Afonso, Uribe, et al., 2017; Duda, Bolin, Meyer, & Kohn, 2008). Hence, the use of cone fossils for calibration has to be done with caution. We tried two different approaches to calibrate the molecular clock: (a) The first known fossil of a cone snail was used to date the divergence between *Tomopleura sp*. (Borsoniidae) and Conidae (57 million years ago –mya–; Tracey, Craig, Belliard, & Gain, 2017) and the age of formation of Sal, the oldest island of Cabo Verde (28 mya; Holm et al., 2008), was used to date the divergence between *Africonus* and (paraphyletic) *Lautoconus*, as the colonization and diversification of *Africonus* species endemic to the different Cabo Verde islands was reported to occur shortly after island emergence (Abalde, Tenorio, Afonso, Uribe, et al., 2017); (b) the two previous references and three well-recognized fossils, which belong to *Conasprella* (Squire, 1987), *Kioconus* (Beu & Maxwell, 1990) and *Cylinder* (Shuto, 1969) lineages that would allow us testing their reliability as calibration points.

The calibration points were included in the analysis as follows. For the date of origin of Conidae (at least, 57 mya), we used a log-normal distribution, enforcing the mean to 58 ($SD = 0.05$, offset = 0.0001). The origin of Sal island (28 mya) was defined by a log-normal distribution, with mean in 24.5 ($SD = 0.05$, offset = 0.7). The three fossils were calibrated using normal distributions, whose means were 44.25 ($SD = 5.4$) for *Conasprella*, 19.5 ($SD = 1.8$) for *Kioconus* and 9.5 ($SD = 3.3$) for *Cylinder*.

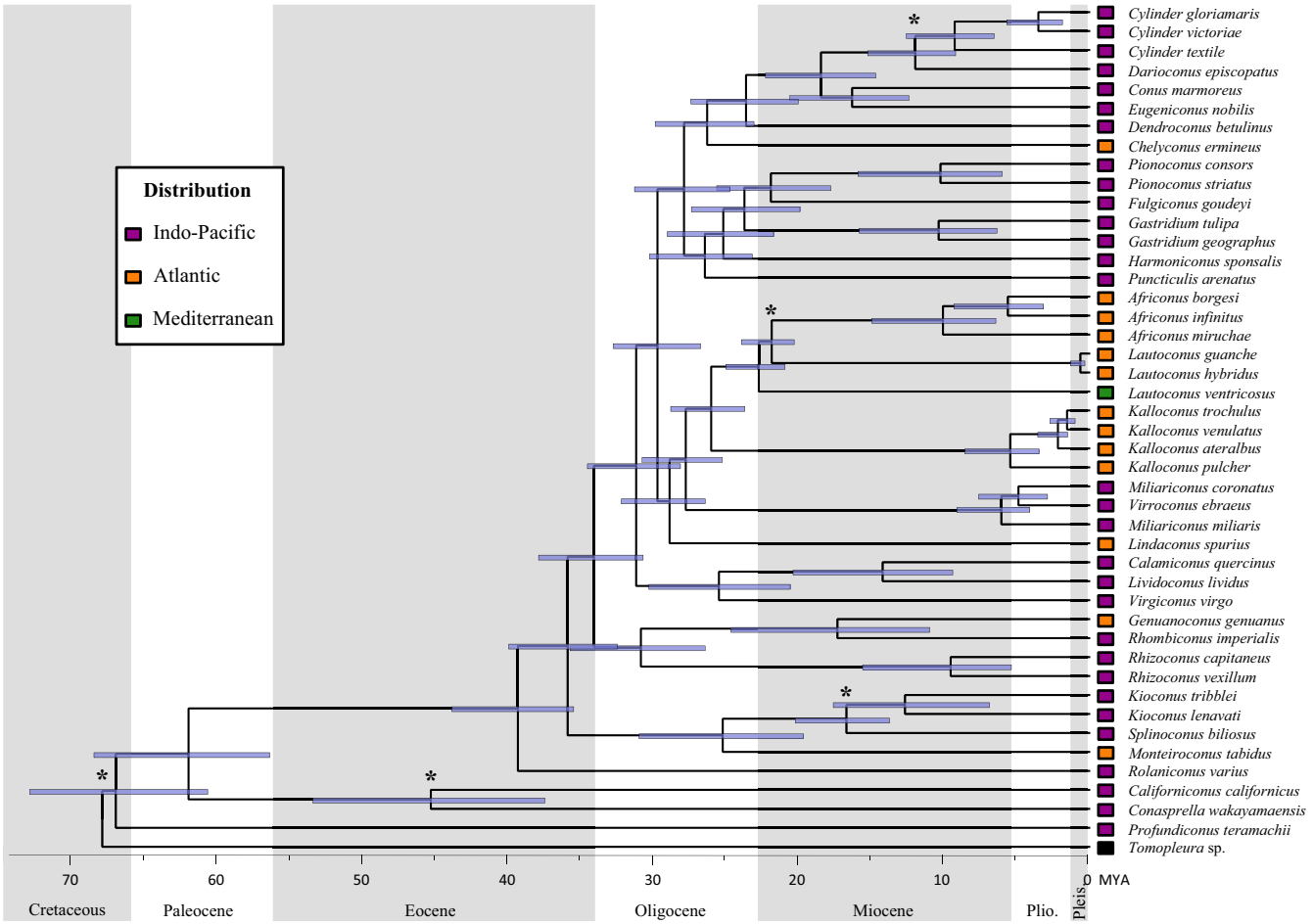


FIGURE 3 Chronogram based on the mitogenome data set (concatenated 13 protein-coding genes analysed as amino acid sequences plus two rRNA genes at nucleotide level) using the topology shown in Figure 1 (only nodes with high statistical support [BPP = 0.95–1] were fixed). A Bayesian uncorrelated relaxed lognormal clock with geographic- and fossil-based calibration priors (denoted by asterisks) was used in BEAST. Horizontal bars represent 95% credible intervals for time estimates; dates are in millions of years. Geological ages are highlighted as gray-white intervals. Square colours indicate cone distributions: purple, Indo-Pacific Ocean; orange, Atlantic Ocean; and green, Mediterranean Sea [Colour figure can be viewed at wileyonlinelibrary.com]

3 | RESULTS

3.1 | Sequencing, assembly and annotation

The nearly complete mt genomes of *F. goudeyi* (length = 15,261 bp; mean coverage = 5,321×; number of reads = 538,094), *G. genuanus* (15,328 bp; 7,301×; 741,250 reads), *L. spurius* (15,329 bp; 2,003×; 203,397 reads) and *M. tabidus* (15,368; 2,821×; 287,156 reads) were sequenced only lacking a fragment including the *trnF*, the control region and the beginning of the *cox3* gene, which was not PCR-amplified. All these mt genomes encode for 13 protein-coding, two rRNA and 22 tRNA genes (but note that the presence of the *trnF* gene could not be determined) and share the same gene order. All genes were encoded by the major strand, except those forming the cluster MYCWQGE (*trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, *trnE*) and the *trnT* gene. The complete annotations of the newly determined mt genomes including start and end of each gene, start and stop codons of protein-coding genes, and the position and length of intergenic sequences are provided in Supporting Information Table S2. The start codon of all protein-coding genes is ATG, except in the case of the *nad4* gene of *F. goudeyi*, *G. genuanus* and *L. spurius*, which is GTG. The stop codon showed more variation among genes (TAG, TAA and TA−).

The transcriptomes downloaded from the SRA database were sequenced using various platforms, which rendered different read depths (details in Supporting Information Table S3). We could identify the 13 protein-coding and two rRNA genes in most cases, but one gene was missing for *C. californicus* and *K. tribblei* (Illumina), two for *Gastridium geographus* and *Rolaniconus varius*, three for *Rhizoconus vexillum*, four for *K. tribblei* (454) and six genes for *Splinoconus biliosus* (Supporting Information Figure S1 and Table S3). Regarding the nuclear data set, the matrix completeness (in terms of presence/absence of genes) was 94%. The species that presented more missing data were *K. tribblei* (454; nine missing genes), *Darioconus episcopatus* (five missing genes) and *C. californicus* and *Splinoconus biliosus* (four missing genes; Supporting Information Figure S1 and Table S3).

For six species, it was possible to compare the sequences of the mt genes assembled from the transcriptomes to the corresponding ones from the mt genomes available in GenBank. The percentage of sequence similarity was above 98% for all genes (slight sequence differences may reflect either distinct geographic origins, individual variability or sequencing errors; see Supporting Information Table S4), which confirms the reliability of our pipeline and the data obtained.

3.2 | Phylogenetic analyses

The phylogenetic relationships within the family Conidae were reconstructed based on three different data sets. The mt genome data set had 13,285 or 5,743 positions depending on whether protein-coding genes were analysed at the nucleotide or amino acid level, respectively; the nuclear data set had 17,124 or 5,708 positions; and the combined data set had 30,266 or 11,456 positions. The best-fit partitions and substitution models according to BIC for each data set can be found in Supporting Information Table S5. A total of 18 phylogenetic trees were reconstructed: Three data sets with protein-coding genes analysed at either the nucleotide or the amino acid level and with either ML or BI (using either site-homogeneous or site-heterogeneous models). The 18 phylogenies are shown in Supporting Information Figure S2, which can be accessed at TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S22475), and their log-likelihoods can be consulted in Supporting Information Table S6. Phylogenetic inferences based on BI or ML using site-homogeneous models arrived at very similar or even identical topologies (BI and ML trees of the mt genome data set analysed at the amino acid level; BI and ML trees of the mt genome data set analysed at the nucleotide level; and BI and ML trees of the combined data set analysed at the nucleotide level; Supporting Information Figure S2). Therefore, the main differences among trees were due to the data set or whether the protein-coding genes were analysed at the amino acid or the nucleotide levels. In general, the different trees (Supporting Information Figure S2) agreed on terminal clades, which generally received strong statistical support in all analyses, and differed mainly on those internal nodes, which lacked statistical support (BP < 70; BPP < 0.95) or which directly rendered a polytomy as was the case of trees based on BI using site-heterogeneous models (Supporting Information Figure S2).

The phylogenetic trees with overall better statistical support along the different nodes were reconstructed using BI, site-homogeneous models, and based on protein-coding genes analysed at either the amino acid or the nucleotide levels. Among these trees, the ones including more taxa (53 tips, 27 genera) were those based on mt genomes. Here, we selected arbitrarily the one using protein-coding genes analysed at the amino acid level as our best working hypothesis for the phylogeny of Conidae (Figure 1), as its differences to the corresponding one with protein-coding genes analysed at the nucleotide level are minimum and restricted to statistically unsupported nodes. According to the selected phylogenetic tree, the Conidae are monophyletic and exhibit a long branch, which separate them from outgroup taxa (Figure 1). All the genera with more than one species were recovered as monophyletic (although with low statistical support in the case of *Kioconus*; BPP = 0.63) except *Miliariconus* (due to *Virroconus*, with BPP = 0.54) and *Lautoconus* (due to *Africonus*, with BPP = 0.98). The genus *Rolaniconus* is sister to all remaining Conidae (BPP = 0.99). Within the latter, a well supported clade (BPP = 1) including *Monteiroconus* sister to *Splinoconus* + *Kioconus*

was the sister group of the remaining taxa (Figure 1). The clade (BPP = 0.99) including *Rhizoconus* sister to *Genuanoconus* + *Rhombiconus* (both preying on worms of the family Amphinomidae) was sister to a clade (BPP = 1) including *Virgiconus* sister to *Lividoconus* + *Calamiconus* and a clade with the remaining analysed Conidae (Figure 1). The latter were arranged into two main clades. One included the Caribbean *Lindaconus* (although with low statistical support; BPP = 0.86) sister to paraphyletic *Miliariconus* + *Virroconus* plus West African cone snails (*Kalloconus* sister to paraphyletic *Lautoconus* + *Africonus*; Figure 1). The paraphyly of *Miliariconus* was due to the close sister group relationship of *Miliariconus coronatus* and *Virroconus ebraeus,* whereas the paraphyly of *Lautoconus* was due to the close sister group relationship of *Lautoconus hybridus* + *Lautoconus guanche* and *Africonus* (Figure 1). The other clade included two lineages: (a) *Harmoniconus* + *Puncticulis* sister to a clade including *Fulgiconus* and piscivorous genera from the Indo-Pacific (*Gastridium* + *Pionoconus*); and (b) the Atlantic and East Pacific piscivorous genus *Chelyconus* sister to *Dendroconus* + the molluscivorous genera (*Eugeniconus* + *Conus* sister to *Darioconus* + *Cylinder*; Figure 1).

The corresponding phylogenetic trees (BI, site-homogeneous models, protein-coding genes at the amino acid level) based on the nuclear and the combined data sets had 24 tips, with all genera but *Cylinder*, *Miliariconus* and *Kioconus* represented by a single species (Figure 2). The combined tree showed more resolution than the nuclear tree and a similar topology to the BI tree based on amino acid mt data, except for the relative positions of *Harmoniconus* (here related to the clade *Virroconus* + paraphyletic *Miliariconus*) and *Puncticulis* (here related to *Chelyconus*). The paraphyly of *Miliariconus* was due to a close sister group relationship between *Miliariconus miliaris* and *V. ebraeus* (Figure 2). There was a general lack of resolution of internal nodes (Figure 2).

The phylogenetic trees based on protein-coding genes using site-homogeneous models and analysed at the nucleotide level showed comparable patterns of resolution and recovered generally similar topologies to the corresponding ones based on amino acid data (Supporting Information Figure S2), although some conflicting nodes were detected: (a) The first offshoot was generally *Rhizoconus*; (b) *Rolaniconus*, the clade *Genuanoconus* + *Rhombiconus* or the clade including *Monteiroconus* sister to *Splinoconus* + *Kioconus* were the second offshoot depending on the analysis; (c) *Harmoniconus* was sister to *Chelyconus* in the BI and ML trees based on the mt genome data set; and (d) the paraphyly of *Lautoconus* was due to the close sister group relationship of *Lautoconus ventricosus* and *Africonus*.

The monophylies of genera preying on snails and on amphinomid worms, respectively, were recovered in all phylogenetic analyses including the corresponding species. However, in the case of the piscivorous cones, the mitochondrial data set regardless of the phylogenetic analysis recovered Atlantic/East Pacific (*Chelyconus*) and Indo-Pacific piscivorous (*Pionoconus* and *Gastridium*) genera as two independent lineages. The same result was obtained with the BI analysis of the combined data set with protein-coding genes analysed at the amino acid level under the site-homogeneous model and with the BI analysis of the combined data set with protein-coding genes analysed at the nucleotide level under the site-heterogeneous model (Supporting Information Figure S2). The remaining phylogenetic analyses based on the combined data set and all those based on the nuclear data set recovered the monophyly of the piscivorous cones. The evolutionary trends of radular morphology and the type of protoconch (shell of the larvae) were inferred under unordered parsimony and mapped onto the tree recovered under the BI analysis based on the mt genome data set with protein-coding genes analysed at the amino acid level (Supporting Information Figure S3). The basal spur (see this and other radular tooth characters in the drawings of Supporting Information Figure S3 and in more detail in Tucker & Tenorio, 2009) is present in vermivorous cones but absent in molluscivorous and piscivorous cones. The posterior fold is absent in *Pionoconus* and *Chelyconus*. The anterior portion of the radular tooth is very long in piscivorous cones and the molluscivorous *Darioconus* and *Cylinder*. The terminating cusp is absent in *Gastridium* and modified into an accessory process in *Pionoconus* and *Chelyconus*. These two latter genera as well as *Splinoconus*, *Lividoconus*, *Calamiconous* and *Fulgiconus* lack serrations. Among the studied cones, only those endemic to Cabo Verde (genera *Kalloconus* and *Africonus*) show multiple rows of serrations. Finally, cones endemic to Cabo Verde and Senegal, *Lautoconus ventricosus*, plus *Eugeniconus nobilis* show paucispiral protoconch, which is a proxy of lecithotrophic larvae (Supporting Information Figure S3).

3.3 | Estimation of divergence times

The topology of our best and most complete working hypothesis for the phylogeny of Conidae (based on the mt genome data set with protein-coding genes analysed at the amino acid level using BI with site-homogeneous models) was used as a reference for inferring divergence times. Those nodes with low statistical support (BPP < 0.95) were not fixed. Major cladogenetic events in the evolutionary history of Conidae were dated using an uncorrelated relaxed molecular clock model, which was calibrated using two alternative approaches as explained above. Regardless of the calibration method, the same topology and divergence times were obtained (Figure 3). The origin of the family Conidae was estimated around 62 (68–56) mya, and diversification of extant lineages started about 39 (43–35) mya. There was an active period of diversification between 30 and 25 mya, when most

genera diverged, and analysed species within each genus appeared about 10 mya.

4 | DISCUSSION

Understanding the relative role of different evolutionary processes leading to the extraordinary morphological, ecological and species diversity of cone snails requires a robust phylogeny, which thus far has been elusive (Aman et al., 2015; Phuong & Mahardika, 2018; Puillandre et al., 2014; Uribe, Puillandre, et al., 2017). Several reasons make particularly challenging the reconstruction of the phylogeny of cone snails, including the difficulty of obtaining thorough taxon samplings and the need of gathering large sequence data sets able to accumulate the phylogenetic signal needed to resolve the typical short nodes associated to evolutionary radiations. In recent years, phylogenetic studies of cone snails were based either on (a) medium (>40 species; Aman et al., 2015) or large taxon samplings (>300 species; Puillandre et al., 2014) but few partial mt gene sequences; or (b) in a shorter taxon sampling of most divergent lineages (14 species) but mt genomes (Uribe, Puillandre, et al., 2017), leading to resolution of relatively shallower or deeper nodes of the Tree of Life of cone snails, respectively.

Here, we propose an intermediate approach similar to that of Phuong and Mahardika (2018) based on the analysis of concatenated exons (>4,000; >500,000 bp) from 32 cone species, but complementary and now possible thanks to the ongoing active sequencing of cone venom gland transcriptomes. We gathered up to 41 different species representing 27 genera of Conidae in the mt genome data set and 22 different species belonging to 19 genera of Conidae in the nuclear and combined data sets. Hence, we included 30% and 21% of the genus diversity (Tucker & Tenorio, 2013) in the phylogenetic analyses, respectively. Moreover, we compiled mt genome and nuclear sequence data sets with 15 and 21 complete genes, and 95% and 94% matrix completeness, respectively, adding up to 30,266 or 11,456 positions when combined for phylogenetic analyses based on protein-coding genes analysed at the nucleotide or amino acid levels, respectively.

4.1 | Differences in gene assembly based on the NGS platform

The mt and nuclear genes were assembled from RNA-Seq raw reads, which were generated using five different platforms (Illumina HiSeq 2000, Illumina Genome Analyzer II, Illumina MiSeq, 454 GS FLX and Ion Torrent PGM). These platforms are known to render important differences in terms of read length, depth and quality, which could affect assembly results (Loman et al., 2012). In this regard, the already discontinued 454 GS FLX provided about one order of magnitude less number of reads. This was reflected in that, for instance, the *K. tribblei* assembly generated with this technique presented more missing mt and nuclear genes than any other species (the comparison is particularly illustrative in the case of the assembly of *K. tribblei* based on Illumina raw data, which were generated within the framework of the same study, and only missed the mt *atp8* gene). Similarly, the second assembly missing more of the studied genes was the one of *Splinoconus biliosus*, which was generated upon Ion Torrent PGM raw data. Altogether, Illumina-derived sequence data rendered best assembly results, with many species having the whole set of intended genes for phylogenetic analyses. Among mt genes, *atp8* and *nad5* (the shortest and longest, respectively) were the genes missing in more taxa, whereas the missing nuclear genes appear to be randomly distributed across taxa. In six instances, we could compare the sequences of mt genes assembled from RNA-Seq raw reads to those obtained from more traditional approaches (long PCR amplification and sequencing of complete mt genomes), demonstrating that the assembly pipeline rendered equivalent results (98%–100% sequence similarity).

4.2 | Phylogenetic relationships of cones based on mt and nuclear sequence data

Up to 18 phylogenetic trees were built based on the mt genome, nuclear and combined data sets using ML (site-homogeneous models) and BI (site-homogenous and site-heterogeneous models) with protein-coding genes analysed at the amino acid and nucleotide levels. The method of phylogenetic inference (BI or ML) had little effect on the final reconstructed tree and differences arose from analysing the different data sets or incorporating protein-coding genes as amino acids or nucleotides in the matrices. The general pattern obtained from the different phylogenetic inferences was that most supported nodes were recovered at the tips of the trees and were consistent regardless of the data set and analyses. In contrast, differences were located mostly at most internal nodes, which varied across analyses, and showed, in general, poor statistical support.

Among all reconstructed trees, we consider the two inferred using BI with site-homogeneous models and based on the mt genome data set as the most taxon-rich and with overall higher support. The topologies of these two trees mostly differed on poorly supported internal nodes, and hence, both could be declared our best and most complete working hypotheses for the phylogeny of Conidae. In order to simplify further comparative and evolutionary studies, we chose arbitrarily the one based on protein-coding genes analysed at the amino acid level as reference (also for discussion). This phylogeny recovered the monophyly of the different genera as proposed by Tucker and Tenorio (2009) and based on morphological characters. The only exceptions were

Miliariconus, which was paraphyletic due to *Virroconus* and *Lautoconus*, which was paraphyletic due to *Africonus*. The same relationships were recovered by Puillandre et al. (2014), although these authors consider *Miliariconus* and *Africonus* as synonyms of *Virroconus* and *Lautoconus*, respectively. The recent study by Phuong and Mahardika (2018) based on multiple exons recovered the reciprocal monophylies of *Miliariconus* and *Virroconus*, supporting the validity of both genera. The paraphyly of *Lautoconus*, which is recovered in all our trees, was already reported in a specific phylogenetic study of West African and Mediterranean cones based on mt genomes and a large taxon sampling of endemic species of the region (Abalde, Tenorio, Afonso, Uribe, et al., 2017). In that study, the paraphyly is due to *L. ventricosus* sister to *Africonus* (see also Puillandre et al., 2014), whereas here that relationship is recovered with the analysis of protein-coding genes at the nucleotide level but not with the analysis at the amino acid level, which favours the species from Senegal, *L. guanche* and *L. hybridus* sister to *Africonus*. In this case, given that these phylogenetic relationships are rather shallow, it is likely that the nucleotide data set may have larger phylogenetic signal than the amino acid data set. In our phylogeny, the monophyly of *Kioconus* with respect to *Splinoconus* was recovered in most (with moderate support) but not all analyses (except the mt genome data set analysed with BI and the site-heterogeneous model). However, using a larger number of species representing both genera, Puillandre et al. (2014) recovered a clade with *Kioconus* species intermixed with *Splinoconus* species and considered the former a synonym of the later.

Beyond the monophyly of the different genera, the phylogenetic trees agreed on some sister group relationships among genera: (a) The genus *Darioconus* was consistently recovered sister to *Cylinder* in all phylogenetic analyses. This relationship was also recovered in Puillandre et al. (2014). The genera *Conus*, *Eugeniconus* and *Dendroconus* were placed as closely related to the clade *Darioconus* + *Cylinder*, as in Puillandre et al. (2014), although these authors somewhat unexpectedly recovered *Eugeniconus* within *Cylinder*.

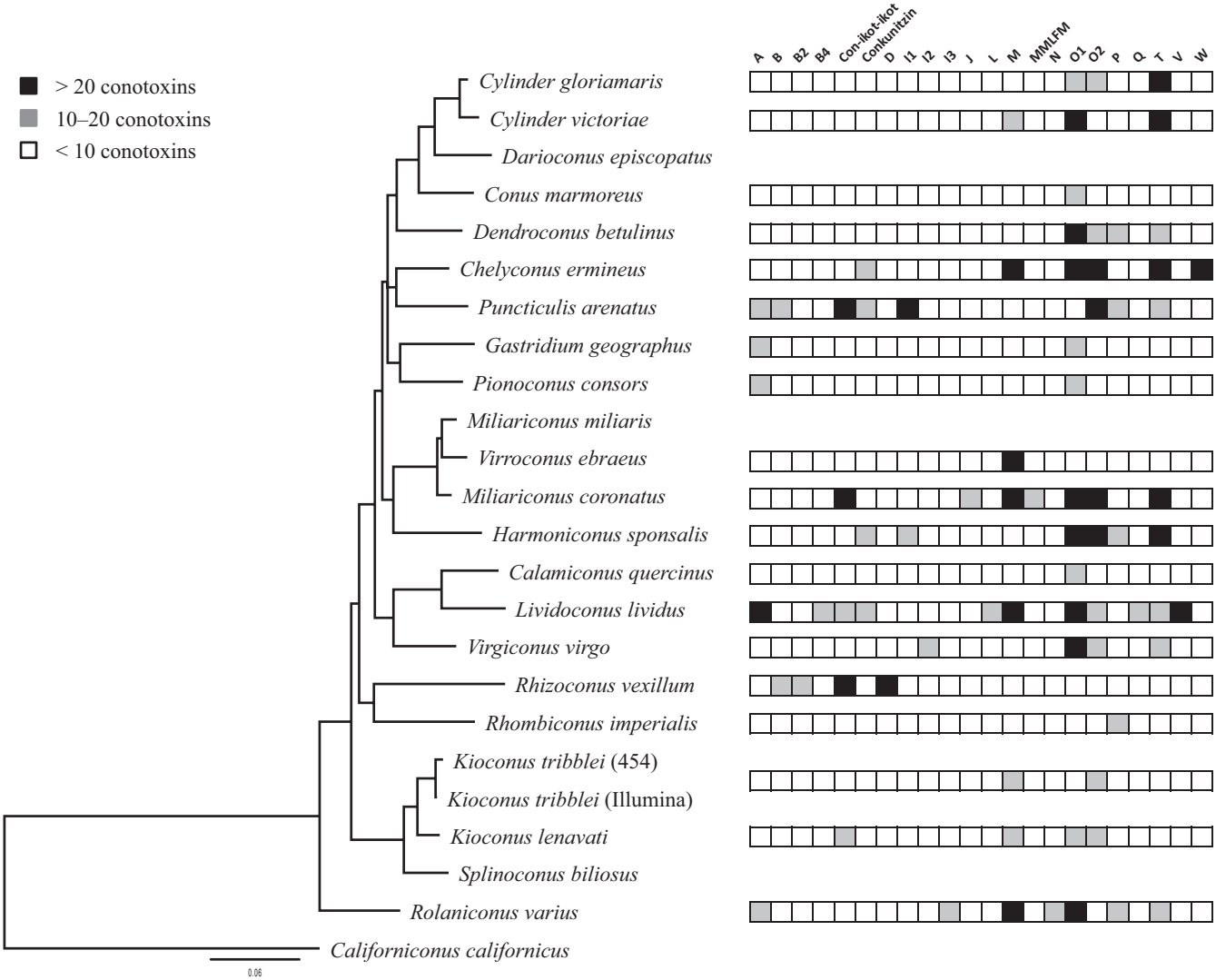


FIGURE 4 Distribution and member diversity of the conotoxin superfamilies across the different studied cone species

Our results are also consistent with the close relationship of *Cylinder* and *Conus* recovered in Aman et al. (2015) and Phuong and Mahardika (2018); (b) the genera *Gastridium* and *Pionoconus* were closely related. This relationship was also recovered in Puillandre et al. (2014), who also included in the clade other genera such as *Phasmoconus* and *Textilia* (see also Aman et al., 2015); (c) the two above-mentioned lineages are related each other and closely related to genera *Fulgiconus*, *Harmoniconus*, *Puncticulis* and *Chelyconus,* although the exact relationship is elusive. The close relationship of all these genera was also found in Puillandre et al. (2014), who synonymized *Fulgiconus* with *Phasmoconus*, and Aman et al. (2015). However, both studies did not include *Fulgiconus goudeyi*, which originally was described as *Phasmoconus goudeyi* (Monnier & Limpalaër, 2012). Puillandre et al. (2014) also included the genera *Lindaconus* and *Virroconus* within this large clade. These two genera were placed as only distantly related in our mt-based phylogenies (although not maximally supported). However, they were included in the large clade in the nuclear-based phylogenies in agreement with Puillandre et al. (2014), but note that the nuclear data set did not include the African/Mediterranean (*Kalloconus*, *Africonus* and *Lautoconus*) cones. The genus *Virroconus* appeared also closely related to *Cylinder* and *Conus* in Aman et al. (2015). A close relationship of *Puncticulis* but not of *Harmoniconus* (see below) to *Pionoconus*, *Cylinder* and *Conus* is recovered in Phuong and Mahardika (2018). In contrast, *Harmoniconus* and *Chelyconus* were sister group in Aman et al. (2015); (d) the close relationship of *Africonus* and *Lautoconus*, and of both to *Kalloconus,* was supported by specific studies on endemic cones of West Africa and the Mediterranean region (Abalde, Tenorio, Afonso, Uribe, et al., 2017) as well as in Puillandre et al. (2014). These three genera were closely related to *Virroconus* and *Miliariconus* in our phylogeny, in contrast to Puillandre et al. (2014). Moreover, *Virroconus* and *Miliariconus* are placed as closely related to *Harmoniconus* in Phuong and Mahardika (2018), as was the case in our phylogenies based on the nuclear and combined data sets, which lacked the West African and Mediterranean cone genera; (e) a strongly supported clade relating *Virgiconus* to *Lividoconus* and *Calamiconus* is recovered in all phylogenetic analyses. This clade is also recovered in the same relative position in the phylogeny of Phuong and Mahardika (2018). In the phylogeny of Puillandre et al. (2014), *Calamiconus quercinus* is considered *Lividoconus quercinus* and recovered as sister to other *Lividoconus*. This genus is sister to *Virgiconus*, and both are placed as closely related to *Kalloconus* and *Lautoconus*, although without support; and (f) *Genuanoconus* and *Rhombiconus* are always recovered as sister taxa. In Puillandre et al. (2014) and Aman et al. (2015), *Rhombiconus imperialis* is considered *Stephanoconus imperialis* and the genus placed as the second offshoot of Conidae after *Fraterconus distans* (and before *Strategoconus*) or sister to *Strategoconus* (and both to *Rhizoconus*), respectively. In Phuong and Mahardika (2018), *R. imperialis* is sister to *Strategoconus* and *Rhizoconus*. With regard to *Genuanoconus genuanus*, Puillandre et al. (2014) consider this species member of *Kalloconus*, as it was recovered deeply nested within this genus. However, a misidentification of the sample in the original work (Cunha et al., 2005) most likely explains this result, and *Genuanoconus* should not be considered a synonym of *Kalloconus* but a distantly related genus.

Most of the differences between inferred trees were concentrated in deepest nodes, affecting the relative position of *Rolaniconus*, *Rhizoconus*, the clade *Rhombiconus* + *Genuanoconus* and the clade including *Monteiroconus* sister to *Splinoconus* + *Kioconus*. In Puillandre et al. (2014) and Phuong and Mahardika (2018), the first diverging lineage of the tree is represented by *F. distans*, a species that we could not incorporate into our analysis. The next diverging lineages in Puillandre et al. (2014) are successively *Stephanoconus* (i.e., *Rhombiconus*), *Strategoconus* (the species *Rolaniconus varius* is considered *Strategoconus varius*) and a clade including, among others, the genus *Turriconus* sister to *Monteiroconus* and *Splinoconus* (including *Kioconus*). Lastly, *Rhizoconus* is recovered in a more derived position as sister to the remaining Conidae. In Phuong and Mahardika (2018), the next diverging lineage includes *Rhombiconus* sister to *Rolaniconus* + *Rhizoconus* (but *Genuanoconus*, *Splinoconus* and *Kioconus* are not included). Finally, in Aman et al. (2015), the first offshoot is *Kioconus* + *Leporiconus*, the second is *Turriconus* and the third is a clade including *Rhizoconus* sister to *Rolaniconus* + *Rhombiconus*. Hence, our phylogenetic analyses and reported trees concur that the above-mentioned species are close to the initial diversification of Conidae but are unable to resolve the exact phylogenetic relationships. Discrepancies among studies could be mainly related to uneven taxon sampling (each study is missing relevant lineages) in the phylogenomic analyses (this work, Phuong & Mahardika, 2018) or lack of enough phylogenetic signal due to relatively small data sets (Aman et al., 2015; Puillandre et al., 2014). The possibility of long branch attraction to the root (Philippe & Laurent, 1998) is less likely as none of these genera shows particularly high evolutionary rates. In any case, it is not possible to shorten the long branch connecting the outgroup and the ingroup as there are no genera more closely related to Conidae than those already included here (Uribe, Puillandre, et al., 2017).

4.3 | Evolution of diet specialization, radular morphology and conotoxin diversity

The reconstructed phylogeny was used as framework to infer the evolution of different traits relevant to the diversification of the group. One key character in Conidae is the feeding mode (Duda, Kohn, & Palumbi, 2001). Here, the ancestor of Conidae was inferred to prey on marine worms

in agreement with previous studies (Duda et al., 2001; Puillandre et al., 2014). The taxonomic and ecological data on which exact worm species are eaten by the different cone species are rather old, scattered and only the family level is determined (Phuong & Mahardika, 2018), so it is not possible to elaborate further on this subject, although subtle differences in radular tooth morphology may point to the existence of some degree of prey specialization. A striking exception is the case of those cone species hunting on fire worms (family Amphinomidae), which show distinct radular teeth (Nybakken, 1970), and in our phylogeny are recovered together as sister taxa, suggesting a single origin for this specialization (Duda et al., 2001). This is particularly remarkable as genera *Genuanoconus* and *Rhombiconus* are from the Eastern Atlantic and Indo-Pacific oceans, respectively. Other cone species preying on amphinomids belong to the Western Atlantic genera *Stephanoconus* and *Tenorioconus*, which are both placed as closely related to *Rhombiconus* in reconstructed phylogenies (Aman et al., 2015; Puillandre et al., 2014). The shift to feed on snails also occurred once in the evolutionary history of the group according to our phylogeny. This result is consistent across phylogenetic studies (Aman et al., 2015; Duda et al., 2001; Puillandre et al., 2014) and further supported by the characteristic (i.e., synapomorphic) curved and slender radular teeth without waist and spur of all molluscivorous species (Nishi & Kohn, 1999), which are repeatedly shot onto each single prey (Kohn, 2003). Finally, according to our phylogenies, the fish-feeding mode, arguably the most complex hunting behaviour among cones (Olivera, Seger, Horvath, & Fedosov, 2015), may have at least two independent origins in the Indo-Pacific and Atlantic/Eastern Pacific regions, respectively. This result was mainly supported by the mitochondrial data. Instead, the monophyly of piscivorous cones was favoured by all phylogenetic analyses based on the nuclear data set and some based on the combined data set. However, these data sets missed key genera to adequately tackle the question. Thus far, all previous phylogenetic studies have recovered piscivorous cones polyphyletic, although with low support (Aman et al., 2015; Duda et al., 2001; Puillandre et al., 2014). While there is no documented evidence that Indo-Pacific *Gastridium* and *Pionoconus* species feed on other prey than fish, Atlantic/Eastern Pacific *Chelyconus* species may also consume other molluscs (Olivera et al., 2015), which may indicate different evolutionary origins of piscivory in these taxa. Moreover, the comparison of the conotoxin repertoires of Indo-Pacific versus Atlantic/Eastern Pacific cones also supported independent origins of piscivory (Abalde et al., 2018). If true, many of the modifications in the radular teeth that are characteristic of *Pionoconus* and *Chelyconus* would be convergent. The studied radular tooth characters that differed between groups were mostly associated with the overall peculiar radular teeth shape of cone species depending on their diet. Hence, the potential differences and limitations in the ancestral character state reconstructions of these characters depending on the reconstructed phylogeny are the same discussed for the diet specializations. Instead, the protoconch evolutionary trends do not vary when different reconstructed phylogenies are considered, as the involved clades are consistently recovered throughout all analyses.

Another important trait in cone diversification and evolution is related with the diversity of the venom cocktails produced by the different species. Conotoxins are organized into superfamilies according to the signal region of the precursor, which is highly conserved (Puillandre, Koua, Favreau, Olivera, & Stöcklin, 2012). When analysing reported venom gland transcriptomes from various cone species, the emerging general pattern is that several conotoxin superfamilies (e.g., O1, M and T) are widespread among cones and constitute the minimal set required for the effective function of the venom, whereas others are restricted to a few lineages (Duda & Remigio, 2008; Puillandre et al., 2012). The different conotoxin superfamilies show diverse degrees of expansion. Here, we obtained listings of conotoxins of the different species directly from the original literature and mapped the number of described conotoxins per superfamily onto the phylogeny and inferred the evolution of conotoxin superfamily expansions (Figure 4). The genera *Chelyconus*, *Puncticulis*, *Miliariconus*, *Lividoconus* and *Rolaniconus* were the ones showing more superfamilies expanded (Figure 4). In contrast, genera *Conus* and *Calamiconus* had only superfamily O1 expanded, genus *Rhombiconus* only superfamily P and genus *Virroconus* only superfamily M (Figure 4). Superfamily O1 showed more than 20 members in many of the studied genera, although had less than 10 in *Puncticulis*, *Rhizoconus*, *Virroconus* and *Rhombiconus* (Figure 4). Similarly, superfamilies M and T were also highly diverse (>20 members) in several genera, although no specific evolutionary trend was inferred (Figure 4). Superfamily O2 was also expanded in many genera, although in most cases, the number of members varied between 10 and 20. The Indo-Pacific piscivorous genera *Pionoconus* and *Gastridium* showed expansions only in superfamilies A and O1, whereas the Atlantic/Eastern Pacific piscivorous genus *Chelyconus* showed expansions in superfamilies O1, O2, M and T as other genera plus in W and conkunitzin but not in A (Figure 4). The conkunitzins were also expanded in *Puncticulis*, *Harmoniconus* and *Lividoconus*, always with the number of members between 10 and 20. The genus *Rhizoconus* showed very specific expansions in superfamilies B, B2, D and con-ikot-ikot. These general patterns of conotoxin diversity distribution across genera are tentative and should be interpreted with caution as venom gland transcriptomes were obtained

using different methodologies and sequencing platforms as well as were assembled using different reference databases. Evolutionary trends on the expansions of superfamilies should be treated at this point as exploratory since key genera (and species) still need to be added to the phylogeny to reach stronger conclusions.

4.4 | Divergence times of major cladogenetic events

Calibration of the molecular clock using alternative approaches rendered the same chronogram indicating that the ages and lineage ascriptions of the three fossils used in the second approach were consistent with the ages used in the first approach. The inferred chronogram showed that first diversification of extant lineages within family Conidae occurred about 40 mya in the Eocene. We did not include *Fraterconus* in our analyses, a genus that has been recovered as the first offshoot in other phylogenetic studies (Phuong & Mahardika, 2018; Puillandre et al., 2014), and thus, it is likely that the diversification started somewhat earlier. In any case, it is evident that a long gap occurred between the origin of Conidae and the diversification of its extant lineages, indicating important lineage extinction events or eventually low speciation rates during the Late Paleocene–Early Eocene. The main lineages in the phylogeny appeared rapidly in only 5–10 my during the Oligocene, and the origin of many current genera was inferred to be relatively old and could be dated back to the Oligocene–Miocene transition, when there was a major radiation, as stated in the fossil record (Kohn, 1990). At that time, there was a global cooling event (Miller, 2005; Zachos, Flower, & Paul, 1997), accompanied by a sea level drop about 50 m (Beddow, Liebrand, Sluijs, Wade, & Lourens, 2016), which likely produced abrupt changes on the coast morphology and on intertidal habitats triggering diversification events as in other marine species (Davis, Hill, Astrop, & Wills, 2016). Other genera appeared steadily during the Miocene, and it appears that genus diversification was completed by the end of this period. We cannot confidently date when bursts of diversification leading to current species richness within each genus occurred as we had few instances in the phylogeny in which more than one species per genus were included. However, all analysed within-genus diversifications were dated in the Pliocene and Pleistocene concurring with glacial–interglacial events (Lisiecki & Raymo, 2005) in agreement with evolutionary studies analysing the radiation of cones in the Cabo Verde archipelago and the Senegal coast (Abalde, Tenorio, Afonso, & Zardoya, 2017; Abalde, Tenorio, Afonso, Uribe, et al., 2017; Cunha et al., 2005; Duda & Rolán, 2005). The early diversification of Conidae occurred in the Indo-Pacific region, which is consistent with the high species richness of this region. The diverse Atlantic lineages originated independently in different clades but almost all simultaneously about 25 mya during the drastic global climate and sea level changes of the Oligocene–Miocene transition.

5 | CONCLUSIONS

The use of mt genomes and several complete nuclear genes assembled from RNA-Seq raw reads allowed us reconstructing a phylogeny of Conidae including representatives of up to 27 of the 89 currently described genera within this group. This phylogeny had good levels of resolution, although the relative position of the early emerging lineages remains uncertain. The reconstructed phylogeny is comparable in resolution to a very recent one of similar number of genera and based on concatenated exons (Phuong & Mahardika, 2018) and agrees at the tips to the most taxon complete published thus far based on partial mt gene sequences (Puillandre et al., 2014). Hence, our results suggest that until sequencing technologies improve, it may be a sufficient (and more economical) compromise selecting a few tens of (mt and nuclear) genes and having complete data matrices instead of gathering thousands of loci in rather incomplete data sets to achieve similar results in terms of resolved trees, particularly when expanding phylogenetic studies to large taxon samplings. The new phylogeny could be further improved in the future by adding new key taxa representing missing genera and by enlarging the nuclear data set. In any case, this phylogeny provides a robust backbone to further understand the evolutionary processes underlying the great diversification of Conidae, supporting, for example, the single origin of the diet shifts to feeding on amphinomid worms and molluscivory, but the likely independent origins of piscivory in the Indo-Pacific and Atlantic/Eastern Pacific cones, respectively.

ACKNOWLEDGEMENTS

We are grateful to Carlos M. L. Afonso as well as to Paulo Vasconcelos and Stiven Pires for providing specimens of *L. spurius* and of *G. genuanus*, respectively. We are indebted to Frank Leterrier, who provided us with *F. goudeyi* tissue from New Caledonia. We thank David Posada and Sara Rocha for their bioinformatic advice in the handling of transcriptomic data. We thank Jesús Marco and Aida Palacios, who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), and member of the Spanish Supercomputing Network, for performing assembly and phylogenetic analyses. We are thankful to Paloma Chamorro for the gastropod and amphinomid silhouettes. This work was supported by the Spanish Ministry of Science, Industry, and Innovation (CGL2016-75255-C2-1-P (AEI/FEDER, UE) to RZ; BES-2014-069575 and EEBB-I-16-11775 to SA; and BES-2011-051469 to JEU).

ORCID

Manuel J. Tenorio https://orcid.org/0000-0003-4088-4958
Rafael Zardoya https://orcid.org/0000-0001-6212-9502

REFERENCES

Abalde, S., Tenorio, M. J., Afonso, C. M. L., Uribe, J. E., Echeverry, A. M., & Zardoya, R. (2017). Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitochondrial genomes. *BMC Evolutionary Biology*, *17*, 231.

Abalde, S., Tenorio, M. J., Afonso, C. M. L., & Zardoya, R. (2017). Mitogenomic phylogeny of cone snails endemic to Senegal. *Molecular Phylogenetics and Evolution*, *112*, 79–87.

Abalde, S., Tenorio, M. J., Afonso, C. M. L., & Zardoya, R. (2018). Conotoxin diversity in *Chelyconus ermineus* (Born, 1778) and the convergent origin of piscivory in Atlantic and Indo-Pacific cones. *Genome Biology and Evolution*, *10*, 2643–2662.

Abascal, F., Zardoya, R., & Telford, M. J. (2010). TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*, *38*, W7–W13. https://doi.org/10.1093/nar/gkq291

Aman, J. W., Imperial, J. S., Ueberheide, B., Zhang, M.-M., Aguilar, M., Taylor, D., …Olivera, B. M. (2015). Insights into the origin of fish hunting in venomous cone snails from studies of *Conus tessulatus*. *Proceedings of the National Academy of Sciences, USA*, *112*, 5087–5092.

Bandyopadhyay, P. K., Stevenson, B. J., Ownby, J.-P., Cady, M. T., Watkins, M., & Olivera, B. M. (2008). The mitochondrial genome of *Conus textile*, *coxI-coxII* intergenic sequences and conoidean evolution. *Molecular Phylogenetics and Evolution*, *46*, 215–223. https://doi.org/10.1016/j.ympev.2007.08.002

Barghi, N., Concepcion, G. P., Olivera, B. M., & Lluisma, A. O. (2015). High conopeptide diversity in *Conus tribblei* revealed through analysis of venom duct transcriptome using two High-Throughput sequencing platforms. *Marine Biotechnology*, *17*, 81–98. https://doi.org/10.1007/s10126-014-9595-7

Barghi, N., Concepcion, G. P., Olivera, B. M., & Lluisma, A. O. (2016). Characterization of the complete mitochondrial genome of *Conus tribblei* Walls, 1977. *Mitochondrial DNA. Part A*, *27*, 4451–4452.

Beddow, H. M., Liebrand, D., Sluijs, A., Wade, B. S., & Lourens, L. J. (2016). Global change across the Oligocene-Miocene transition: High-resolution stable isotope records from IODP Site U1334 (equatorial Pacific Ocean). *Paleoceanography*, *31*, 81–97. https://doi.org/10.1002/2015PA002820

Beu, A. G., & Maxwell, P. A. (1990). Cenozoic Mollusca of New Zealand. *New Zealand Geological Survey of Paleontology Bulletin*, *58*, 1–518.

Bleidorn, C., Podsiadlowski, L., Zhong, M., Eeckhaut, I., Hartmann, S., Halanych, K. M., & Tiedemann, R. (2009). On the phylogenetic position of Myzostomida: Can 77 genes get it wrong? *BMC Evolutionary Biology*, *9*, 150. https://doi.org/10.1186/1471-2148-9-150

Boore, J. L., Macey, J. R., & Medina, M. (2005). Sequencing and comparing whole mitochondrial genomes of animals. *Methods in Enzymology*, *395*, 311–348.

Brauer, A., Kurz, A., Stockwell, T., Baden-Tillson, H., Heidler, J., Wittig, I., … Remm, M. (2012). The mitochondrial genome of the venomous cone snail *Conus consors*. *PLoS ONE*, *7*, e51528. https://doi.org/10.1371/journal.pone.0051528

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Chen, P. W., Hsiao, S. T., Chen, K. S., Tseng, C. T., Wu, W. L., & Hwang, D. F. (2016a). The complete mitochondrial genome of *Conus capitaneus* (Neogastropoda: Conidae). *Mitochondrial DNA. Part B*, *1*, 520–521.

Chen, P. W., Hsiao, S. T., Chen, K. S., Tseng, C. T., Wu, W. L., & Hwang, D. F. (2016b). The complete mitochondrial genome of *Conus striatus* (Neogastropoda: Conidae). *Mitochondrial DNA. Part B*, *1*, 493–494.

Chen, P. W., Hsiao, S. T., Huang, C. W., Chen, K. S., Tseng, C. T., Wu, W. L., & Hwang, D. F. (2016). The complete mitochondrial genome of *Conus tulipa* (Neogastropoda: Conidae). *Mitochondrial DNA. Part A*, *27*, 2738–2739.

Cunha, R. L., Castilho, R., Rüber, L., & Zardoya, R. (2005). Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde islands. *Systematic Biology*, *54*, 634–650. https://doi.org/10.1080/106351591007471

Cunha, R., Grande, C., & Zardoya, R. (2009). Neogastropod phylogenetic relationships based on entire mitochondrial genomes. *BMC Evolutionary Biology*, *9*, 210. https://doi.org/10.1186/1471-2148-9-210

Davis, K. E., Hill, J., Astrop, T. I., & Wills, M. A. (2016). Global cooling as a driver of diversification in a major marine clade. *Nature Communications*, *7*, 13003. https://doi.org/10.1038/ncomms13003

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214. https://doi.org/10.1186/1471-2148-7-214

Duda, T. F., Bolin, M. B., Meyer, C. P., & Kohn, A. J. (2008). Hidden diversity in a hyperdiverse gastropod genus: Discovery of previously unidentified members of a *Conus* species complex. *Molecular Phylogenetics and Evolution*, *49*, 867–876.

Duda, T. F., & Kohn, A. J. (2005). Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. *Molecular Phylogenetics and Evolution*, *34*, 257–272. https://doi.org/10.1016/j.ympev.2004.09.012

Duda, T. F., Kohn, A. J., & Palumbi, S. R. (2001). Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. *Biological Journal of the Linnean Society*, *73*, 391–409. https://doi.org/10.1111/j.1095-8312.2001.tb01369.x

Duda, T. F., & Remigio, E. A. (2008). Variation and evolution of toxin gene expression patterns of six closely related venomous marine snails. *Molecular Ecology*, *17*, 3018–3032. https://doi.org/10.1111/j.1365-294X.2008.03804.x

Duda, T. F., & Rolán, E. (2005). Explosive radiation of Cape Verde *Conus*, a marine species flock. *Molecular Ecology*, *14*, 267–272. https://doi.org/10.1111/j.1365-294X.2004.02397.x

Dunn, C. W., Howison, M., & Zapata, F. (2013). Agalma: An automated phylogenomics workflow. *BMC Bioinformatics*, *14*, 330. https://doi.org/10.1186/1471-2105-14-330

Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., … Lewis, R. J. (2014). Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nature Communications*, *5*, 3521. https://doi.org/10.1038/ncomms4521

Espeland, M., Breinholt, J., Willmott, K. R., Warren, A. D., Vila, R., Toussaint, E. F. A., … Kawahara, A. Y. (2018). A comprehensive and dated phylogenomic analysis of butterflies. *Current Biology*, *28*, 1–9. https://doi.org/10.1016/j.cub.2018.01.061

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*, 368–376. https://doi.org/10.1007/BF01734359

Gao, B., Peng, C., Chen, Q., Zhang, J., & Shi, Q. (2018). Mitochondrial genome sequencing of a vermivorous cone snail *Conus quercinus* supports the correlative analysis between phylogenetic relationships and dietary types of *Conus* species. *PLoS ONE*, *13*, e0193053. https://doi.org/10.1371/journal.pone.0193053

Glez-Pena, D., Gomez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010). ALTER: Program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research*, *38*, W14–W18. https://doi.org/10.1093/nar/gkq321

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652. https://doi.org/10.1038/nbt.1883

Haas, B. J. (2016). *TransDecoder*. Retrieved from https://github.com/TransDecoder/

Hedtke, S. M., Townsend, T. M., & Hillis, D. M. (2006). Resolution of phylogenetic conflict in larga data sets by increased taxon sampling. *Systematic Biology*, *55*, 522–529.

Holm, P. M., Grandvuinet, T., Friis, J., Wilson, J. R., Barker, A. K., & Plesner, S. (2008). An 40Ar-39Ar study of the Cape Verde hot spot: Temporal evolution in a semistationary plate environment. *Journal of Geophysical Research*, *113*, B08201.

Hu, H., Bandyopadhyay, P. K., Olivera, B. M., & Yandell, M. (2012). Elucidation of the molecular envenomation strategy of the cone snail *Conus geographus* through transcriptome sequencing of its venom duct. *BMC Genomics*, *13*, 284. https://doi.org/10.1186/1471-2164-13-284

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.-Y., Kupfer, A., … Philippe, H. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nature Ecology and Evolution*, *1*, 1370–1378. https://doi.org/10.1038/s41559-017-0240-5

Jeffroy, O., Brinkmann, H., Delsuc, F., & Philippe, H. (2006). Phylogenomics: The beginning of incongruence? *Trends in Genetics*, *22*, 225–231. https://doi.org/10.1016/j.tig.2006.02.003

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kohn, A. J. (1959). The ecology of *Conus* in Hawaii. *Ecological Monographs*, *29*, 47–90. https://doi.org/10.2307/1948541

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

Kohn, A. J. (2003). The feeding process in *Conus victoriae*. In F. E. Wells, D. I. Walker, & D. S. Jones (Eds.), *The marine flora and fauna of dampier, Western Australia* (pp. 101–108). Perth, WA: Western Australian Museum.

Kohn, A. J. (2014). *Conus of the Southeastern United States and Caribbean*. Princeton, NJ: Princeton University Press.

Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J., & Zardoya, R. (2015). TRUFA: A user-friendly web server for *de novo* RNA-seq analysis using cluster computing. *Evolutionary Bioinformatics*, *11*, 97–104.

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). Partitionfinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*, 1695–1701. https://doi.org/10.1093/molbev/mss020

Lartillot, N., Rodrigue, N., Stubbs, D., & Richer, J. (2013). PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, *62*, 611–615. https://doi.org/10.1093/sysbio/syt022

Lavergne, V., Harliwong, I., Jones, A., Miller, D., Taft, R. J., & Alewood, P. F. (2015). Optimized deep-targeted proteotranscriptomic profiling reveals unexplored *Conus* toxin diversity and novel cysteine frameworks. *Proceedings of the National Academy of Sciences USA*, *112*, E3782–E3791. https://doi.org/10.1073/pnas.1501334112

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. *Journal of Proteome Research*, *11*, 5046–5058.

Li, Q., Barghi, N., Lu, A., Fedosov, A. E., Bandyopadhyay, P. K., Lluisma, A. O., … Safavi-Hemami, H. (2017). Divergence of the venom exogene repertoire in two sister species of *Turriconus*. *Genome Biology and Evolution*, *9*, 2211–2225. https://doi.org/10.1093/gbe/evx157

Lisiecki, L. E., & Raymo, M. E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography*, *20*, PA1003.

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., & Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*, *30*, 434–439. https://doi.org/10.1038/nbt.2198

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, *44*, W54–W57. https://doi.org/10.1093/nar/gkw413

Maddison, W. P., & Maddison, D. R. (2018). *Mesquite: A modular system for evolutionary analysis*. Version 3.40. Retrieved from https://mesquiteproject.org

Miljanich, G. P. (2004). Ziconotide: Neuronal calcium channel blocker for treating severe chronic pain. *Current Medicinal Chemistry*, *11*, 3029–3040.

Miller, K. G. (2005). The Phanerozoic record of global sea-level change. *Science*, *310*, 1293–1298. https://doi.org/10.1126/science.1116412

MolluscaBase (2018). *Conidae J. Fleming, 1822. Accessed through: World Register of Marine Species*. Retrieved from https://www.marinespecies.org/aphia.php?p=taxdetails&xml:id=14107 Accessed 2018-07-12.

Monnier, E., & Limpalaër, L. (2012). *Phasmoconus goudeyi* (Gastropoda: Conidae) a new species from New Caledonia. *Visaya*, *3*, 41–44.

Monnier, E., Tenorio, M. J., Bouchet, P., & Puillandre, N. (2018). The cones (Gastropoda) from Madagascar "Deep South": Composition, endemism and new taxa. *Xenophora Taxonomy*, *19*, 25–75.

Nishi, M., & Kohn, A. J. (1999). Radular teeth of Indo-Pacific molluscivorous species of *Conus*: A compartive analysis. *Journal of Molluscan Studies*, *65*, 483–497.

Norton, R. S., & Olivera, B. M. (2006). Conotoxins down under. *Toxicon*, *48*, 780–798. https://doi.org/10.1016/j.toxicon.2006.07.022

Nybakken, J. W. (1970). Correlation of radula tooth structure and food habits of three vermivorous species of *Conus*. *The Veliger*, *12*, 316–318.

Olivera, B. M., Seger, J., Horvath, M. P., & Fedosov, A. E. (2015). Prey-capture strategies of fish-hunting cone snails: Behavior, neurobiology and evolution. *Brain, Behavior and Evolution*, *86*, 58–74. https://doi.org/10.1159/000438449

Osca, D., Templado, J., & Zardoya, R. (2015). Caenogastropoda mitogenomics. *Molecular Phylogenetics and Evolution*, *93*, 118–128.

Peng, C., Yao, G., Gao, B.-M., Fan, C.-X., Bian, C., Wang, J., … Shi, Q. (2016). High-throughput identification of novel conotoxins from the Chinese tubular cone snail (*Conus betulinus*) by multitranscriptome sequencing. *GigaScience*, *5*, 17. https://doi.org/10.1186/s13742-016-0122-9

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biology*, *9*, e1000602. https://doi.org/10.1371/journal.pbio.1000602

Philippe, H., & Laurent, J. (1998). How good are deep phylogenetic trees? *Current Opinion in Genetics & Development*, *8*, 616–623. https://doi.org/10.1016/S0959-437X(98)80028-2

Phuong, M. A., & Mahardika, G. N. (2018). Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. *Molecular Biology and Evolution*, *35*, 1210–1224. https://doi.org/10.1093/molbev/msy034

Phuong, M. A., Mahardika, G. N., & Alfaro, M. E. (2016). Dietary breadth is positively correlated with venom complexity in cone snails. *BMC Genomics*, *17*, 401. https://doi.org/10.1186/s12864-016-2755-6

Pin, M., & Leung Tack, K. D. (1995). The cones of Senegal. *Conchiglia*, *277*, 1–56.

Prashanth, J. R., Dutertre, S., Jin, A. H., Lavergne, V., Hamilton, B., Cardoso, F. C., ... Lewis, R. J. (2016). The role of defensive ecological interactions in the evolution of conotoxins. *Molecular Ecology*, *25*, 598–615. https://doi.org/10.1111/mec.13504

Puillandre, N., Bouchet, P., Duda, T. F. Jr, Kauferstein, S., Kohn, A. J., Olivera, B. M., … Meyer, C. (2014). Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). *Molecular Phylogenetics and Evolution*, *78*, 290–303. https://doi.org/10.1016/j.ympev.2014.05.023

Puillandre, N., Koua, D., Favreau, P., Olivera, B. M., & Stöcklin, R. (2012). Molecular phylogeny, classification and evolution of conopeptides. *Journal of Molecular Evolution*, *74*, 297–309. https://doi.org/10.1007/s00239-012-9507-2

Puillandre, N., & Tenorio, M. J. (2017). A question of rank: DNA sequences and radula characters reveal a new genus of cone snails (Gastropoda: Conidae). *Journal of Molluscan Studies*, *83*, 200–210. https://doi.org/10.1093/mollus/eyx011

Rambaut, A., & Drummond, A. J. (2007). *Tracer v1.4*. Retrieved from https://beast.bio.ed.ac.uk/Tracer

Robinson, S. D., Li, Q., Lu, A., Bandyopadhyay, P. K., Yandell, M., Olivera, B. M., & Safavi-Hemami, H. (2017). The venom repertoire of Conus gloriamaris (Chemnitz, 1777), the glory of the sea. *Marine Drugs*, *15*, 145. https://doi.org/10.3390/md15050145

Robinson, S. D., Safavi-Hemami, H., McIntosh, L. D., Purcell, A. W., Norton, R. S., & Papenfuss, A. P. (2014). Diversity of conotoxin gene superfamilies in the venomous snail, Conus victoriae. *Plos One*, *9*, e87648. https://doi.org/10.1371/journal.pone.0087648

Rokas, A., Williams, B. L., King, N., & Carroll, S. B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, *425*, 798–804. https://doi.org/10.1038/nature02053

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, *27*, 863–864. https://doi.org/10.1093/bioinformatics/btr026

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464. https://doi.org/10.1214/aos/1176344136

Sharma, P. P., Baker, C. M., Cosgrove, J. G., Johnson, J. E., Oberski, J. T., Raven, R. J., … Giribet, G. (2018). A revised dated phylogeny of scorpions: Phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. *Molecular Phylogenetics and Evolution*, *122*, 37–45. https://doi.org/10.1016/j.ympev.2018.01.003

Shuto, T. (1969). Neogene gastropods from Panay Island, the Philippines. Contributions to the geology and paleontology of Southeast Asia, LXVII. *Memoirs of the Faculty of Science, Kyushu University, Series D, Geology*, *19*, 1–250.

Smith, S. A., Wilson, N. G., Goetz, F. E., Feehery, C., Andrade, S. C. S., Rouse, G. W., … Dunn, C. W. (2011). Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature*, *480*, 364–367. https://doi.org/10.1038/nature10526

Squire, R. L. (1987). Eocene molluscan paleontology of the Whitaker Peak area, Los Angeles and Ventura Counties, California. *Natural History Museum of Los Angeles County Contributions in Science*, *388*, 50–51.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Tracey, S., Craig, B., Belliard, L., & Gain, O. (2017). One, four or forty species? – Early Conidae (Mollusca, Gastropoda) that led to a radiation and biodiversity peak in the late Lutetian Eocene of the Cotentin, NW France. *Carnets De Voyages Paléontologiques Dans Le Bassin Anglo-Parisien*, *3*, 1–38.

Tucker, J. K., & Tenorio, M. J. (2009). *Systematic classification of recent and fossil conoidean gastropods, with keys to the genera of cone shells*. Harxheim, Germany: ConchBooks.

Tucker, J. K., & Tenorio, M. J. (2013). *Illustrated catalog of the living cone shells*. Wellington, FL: MdM Publishing.

Uribe, J. E., Puillandre, N., & Zardoya, R. (2017). Beyond *Conus*: Phylogenetic relationships of Conidae based on complete mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *107*, 142–151.

Uribe, J. E., Williams, S. T., Templado, J., Abalde, S., & Zardoya, R. (2017). Denser mitogenomic sampling improves resolution of the phylogeny of the superfamily Trochoidea (Gastropoda: Vetigastropoda). *Journal of Molluscan Studies*, *83*, 111–118.

Vargas, P., & Zardoya, R. (2014). *The tree of lIfe*. Sunderland, MA: Oxford University Press.

Weese, D. A., & Duda, T. F. (2015). Transcriptomic resources for three populations of *Conus miliaris* (Mollusca: Conidae) from Easter Island, American Samoa and Guam. *Molecular Ecology Resources*, *15*, 228–229. https://doi.org/10.1111/1755-0998.12340

Weigert, A., Helm, C., Meyer, M., Nickel, B., Arendt, D., Hausdorf, B., … Struck, T. H. (2014). Illuminating the base of the annelid

tree using transcriptomics. *Molecular Biology and Evolution*, *31*, 1391–1401. https://doi.org/10.1093/molbev/msu080

Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., … Leebens-Mack, J. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences, USA*, *111*, E4859. https://doi.org/10.1073/pnas.1323926111

Yang, M., Zhao, S., Min, X., Shao, M., Chen, Y., Chen, Z., & Zhou, M. (2017). A novel mu-conotoxin from worm-hunting *Conus tessulatus* that selectively inhibit rat TTX-resistant sodium currents. *Toxicon*, *130*, 11–18.

Zachos, J. C., Flower, B. P., & Paul, H. (1997). Orbitally paced climate oscillations across the Oligocene/Miocene boundary. *Nature*, *388*, 567–570. https://doi.org/10.1038/41528

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Abalde S, Tenorio MJ, Uribe JE, Zardoya R. Conidae phylogenomics and evolution. *Zool Scr*. 2019;48:194–214. https://doi.org/10.1111/zsc.12329

3.2. – Chapter II: “Mitogenomic phylogeny of cone snails endemic to Senegal”

3.2. - Capítulo II: “Filogenia basada en mitogenomes de los conos endémicos de Senegal”

Phylogenetics and Evolution, 112: 79-87

Los conos alcanzan en Senegal uno de sus picos más altos de diversidad de especies en toda África occidental. Se han descrito un total de 15 especies endémicas, todas asignadas al género *Lautoconus*. Si bien hay amplios datos sobre la morfología de la concha y el diente radular en estas especies, prácticamente no se sabe nada sobre la diversidad genética y las relaciones filogenéticas de uno de los grupos de conos más amenazados. En este trabajo, determinamos el genoma mitocondrial (mt) completo o casi completo (sólo faltaría la región de control) de 17 especímenes, que representan 11 especies endémicas (*Lautoconus belairensis*, *Lautoconus bruguieresi*, *Lautoconus cacao*, *Lautoconus cloveri*, *Lautoconus* cf. *echinophilus*, *Lautoconus guinaicus*, *Lautoconus hybridus*, *Lautoconus senegalensis*, *Lautoconus mercator, Lautoconus taslei* y *Lautoconus unifasciatus*). También secuenciamos el genoma mt completo de *Lautoconus guanche* de las Islas Canarias, que ha sido relacionado con los conos endémicos de Senegal. Todos los genomas mt comparten el mismo orden génico, que se ajusta al consenso publicado para Conidae, Neogastropoda y Caenogastropoda. Los análisis filogenéticos utilizando métodos probabilísticos recuperaron tres clados principales, cuya divergencia coincidió en el tiempo con cambios en el nivel del mar y las corrientes oceánicas, así como con variaciones de temperatura durante la crisis salina del Messiniense y la transición del Plioceno al Pleistoceno. Además, los tres clados se corresponden con distintos tipos de dientes radulares (robusto, pequeño y alargado), lo que sugiere que la especialización trófica podría ser otro factor promoviendo la diversificación de los conos endémicos de Senegal. La filogenia reconstruida mostró varias casos de convergencia fenotípica (especies crípticas) y cuestiona la validez de algunas especies (ecotipos o plasticidad fenotípica). Ambos resultados tienen importantes consecuencias taxonómicas y de conservación.

Molecular Phylogenetics and Evolution 112 (2017) 79–87

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Mitogenomic phylogeny of cone snails endemic to Senegal

Samuel Abalde [a], Manuel J. Tenorio [b], Carlos M.L. Afonso [c], Rafael Zardoya [a,*]

[a] Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain
[b] Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cadiz, 11510 Puerto Real, Cádiz, Spain
[c] Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

ARTICLE INFO

Article history:
Received 24 November 2016
Revised 21 April 2017
Accepted 22 April 2017
Available online 24 April 2017

Keywords:
Mitochondrial genomes
Radula
Conus
Lautoconus
Messinian Salinity Crisis
Plio-Pleistocene transition
Convergence

ABSTRACT

Cone snails attain in Senegal one of their highest peaks of species diversity throughout the continental coast of Western Africa. A total of 15 endemic species have been described, all placed in the genus *Lautoconus*. While there is ample data regarding the morphology of the shell and the radular tooth of these species, virtually nothing is known regarding the genetic diversity and phylogenetic relationships of one of the most endangered groups of cones. In this work, we determined the complete or near-complete (only lacking the control region) mitochondrial (mt) genomes of 17 specimens representing 11 endemic species (*Lautoconus belairensis*, *Lautoconus bruguieresi*, *Lautoconus cacao*, *Lautoconus cloveri*, *Lautoconus* cf. *echinophilus*, *Lautoconus guinaicus*, *Lautoconus hybridus*, *Lautoconus senegalensis*, *Lautoconus mercator*, *Lautoconus taslei*, and *Lautoconus unifasciatus*). We also sequenced the complete mt genome of *Lautoconus guanche* from the Canary Islands, which has been related to the cones endemic to Senegal. All mt genomes share the same gene arrangement, which conforms to the consensus reported for Conidae, Neogastropoda and Caenogastropoda. Phylogenetic analyses using probabilistic methods recovered three major lineages, whose divergence coincided in time with sea level and ocean current changes as well as temperature fluctuations during the Messinian salinity crisis and the Plio-Pleistocene transition. Furthermore, the three lineages corresponded to distinct types of radular tooth (robust, small, and elongated), suggesting that dietary specialization could be an additional evolutionary driver in the diversification of the cones endemic to Senegal. The reconstructed phylogeny showed several cases of phenotypic convergence (cryptic species) and questions the validity of some species (ecotypes or phenotypic plasticity), both results having important taxonomic and conservation consequences.

1. Introduction

Cone snails (Conidae, Caenogastropoda) represent the paradigm of a species rich clade of marine animals (Röckel et al., 1995; Duda and Kohn, 2005; Tucker and Tenorio, 2013; Kohn, 2014; Puillandre et al., 2014), and therefore are an excellent group for studying the evolutionary processes underlying biological diversification. The more than 800 described species of cone snails (WoRMS, accessed October 2016; Bouchet and Gofas, 2010) are widely found in all tropical and subtropical seas from intertidal zones to deep waters associated to rocky shores, coral reefs, and sandy bottoms, preying on marine worms, snails, and fishes (Tucker and Tenorio, 2013). Cone snails are best known for their harpoon-like radular teeth and for having one of the most sophisticated venom strategies of the animal kingdom (Olivera et al., 2012): within a specialized venom gland, cones produce a cocktail composed of small peptides named conotoxins with both predatory and defensive functions (Dutertre et al., 2014).

The species diversity of cones is highest in the Indo-West Pacific region (Röckel et al., 1995; Duda and Kohn, 2005; Puillandre et al., 2014), and consequently for many years, studies on ecology, natural history, and conotoxin diversity of cone snails focused on species from this area (e.g., Duda et al., 2001) to the detriment of others such as Western Africa (e.g., Monteiro et al., 2004; Cunha et al., 2005; Duda and Rolan, 2005) or the Western Atlantic (Kohn, 2014). Cone species in the Indo-West Pacific region attain maximum diversity in the tropics and tend to show relatively widespread distributions (Cunha et al., 2014). In contrast, studies focused on Western African cones have revealed high levels of endemicity and peaks of species diversity concentrated in subtropical areas around Senegal and Angola in the continent, and most prominently in the Cabo Verde archipelago, which may harbor about 10% of cone species diversity worldwide (Cunha et al., 2014). These remarkable differences in species richness distribution indicate that distinct diversification processes may be acting

* Corresponding author.
E-mail address: rafaz@mncn.csic.es (R. Zardoya).

http://dx.doi.org/10.1016/j.ympev.2017.04.020
1055-7903/

in both regions, provided that the observed patterns rely on a consistent estimation of the number of cone species in both areas. Thus far, however, species delimitation in cones has been mostly based on the shape, color, and banding patterns of the shell, which may show in some instances important levels of homoplasy (Duda et al., 2008). In many cases, it is unclear whether different shell morphotypes represent distinct species or different forms of the same species (Duda and Palumbi, 1999). Therefore, determination of genetic variation and inference of phylogenetic relationships are timely in this hyperdiverse group in order to discern in which instances conchology can be used as basis for the recognition of species (Duda et al., 2008).

In this study, we focus on cone snails endemic to Senegal (Monteiro et al., 2004). This country, and in particular the Cape Verde peninsula (now entirely occupied by the urban growth of Dakar) has one of the highest peaks of diversity of cone snails in the Western African coast (Cunha et al., 2014). Although the cones of Senegal were already known in the times of Linnaeus, it was not until recently that a comprehensive monograph was produced upon exhaustive sampling and detailed morphological comparisons (Pin and Leung Tack, 1995). A total of 11 different endemic species were identified (see Table 1), all belonging to the subgenus *Lautoconus* (Puillandre et al., 2015), which some authors have elevated to the generic status (Tucker and Tenorio, 2009); we follow herein the latter taxonomic proposal. Moreover, *Lautoconus taslei* from the Petite-Côte region of Senegal was not considered in the revision. Afterwards, *Lautoconus trencarti* (Nolf and Verstraeten, 2008), *Lautoconus tacomae* (Boyer and Pelorce, 2009) and *Lautoconus dorotheae* (Monnier and Limpalaër, 2010) that live in deeper waters were added to the list. Recently, *Lautoconus senegalensis* (Gulden et al., 2017) was described. It corresponds to specimens previously known as *Lautoconus* cf. *mediterraneus* from Senegal (Pin and Leung Tack, 1995). We will use the new name henceforth. Importantly, seven species of *Lautoconus* from Senegal are considered endangered and another three vulnerable according to the IUCN Red List (Peters et al., 2013). In addition, several non-endemic species are found in Senegal including *Genuanoconus genuanus*, *Kalloconus pulcher*, *Kalloconus byssinus*, *Monteiroconus tabidus*, *Monteiroconus ambiguus*, and the amphi-Atlantic *Chelyconus ermineus* (Pin and Leung Tack, 1995).

Thus far, several molecular phylogenies have been reported either for the family Conidae (Puillandre et al., 2014; Uribe et al., 2017) or focused in particular geographic regions including the Indo-West Pacific (Duda and Palumbi, 1999; Duda and Kohn, 2005), Cabo Verde archipelago (Cunha et al., 2005; Duda and Rolan, 2005; Cunha et al., 2008), Canary Islands (Cunha et al., 2014), Mozambique (Pereira et al., 2010) and Saint Helena Island (Tenorio et al., 2016), but none has studied in a comprehensive fashion the cone snails endemic to Senegal.

In this study, we used complete mitochondrial (mt) genomes, which have proven to be very useful in reconstructing relatively highly resolved phylogenies of different gastropod groups including Neogastropoda (Cunha et al., 2009) and in particular, Conidae (Uribe et al., 2017). At present, the complete or near-complete mt genomes of 13 species belonging to the family Conidae are publicly available. Here, we sequenced the complete or nearly complete mt genomes of 17 individuals representing different populations and species of *Lautoconus* endemic to Senegal. In addition, we sequenced the complete mt genome of *Lautoconus guanche* from the Canary Islands (also occurring from Northern Mauritania to Morocco), which is included in the same genus (Tucker and Tenorio, 2009). We aimed to: (1) reconstruct a robust phylogeny of cones endemic to Senegal; (2) study radular tooth evolution in the group; (3) provide a first genetic hypothesis of species delimitation within the group; and (4) date major events in the diversification of Senegal endemic cones.

2. Materials and methods

2.1. Samples and DNA extraction

The complete list of specimens analyzed in this study corresponding to different populations and species of *Lautoconus* from Senegal and the Canary Islands (Spain) is shown in Table 1, along with details on the respective sampling localities and museum vouchers. Specimens were collected by snorkel at 1–3 m depth, or picked by hand at low tide. All samples were stored in 100% ethanol, and total DNA was isolated from 5–10 mg of foot tissue following a standard phenol-chloroform extraction (Sambrook et al., 1989).

Table 1
New mitochondrial (mt) genomes analyzed in this study.

ID	Species	Location	Coverage		Length (bp)	GenBanc Acc. No	Voucher (MNCN/ADN)	Voucher (shell) (MNCN)
			No. reads	Mean depth				
1258	*Lautoconus mercator* (Linnaeus, 1758)	Les Almadies, Dakar, Senegal	193,127	1902	15,332	KY801864	91278	15.05/78419
1266	*Lautoconus hybridus* (Kiener, 1845)[a]	NGor, Dakar, Senegal	123,000	1208	15,507	KY801863	91279	15.05/78427
1278	*Lautoconus mercator* (Linnaeus, 1758)	NGor, Dakar, Senegal	180,207	1774	15,329	KY801862	91280	15.05/78439
1282	*Lautoconus guinaicus* (Hwass, 1792)	Ndayane, Senegal	163,070	1608	15,316	KY801861	91281	15.05/78443
1290	*Lautoconus unifasciatus* (Kiener, 1845)[a]	Ndayane, Senegal	144,388	1424	15,506	KY801860	91282	15.05/78451
1296	*Lautoconus cloveri* (Walls, 1978)	Ndayane, Senegal	67,122	659	15,323	KY801859	91283	15.05/78457
1301	*Lautoconus cacao* (Ferrario, 1983)	Ndayane, Senegal	179,361	1767	15,318	KY801858	91284	15.05/78462
1302	*Lautoconus cacao* (Ferrario, 1983)	Ndayane, Senegal	19,742	192	15,327	KY801857	91285	15.05/78463
1312	*Lautoconus senegalensis* (Gulden et al., 2017)	Ndayane, Senegal	200,965	1806	15,317	KY801856	91286	15.05/78473
1315	*Lautoconus taslei* (Kiener, 1845)	Joal-Fadiouth, Senegal	116,067	1202	15,314	KY801855	91287	15.05/78476
1321	*Lautoconus mercator* (Linnaeus, 1758)	Île de Gorée, Dakar, Senegal	107,938	1063	15,328	KY801854	91288	15.05/78482
1335	*Lautoconus guinaicus* (Hwass, 1792)[a]	Île de Gorée, Dakar, Senegal	124,435	1227	15,506	KY801853	91289	15.05/78496
1336	*Lautoconus bruguieresi* (Kiener, 1845)	Île de Gorée, Dakar, Senegal	86,860	910	15,318	KY801852	91290	15.05/78497
1338	*Lautoconus bruguieresi* (Kiener, 1845)	Île de Gorée, Dakar, Senegal	144,761	1424	15,340	KY801851	91291	15.05/78499
1341	*Lautoconus* cf. *echinophilus* (Petuch, 1975)	Île de Gorée, Dakar, Senegal	219,541	2164	15,319	KY801850	91292	15.05/78502
1343	*Lautoconus belairensis* (Pin and Leung Tack, 1989)	Terrou-Bi, Dakar, Senegal	74,962	738	15,321	KY801849	91293	15.05/78504
1350	*Lautoconus guinaicus* (Hwass, 1792)	Terrou-Bi, Dakar, Senegal	148,308	1456	15,323	KY801848	91294	15.05/78511
CG13	*Lautoconus guanche* (Lauer, 1993)[a]	Lanzarote, Canary Islands, Spain	318,448	2915	15,506	KY801847	91295	–

[a] Complete.

2.2. Radular tooth preparation

The radular sac was dissected from the cone snail and soft parts were digested in concentrated aqueous potassium hydroxide for 24 hours. The resulting mixture was then placed in a petri dish and examined with a binocular microscope. The entire radula was removed with fine tweezers and rinsed with distilled water, then mounted on a slide using Aquatex (Merck, Germany) mounting medium, and examined under a compound microscope. Photographs were obtained with a charge-coupled device (CCD) camera attached to the microscope. Terminology for radular morphology follows Tucker and Tenorio (2009), with abbreviations following Kohn et al. (1999).

2.3. PCR amplification and sequencing

Complete or near-complete (without the control region; see results) mt genomes were amplified through a combination of standard and long PCRs using the primers and following the protocols of Uribe et al. (2017). Standard-PCR products were sequenced using Sanger technology. Long-PCR products were subjected to massive parallel sequencing. Briefly, products were purified by ethanol precipitation and amplified fragments from the same mt genome were pooled together in equimolar concentrations. For each mt genome a separate indexed library was constructed using the NEXTERA XT DNA library prep kit (Illumina, San Diego, CA, USA) and run in an Illumina MiSeq platform (v.2 chemistry; 2×150 paired-end) at Sistemas Genómicos (Valencia, Spain).

2.4. Genome assembly and annotation

The reads corresponding to the different PCR amplified mt genomes were sorted using the corresponding library indices, and assembly of the different mt genomes was performed in the TRUFA webserver (Kornobis et al., 2015). Briefly, adapters were removed using SeqPrep (StJohn, 2011), quality of the reads was checked using FastQC v.0.10.1 (Andrews, 2010), and raw sequences were trimmed and filtered out according to their quality scores using PRINSEQ v.0.20.3 (Schmieder and Edwards, 2011). Filtered reads were used for *de novo* assembly of each mt genome using default settings (minimum contig length: 200; sequence identity threshold: 0.95) of Trinity r2012-06-08 (Grabherr et al., 2011) in TRUFA, and only retaining contigs with a minimum length of 3 kb. These contigs were finally overlapped in Sequencher 5.0.1 to render the different complete or nearly complete mt genomes included within each index. In order to estimate mean coverage, each assembled mt genome was used as a reference to map the original (raw) reads with a minimum identity of 100% using Geneious® 8.0.3.

The newly determined mt genomes were annotated with the MITOS webserver (Bernt et al., 2013) using the *Africonus borgesi* mt genome (Cunha et al., 2009) as a reference. Annotations of the 13mt protein-coding genes were corroborated manually identifying the corresponding open reading frames using the invertebrate mitochondrial code. The transfer RNA (tRNA) genes were further identified with tRNAscan-SE 1.21 (Schattner et al., 2005), which infer cloverleaf secondary structures (with a few exceptions that were determined manually). The ribosomal RNA (rRNA) genes were identified by sequence comparison with the *A. borgesi* mt genome (Cunha et al., 2009), and assumed to extend to the boundaries of adjacent genes (Boore et al., 2005). GenBank accession numbers of each mt genome are provided in Table 1.

2.5. Sequence alignment and phylogenetic analyses

The newly sequenced complete or nearly complete mt genomes were aligned with the mt genomes of *A. borgesi* (Cunha et al., 2009) from Boa Vista Island, Republic of Cabo Verde, and *Lautoconus ventricosus* (Uribe et al., 2017) from Faro, Portugal, which were used as outgroup taxa based on the phylogeny reported by Puillandre et al. (2014). A sequence data set was constructed concatenating the nucleotide sequences of the 13 mt protein-coding and two rRNA genes. The deduced amino acid sequences of the 13 mt protein-coding genes were aligned separately and used to guide the alignment of the corresponding nucleotide sequences with Translator X (Abascal et al., 2010). Nucleotide sequences of the mt rRNA genes were aligned separately using MAFFT v7 (Katoh and Standley, 2013) with default parameters. Ambiguously aligned positions were removed using Gblocks, v.0.91b (Castresana, 2000) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: No. Finally, the different single alignments were concatenated using Geneious® 8.0.3. Sequences where format converted for further analyses using the ALTER webserver (Glez-Peña et al., 2010).

Phylogenetic relationships were inferred using maximum likelihood (ML, Felsenstein, 1981) and Bayesian inference (BI, Huelsenbeck and Ronquist, 2001). For ML, we used RAxML v8.1.16 (Stamatakis, 2006) with the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates (BP). BI analyses were conducted with MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003), running four simultaneous Markov chains for 10 million generation, sampling every 1000 generations, and discarding the first 25% generations as burn-in (as judged by plots of ML scores and low SD of split frequencies) to prevent sampling before reaching stationarity. Two independent Bayesian inference runs were performed to increase the chance of adequate mixing of the Markov chains and to increase the chance of detecting failure to converge, as determined using Tracer v1.6 (Rambaut and Drummond, 2007). The effective sample size (ESS) of all parameters was above 200. Node support was assessed based on Bayesian Posterior Probabilities (BPP). A node was considered highly supported with BP and BPP values above 70% and 0.95, respectively.

The best partition schemes and best-fit models of substitution for the data set were identified using Partition Finder (Lanfear et al., 2012) with the Akaike information criterion (AIC, Akaike, 1973). For the protein-coding genes, the partitions tested were: all genes grouped; all genes separated (except *atp6-atp8* and *nad4-nad4L*); and genes grouped by subunits (*atp*, *cob, cox*, and *nad*). In addition, these three partitions schemes were tested taking into account separately the three codon positions. The rRNA genes were tested with two different schemes, genes separated or combined.

2.6. Estimation of divergence times

The program BEAST v.1.7 (Drummond and Rambaut, 2007) was used to perform a Bayesian estimation of divergence times. An uncorrelated relaxed molecular clock was used to infer branch lengths and nodal ages. The tree topology was fixed using the one recovered by the ML and BI analyses. For the clock model, the lognormal relaxed-clock model was selected, which allows rates to vary among branches without any a priori assumption of autocorrelation between adjacent branches. For the tree prior, a Yule process of speciation was employed. Only the protein-coding genes were used. The partitions and models selected by Partition Finder were applied (see results; except for second codon positions that we used the HKY+I model instead of the GTR+I model, as the latter could not converge). The final Markov chain was run twice for 100 million generations, sampling every 10,000 generations, and the first 1000 trees were discarded as part of the burn-in process, according to the convergence of chains

checked with Tracer v.1.5. (Rambaut and Drummond, 2007). The ESS of all parameters was above 200.

Despite the fact that there are many fossils of Conidae, it is difficult in many instances to be certain about species identifications given the important levels of homoplasy in shell shape (Duda et al., 2008). Hence, although there are fossils attributed to *L. ventricosus* (Sacco, 1893) and *L. mercator* (Glibert, 1960), we opted to calibrate the clock using a biogeographical event. The posterior distribution of the estimated divergence times was obtained by specifying one calibration point as prior for the divergence time of the split between *L. ventricosus* and *A. borgesi* in the outgroup. The latter species is endemic to Boa Vista, and we used the age of formation of this island (16.5 Mya; Dyhr and Holm, 2010) as biogeographical calibration point. We applied a log-normal distribution as the prior model for the calibration and enforced the median divergence time to equal 16.5 (s.d. = 0.05, offset = 0.5).

3. Results

3.1. Sequencing and assembly

The nucleotide sequences for the mt genomes of *Lautoconus hybridus*, three *Lautoconus guinaicus, Lautoconus unifasciatus, Lautoconus belairensis*, and *L. guanche* were determined to be complete whereas those of three *Lautoconus mercator*, two *Lautoconus bruguieresi*, two *Lautoconus cacao*, *Lautoconus cloveri, Lautoconus senegalensis, Lautoconus* cf. *echinophilus*, and *L. taslei* lacked the *trnF* gene, the control region, and the start of the *cox3* gene because the corresponding fragment could not be PCR amplified (Fig. 1). The number of reads, mean coverage, and length of each mt genome are provided in Table 1. The mt genomes of *L. cacao* from Ndayane and *L. guanche* received the minimum (19,742) and maximum (318,448) number of reads, respectively. The same samples received the minimum (192×) and maximum (2915×) coverage, respectively (Table 1).

3.2. Genome organization and sequence divergence

All sequenced *Lautoconus* mt genomes encode for 13 protein-coding, 2 rRNA and 22 tRNA genes (but note that the *trnF* gene could not be determined in the incomplete mt genomes; see Appendix A). They all share the same genome organization: the major strand encodes all genes, except those forming the cluster MYCWQGE (*trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, *trnE*) and the *trnT* gene (Fig. 1). The genes *nad4*/*nad4L* overlapped in seven nucleotides in all the mt genomes. All protein-coding genes start with ATG (but note that the beginning of *cox3* could not be determined in the incomplete mt genomes; see Appendix A). The stop codons were variable between TAA and TAG, depending on the gene and the species. There were six genes (*cox1*, *cox3*, *nad1*, *nad3*, *nad5* and *nad6*) whose stop codon is TAA. Two genes (*cob* and *nad4L*) ended with TAG. Two genes (*nad2* and *nad4*) had incomplete stop codons (TA-) that become functional after polyadenylation (Chang and Tong, 2012). The remaining three genes (*atp6*, *atp8*, and *cox2*) varied in their stop codon depending on the species (see Appendix A).

Pairwise uncorrected sequence divergences between *L.* cf. *echinophilus*, *L. bruguieresi*, *L. cloveri*, *L. mercator* from Les Almadies and Gorée Island, *L. senegalensis*, and *L. cacao* from Ndayane (specimen #1302) *versus* the remaining analyzed samples averaged 4% (see Appendix A). Pairwise uncorrected sequence divergences between *L. belairensis*, *L. cacao* from Ndayane (specimen #1301) and *L. mercator* from NGor *versus L. unifasciatus*, *L. guanche*, *L. guinaicus*, *L. hybridus*, and *L. taslei* averaged 2.6%. The pairwise uncorrected sequence divergences between *L. guanche versus L. guinaicus*, *L. hybridus*, and *L. taslei* varied 0.4–0.5%. Pairwise uncorrected sequence comparisons (1) between *L.* cf. *echinophilus* and *L. bruguieresi*; (2) among *L. mercator* of Gorée Island, *L. senegalensis*, and *L. cacao* from Ndayane (specimen #1302); (3) among *L. taslei*, and *L. guinaicus* from Terrou-Bi and Ndayane; and (4) between *L. hybridus* and *L. guinaicus* from Gorée Island showed almost no sequence divergence (<0.1%). The mt genomes of *L. bruguieresi* from Gorée Island and Les Almadies had exactly the same sequence (see Appendix A).

3.3. Phylogenetic relationships within Lautoconus

Phylogenetic relationships of cones endemic to Senegal were reconstructed based on the nucleotide sequences of the concatenated 13 mt protein-coding and two rRNA genes using probabilistic methods. The final matrix was 13,582 positions long. According to the AIC, the best partition scheme for the protein-coding genes was the one combining all these genes but analyzing each codon position separately. The best substitution models were GTR+I+G for the first codon position, GTR+I for the second position, and GTR+G for the third codon position. For the rRNA genes, the best scheme had both genes combined under the GTR+I model. Both, ML ($-lnL = 31{,}523.23$) and BI ($-\ln L = 31{,}541.77$ for run 1; $-lnL = 31{,}543.66$ for run 2) arrived at identical topology (Fig. 2), with only slightly differences in branch lengths. All nodes in the reconstructed phylogeny received high statistical support (Fig. 2).

Up to three major lineages could be distinguished within the reconstructed phylogeny (Fig. 2). One lineage included different populations of *L. bruguieresi* and *L.* cf. *echinophilus* sister to a clade that included *L. cloveri* from Ndayane and a group with specimens of *L. mercator* from Gorée Island and Les Almadies, *L. senegalensis* from Ndayane, and one specimen (#1302) of *L. cacao* from Ndayane (Fig. 2). A second lineage included *L. belairensis* from Terrou-Bi sister to *L. mercator* from NGor and another specimen (#1301) of *L. cacao* from Ndayane (Fig. 2). This lineage was sister to another one including the following clades: (1) *L. guinaicus* from Ndayane and Terrou-Bi and *L. taslei* from Joal-Fadiouth sister to (2) *L. hybridus* from NGor and *L. guinaicus* from Gorée Island to the exclusion of (3) *L. unifasciatus* from Ndayane and *L. guanche* from Canary Islands (Fig. 2).

The above-mentioned three major *Lautoconus* lineages corresponded to three distinct types (robust, small, and elongated, respectively) of radular tooth (Fig. 2), all of them of the vermivorous kind. The robust and elongated types were of medium relative size (Shell Length/Tooth Length = 37–42), with a short, pointed barb. They differed in the size of the anterior section of the tooth, which is equal or slightly shorter than the posterior section of the

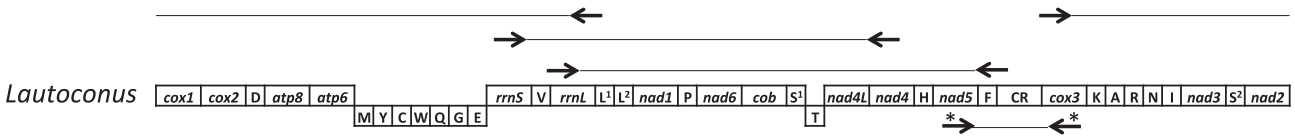


Fig. 1. Mitochondrial gene order of *Lautoconus*. All newly determined mt genomes shared identical genome organization. The genes encoded in the major and minor strands are shown in the top and bottom lines, respectively. The relative position of the primers for long PCR is shown. The two primers labeled with an asterisk were designed to amplify the control region, and failed in several cases rendering incomplete mt genomes.

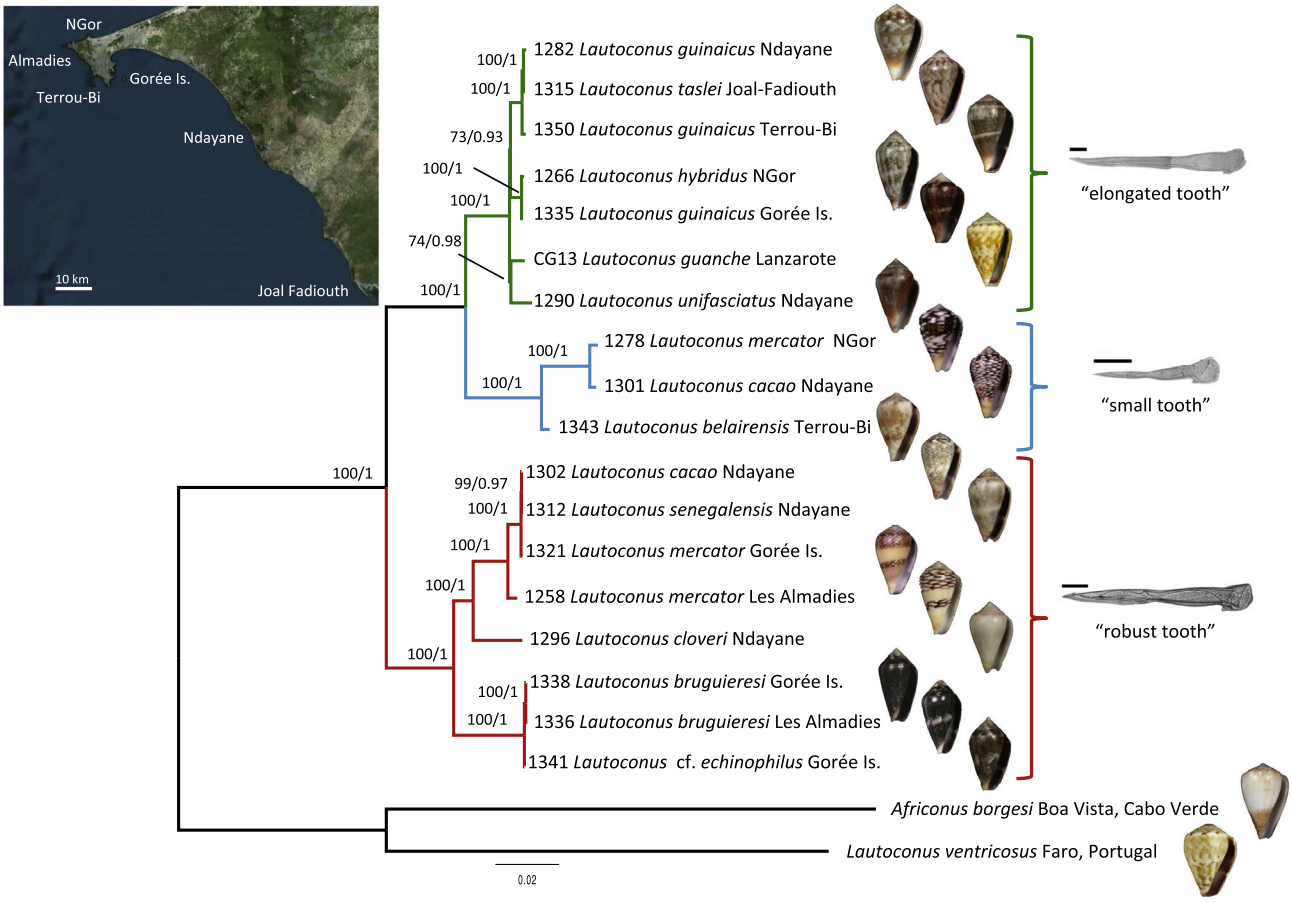


Fig. 2. Phylogenetic relationships of *Lautoconus* based on complete mt genomes (concatenated protein coding plus rRNA genes analyzed at the nucleotide level). The reconstructed ML phylogram using *A. borgesi* from Boa Vista, Cabo Verde archipelago, and *L. ventricosus* from Faro, Portugal as outgroup taxa is shown. Number of specimen, initial species assignment, locality, and a ventral picture of the shell are provided. Three major clades are indicated with different colors (red, blue, and green) that corresponded to three different types of radular tooth: robust, small and elongated, respectively (scale bar equals 0.1 mm; see all radular teeth in Appendix A). A map of Senegal as inset is provided with sampling localities. Numbers at nodes are statistical support values for ML (bootstrap proportions)/BI (posterior probabilities). Scale bar indicates substitutions/site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tooth for the robust type (Tooth Length/Anterior Portion Length = 2.0–2.1), but much longer for the elongated type (Tooth Length/Anterior Portion Length = 1.6–1.7). They also differed in the extension of the anterior portion covered by the blade (80–85% in the robust type *versus* 40–46% in the elongated type) and the number of denticles present in the serration (19–30 in the robust type, but more than 40 in the elongated type). The radular tooth of the small type had, as the name indicates, a much smaller relative size (Shell Length/Tooth Length = 85–96). In this case, the anterior section of the tooth was significantly shorter than the posterior section (Tooth Length/Anterior Portion Length = 2.3–2.5). The blade covered 64–70% of the anterior portion, and there were 15–20 small denticles in the serration arranged on a single row. The base of the tooth was large and broad (see Appendix A).

3.4. Divergence times

Major cladogenetic events within *Lautoconus* were dated using an uncorrelated relaxed molecular clock model, which was calibrated using the age of Boa Vista Island at the divergence of *A. borgesi* and *L. ventricosus*. The first event of diversification within the crown group of *Lautoconus* was estimated at a mean of 5.6 (4.5–7.0, credible interval) million years ago (Mya) separating the species with robust radular teeth from the ancestor of the species with elongated and small radular teeth (Fig. 3). The second main split was estimated to occur about 2.8–2.9 (2.1–3.8) Mya separating (1) cones with elongated radular teeth from those with small radular teeth and (2) *L. bruguieresi* and allies from the remaining cones with robust radular teeth (Fig. 3). Finally, it is interesting to note that main speciation events occurred between 0.7–2.0 (0.5–2.7) Mya whereas population divergences were dated about 0.03–0.19 Mya (see discussion regarding species boundaries; Fig. 3).

4. Discussion

The region around Dakar is considered a hotspot of diversity for cone snails off the Western Africa continent (Cunha et al., 2014). However, rapid expansion of the metropolitan area is seriously compromising the conservation of this extraordinary species richness mainly due to the loss of adequate habitats (e.g., the neighborhood of Bel-Air has a rocky plateau now heavily polluted by the adjacent Port of Dakar. Bel-Air is the type locality of *L. belairensis*, which is not found there anymore). In fact, all 14 cone species included in the highest categories of Critically Endangered and Endangered in the IUCN Red List are endemic to either Senegal or the Cabo Verde archipelago (Peters et al., 2013). Therefore, it is urgent to have a better understanding of the exact number of cone species and their genetic diversity in the region. Moreover, any further comparative analysis aimed at understanding the evo-

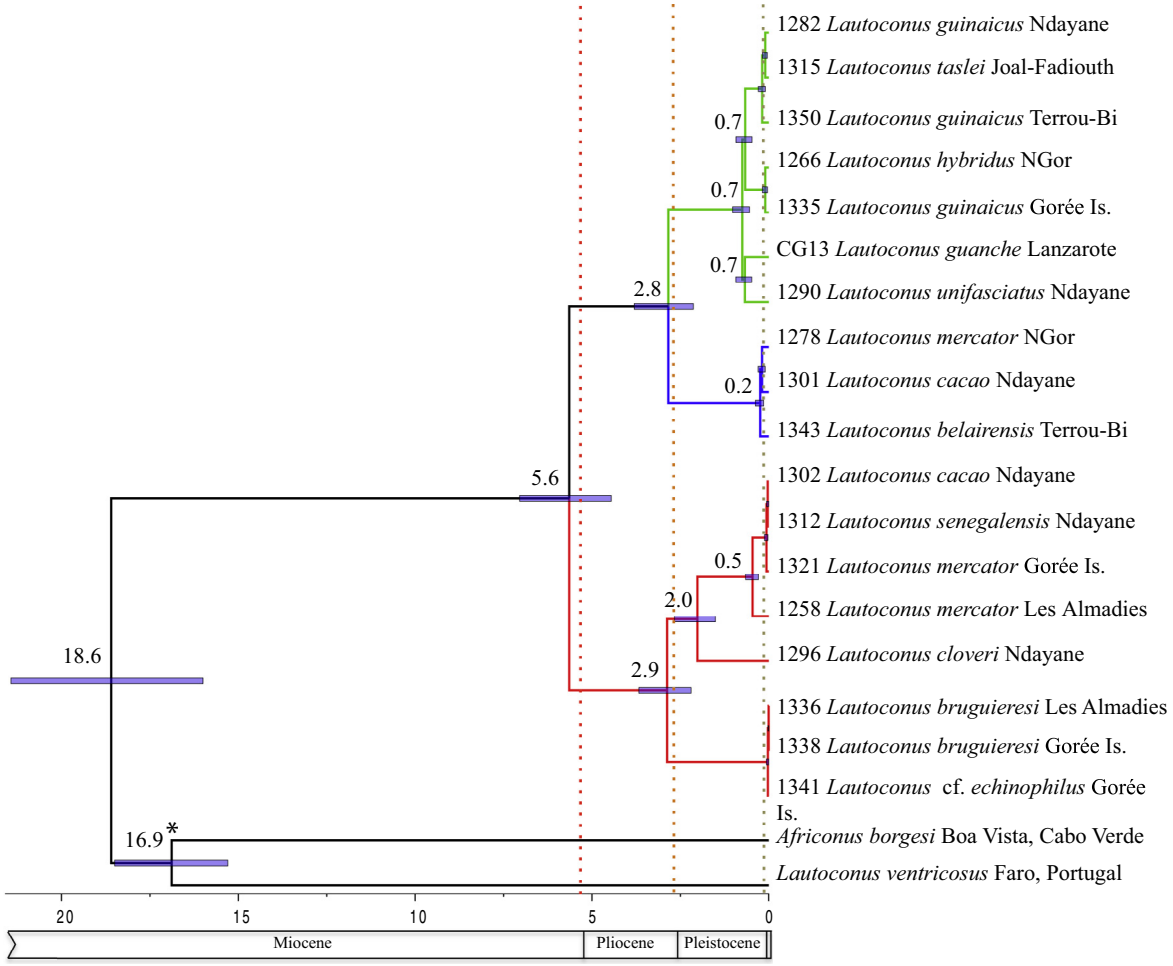


Fig. 3. Chronogram of *Lautoconus* based on complete mt genomes (concatenated protein coding plus rRNA genes analyzed at the nucleotide level) and using the fixed topology of the ML tree shown in Fig. 2. A Bayesian uncorrelated relaxed lognormal clock with a fossil/geographic-based calibration prior (denoted by an asterisk) was used in BEAST. Horizontal bars represent 95% credible intervals for time estimates; dates are in millions of years. The dotted lines represent the age of the Messinian Salinity Crisis (red), the Plio-Pleistocene transition (orange), and the beginning of the Holocene (brown), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lutionary process underlying cone diversification in this area needs to be based on a robust phylogeny (Harvey and Pagel, 1991). This primary information was lacking for the cones endemic to Senegal.

Here, we amplified and sequenced complete or near-complete (without the control region) mt genomes of 18 specimens initially assigned to 10 out of the 14 species of genus *Lautoconus* described as endemic to Senegal plus *L. guanche* from Canary Islands. We had no individuals from the four species living in deeper waters (*L. dorotheae*, *L. pineaui*, *L. tacomae* and *L. trencarti*), although *cox1* partial sequences are available for the first two species (Puillandre et al., 2014). The newly determined mt genomes share identical gene arrangement, which conforms to the consensus genome organization reported for Conidae (Uribe et al., 2017), Neogastropoda (Cunha et al., 2009), and Caenogastropoda (Osca et al., 2015).

The reconstructed phylogenies using probabilistic methods were fully resolved (all nodes had high statistical support), allowing several further evolutionary inferences. The first striking outcome of our analyses is that cones endemic to Senegal do not form a monophyletic group because *L. guanche* from Canary Islands is recovered as sister to *L. unifasciatus*, nested deep within the reconstructed phylogeny. Its close phylogenetic relationship to cones endemic to Senegal was already hypothesized based on radular tooth comparisons, which prompted the inclusion of the species into the genus *Lautoconus* (Tucker and Tenorio, 2009). This species was also placed along with Angola and Senegal cone endemics in the *cox1* phylogeny of Puillandre et al. (2014), although its exact phylogenetic position could not be fully resolved. Even though *L. guanche* has non-planktonic development (see Appendix A; Cunha et al., 2014), and thus a supposedly limited capacity of dispersal, this species is not endemic to Canary Islands but also found in the Western African coast from Northern Mauritania to Morocco. Phylogenetic studies focused on *L. guanche* showed no differentiation of the species within the Canary archipelago or between the islands and the Western African coast indicating recurrent gene flow (Cunha et al., 2014), which could result from seasonal changes in the direction of the Canary current (Stramma and Siedler, 1988; Navarro-Pérez and Barton, 2001). These findings indicate clearly that the recently introduced taxon *Lautoconus saharicus* (Petuch and Berschauer, 2016) represents the local form of *L. guanche* from Dahkla Bay, Western Sahara. It lacks taxonomical value and must be considered a junior synonym.

The cones endemic to Senegal could be grouped into three main lineages according to our phylogenetic analyses. The members within each lineage share a distinct radular tooth (robust, small, and elongated, respectively). Hence, the type of radular tooth could be used as predictor of the relative position of the missing taxa in our phylogeny. Accordingly, *L. dorotheae*, *L. pineaui*, and *L. trencarti* could be placed within the "elongated tooth" clade, as suggested

by their radular tooth shape (Pin and Leung Tack, 1995; Boyer and Pelorce, 2009; Monnier and Limpalaër, 2010; see Appendix A). Likewise, the radular tooth of *L. tacomae* (Boyer and Pelorce, 2009; see Appendix A) is of the robust type, similar to that of *L. bruguieresi*; we could therefore assume that these two species are closely related.

The inferred chronogram suggests that early diversification in *Lautoconus* was driven by profound changes in the paleoecosystems of continental Western Africa as the two main divergence events in the group were related to drastic shifts in past global climate. The first divergence episode was dated during the Messinian Salinity Crisis (MSC), when the Mediterranean Sea desiccated at the end of the Miocene from 5.96 to 5.33 Mya (Krijgsman et al., 1999), whereas the second one corresponded to the transition between the Pliocene and the Pleistocene about 2.8 Mya. Plate tectonic movements are believed to be responsible of the onset of the MSC (Duggen et al., 2003), but simultaneously there was a glacial period that lasted from 6.26 to 5.50 Mya, and produced an eustatic sea level drop between −10 and −30 m (Hodell et al., 2001). The closure of marine gateways between the Atlantic Ocean and the Mediterranean Sea had likely a great influence on the Canary current and the Atlantic meridional overturning circulation (Ivanovic et al., 2014). Altogether, the glacially driven eustatic sea level changes combined with modifications in surface and deep ocean currents could have seriously affected dispersal patterns in *Lautoconus* given the non-planktonic nature of their larvae (see Appendix A), thus promoting instances of isolation in refugia, restricted gene flow, and the consequent diversification processes. The transition of the Pliocene to the Pleistocene was associated to a shift from a warm to a cold climate, and resulted in oceans completely different in terms of circulation (Filippelli and Flores, 2009). Moreover, this transition marked a pronounced change from a somewhat stable pattern in sea levels to the onset of extreme oscillations concurring with glacial-interglacial periods (Lisiecki and Raymo, 2005). This combination of changes in the sea realm during the Plio-Pleistocene boundary and afterwards could have been determinant for diversification of *Lautoconus* following population expansion-contraction cycles, as has been suggested for other marine organisms (Marko et al., 2010; Shen et al., 2011).

For many years, allopatry was not considered a predominant mode of speciation in the sea due to the high potential for dispersal of many marine organisms (adults and larvae), and the general lack of geographical barriers in the marine realm. However, this view has considerably changed in the recent years due to genetic studies revealing different factors limiting gene flow in the sea, which ultimately provoke population subdivision and genetic differentiation (Palumbi, 1994; Williams and Reid, 2004; Shen et al., 2011). Our results support that allopatry is the main mode of speciation for cone snails having non-planktotrophic larvae, as is the case of *Lautoconus* species (this study) and as it has been suggested for *Africonus* and *Trovaoconus* of the Cabo Verde archipelago, although for these two genera, low sea levels during glacial maxima promoted connection between islands (i.e., gene flow) and posterior sea level rises induced isolation (Cunha et al., 2005; Cunha et al., 2008).

It has been shown that vermivorous, molluscivorous, and piscivorous cone snails have distinct types of radular tooth (Duda et al., 2001). As the type of radular tooth closely correlated with the earliest cladogenetic events in the reconstructed phylogeny, it could be argued that ecological (dietary) adaptation was an additional evolutionary process triggering diversification in the cones endemic to Senegal. It is plausible that ecosystem changes during the MSC and the Plio-Pleistocene transition also enhanced diversification in other marine groups, and polychaetes in particular, triggering trophic specializations in *Lautoconus*. However, within vermivore cone snails, radular tooth specialization has been only clearly documented for those species such as e.g., *Stephanoconus regius* that prey on amphinomids (Duda et al., 2001). Hence, more ecological studies are needed to assess whether a particular tooth shape is directly related to the types of worms that are eaten by a given species or group of species.

Finally, it is important to note that within each of the three main lineages, the different species occupy complementary geographic distributions that cover as a whole all suitable habitats available in the territory (i.e., the different locations in the Cape Verde peninsula, Gorée Island, Ndayane and Joal-Fadiouth). Hence, niche segregation might be a final evolutionary process also contributing to the diversification of the group. Patterns of speciation have been predominantly documented in adaptive radiations, which show how the combination of different evolutionary processes acting successively can promote increasing levels of diversification in relatively short periods of time (Danley and Kocher, 2001; Streelman et al., 2002; Rüber et al., 2003). In addition to the evolutionary processes (vicariance, trophic specialization, niche segregation) that could have been important in generating the diversity of cone snails endemic to Senegal (as here inferred), others have been proposed in the case of adaptive radiations such as ecomorphological and behavioral adaptations, which still need to be investigated in detail for cone snails.

Species delimitation in cone snails has been traditionally based on shell shape and color banding patterns (Tucker and Tenorio, 2013), largely ignoring genetic data. However, shell morphology of snails could be in many cases convergent, reflecting adaptation of genetically distinct populations (ecotypes) or species (sibling or cryptic) to local environments (Knowlton, 1993; Hollander and Butlin, 2010; Dowle et al., 2015). Moreover, phenotypic variation may result not only from genetic differences but also from phenotypic plasticity i.e., the capacity of one genotype to generate different phenotypes in response to distinct environments (Hollander and Butlin, 2010; Dowle et al., 2015). Convergence and phenotypic plasticity may confound taxonomists and could result in under- and overestimations of the number of species in a group, respectively. Hence, the need of detecting such evolutionary processes using robust phylogenies and the comparative method (Harvey and Pagel, 1991), as part of a multidisciplinary approach to species delimitation.

In the case of the cones endemic to Senegal, the history of their taxonomy already reflects controversial decisions regarding the species status of some of the taxa. For instance, some authors considered *L. cacao* a synonym of *L. mercator*, and others proposed that *L. echinophilus* could be a juvenile of *L. bruguieresi* (Pin and Leung Tack, 1995). Moreover, the phenotypic similarity of *L. senegalensis* (Senegal) and *L. ventricosus* (Mediterranean Sea) despite their disjoint geographic distribution is also striking, and calls for a case of potential cryptic species (Bandel and Wils, 1977). According to the reconstructed phylogeny, it is possible to detect two clear cases of phenotypic convergence: (1) the specimen #1335 from Gorée Island initially identified as *L. guinaicus* was recovered as sister to *L. hybridus*, and thus should belong to this latter species. Given the striking shell similarity of specimen #1335 to *L. guinaicus* and in order to discard potential contamination, we sequenced the universal *cox1* gene fragment (Folmer et al., 1994) of extra specimens of *L. guinaicus* from Gorée Island confirming this result (not shown); (2) the different specimens initially identified as *L. mercator* and *L. cacao* were distributed in two distinct clades. The taxonomic implications in this case are more complex (see below).

In contrast, several specimens initially attributed to different species, showed little (<0.1%) or no genetic divergence at all (the time tree dates these divergences less than 100 k years ago). These were the cases of: (1) *L. taslei* and two populations (Terrou-Bi and Ndayane) of *L. guinaicus*; thus, despite its disjoint distribution (southern coast of Senegal), *L. taslei* cannot be considered a valid species (see Appendix A). (2) *L. mercator* from NGor and one spec-

imen (#1301) of *L. cacao* from Ndayane; (3) *L. mercator* from Gorée Island (and possibly from Les Almadies), *L. senegalensis* from Ndayane, and one specimen (#1302) of *L. cacao* from Ndayane; (4) *L.* cf. *echinophilus* and two populations (Les Almadies and Gorée Island) of *L. bruguieresi*. There are several alternative explanations for this pattern including phenotypic plasticity, recent speciation, and mtDNA introgression. Unfortunately, it is not possible to distinguish between them without having the nuclear counterpart. However, for some of these cases there are some independent lines of evidence that could help resolving the conundrum. As mentioned above, our results would be in favor of considering that *L. echinophilus* is a form of *L. bruguieresi* (see Appendix A). However, it is important to note that the specimen that was sequenced in this study was identified as *L.* cf. *echinophilus* as it did not fully match the description of the type specimen of the species. Hence, new specimens ascribed to this species need to be sequenced to resolve the taxonomic status of this species. In the case of *L. cacao* and *L. mercator*, taxonomic problems of synonymy are mixed with evolutionary convergence. The shell morphology of the population of *L. mercator* from Gorée Island matches that of the lectotype of *Conus mercator* L., 1758, whereas specimen #1302 of *L. cacao* from Ndayane compares well with the lectotype of *Conus cacao* Ferrario, 1983 (see Appendix A). According to our results, *L. senegalensis* is a form of *L. cacao*, which in turn becomes a junior synonym of *L. mercator*. On the other hand, specimen #1301 of *L. cacao* from Ndayane was initially considered a juvenile of this species, but the phylogeny indicates that it is actually not related to *L. cacao* (= *mercator*). The shell morphology of this population matches that of the lectotype of *Conus reticulatus* (Born, 1778), which was traditionally considered a junior synonym of *Conus mercator* (see Appendix A). Our data suggest that the taxon *Lautoconus reticulatus* is actually a valid species, distinct from *L. mercator*, and must be reinstated as such. The name *L. reticulatus* (Born, 1778) should be also used for the population of *mercator*-like specimens from NGor. Moreover, *L. belairensis* also stands as a valid species, distinct from *L. mercator* and sister to *L. reticulatus*. Finally, the population of *L. mercator* from Les Almadies is in the limit of sequence divergence that could be associated to speciation events, assuming as threshold for the species status that *L. guanche* from Canary Islands is a valid species given its allopatric geographic distribution, and that it has 0.4–0.5% sequence divergence to closely related endemic species from Senegal in the same clade (more than half My of independent evolution in the chronogram). This threshold lies well within the so-called grey zone of speciation between 0.5–2% (Roux et al., 2016).

5. Conclusions

We reconstructed a robust phylogeny of cone snails endemic to Senegal (genus *Lautoconus*) using complete or near-complete mt genomes. The dating of this phylogeny revealed that major changes in the marine realm during the MSC and the Plio-Pleistocene transition could have produced vicariant events promoting diversification in these cone snails, which have a non-planktonic larval stage. Diversification was further accompanied by radular tooth specializations (which may correlate with dietary adaptations), and followed by speciation in allopatry (an extreme case would be *L. guanche*, whose distribution does not overlap with any cone endemic to Senegal). The reconstructed phylogeny together with sequence divergence (uncorrected *p* distances) data helped recognizing instances of shell convergence and questioned the validity of some species given their little genetic differentiation, although this result needs to be fully validated by sequencing nuclear genes. In any case, our study calls for a thorough revision of species delimitation in the family Conidae using genetic data. Moreover, by providing an evolutionary framework, the results here obtained are particularly important for designing a conservation strategy for Western African cone snails that face serious threats.

Acknowledgements

We are indebted to Amadou Gaye and Pier Luigi Tamagnini in Senegal and to Francisco Sicilia in Lanzarote, Canary Islands for their valuable help during sampling. We thank William J. Fenzan and two anonymous reviewers for their insightful comments on a previous version of the manuscript. We are grateful to Jesús Marco and Luis Cabellos, who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing phylogenetic analyses, as well as to Anita Eschner (Naturhistorisches Museum Wien, Austria), Virginie Héros (Muséum National d'Histoire Naturelle, Paris, France) and the Linnean Society of London for the photos of type material hosted at their respective institutions. Special thanks to Emilio Rolán (Vigo, Spain) and Yuri Kantor (Russian Academy of Sciences, Moscow, Russia) for supplying images of some radular teeth and protoconchs. This work was supported by the Spanish Ministry of Science and Innovation, Spain (CGL2013-45211-C2-2-P to RZ; BES-2014-069575 to SA).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.04.020.

References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38, W7–W13.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), 2nd International symposium on information theory akademiai Kiado. Budapest, Hungary, pp. 267–281.

Andrews, S., 2010. FastQC. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bandel, K., Wils, E., 1977. On *Conus mediterraneus* and *Conus guinaicus*. Basteria 41, 33–45.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319.

Boore, J.L., Macey, J.R., Medina, M., 2005. Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol. 395, 311–348.

Bouchet, P., Gofas, S., 2010. Conidae Fleming, 1822. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://marinespecies.org/aphia.php?p=taxdetails&id=14107 on 2016-10-21.

Boyer, F., Pelorce, J., 2009. Description d'un nouveau *Conus* (Gastropoda: Conidae) du Sénégal dans le groupe *Conus mediterraneus*". Novapex 10, 25–32.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Chang, J.H., Tong, L., 2012. Mitochondrial poly(A) polymerase and polyadenylation. Biochim. Biophys. Acta. 1819, 992–997.

Cunha, R., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evol. Biol. 9, 210.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde Islands. Syst. Biol. 54, 634–650.

Cunha, R.L., Lima, F.P., Tenorio, M.J., Ramos, A.A., Castilho, R., Williams, S.T., 2014. Evolution at a different pace: distinctive phylogenetic patterns of cone snails from two ancient oceanic archipelagos. Syst. Biol. 63, 971–987.

Cunha, R.L., Tenorio, M.J., Afonso, C., Castilho, R., Zardoya, R., 2008. Replaying the tape: recurring biogeographical patterns in Cape Verde *Conus* after 12 million years. Mol. Ecol. 17, 885–901.

Danley, P.D., Kocher, T.D., 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. Mol. Ecol. 10, 1075–1086.

Dowle, E.J., Morgan-Richards, M., Brescia, F., Trewick, S.A., 2015. Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. Mol. Ecol. 24, 4205–4221.

Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Duda, T.F., Bolina, M.B., Meyer, C.P., Kohn, A.J., 2008. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. Mol. Phylogenet. Evol. 49, 867–876.
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.
Duda, T.F., Kohn, A.J., Palumbi, S.R., 2001. Origins of diverse feeding ecologies within Conus, a genus of venomous marine gastropods. Biol. J. Linn. Soc. 73, 391–409.
Duda, T.F., Palumbi, S.R., 1999. Developmental shifts and species selection in gastropods. Proc. Natl. Acad. Sci. USA 96, 10272–10277.
Duda, T.F., Rolan, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Mol. Ecol. 14, 267–272.
Duggen, S., Hoernle, K., van den Bogaard, P., Rupke, L., Phipps Morgan, J., 2003. Deep roots of the Messinian salinity crisis. Nature 422, 602–606.
Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nat. Comm. 5, 3521.
Dyhr, C.T., Holm, P.M., 2010. A volcanological and geochemical investigation of Boa Vista, Cape Verde Islands; $^{40}Ar/^{39}Ar$ geochronology and field constraints. J. Volcanol. Geotherm. Res. 189, 19–32.
Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.
Filippelli, G.M., Flores, J.-A., 2009. From the warm pliocene to the cold pleistocene: a tale of two oceans. Geology 37, 959–960.
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.
Glibert, M., 1960. Les Conacea fossiles du Cénozoïque étranger des collections de l'Institut Royal des Sciences Naturelles de Belgique. Institut Royal des Sciences Naturelles de Belgique Mémoire, série 2 (64), 1–132.
Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., Posada, D., 2010. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res. 38, W14–W18.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotech. 29, 644–652.
Gulden, G.J., Moolenbeek, R.G., Goud, J., 2017. A new *Lautoconus* (Gastropoda: Conidae) from Senegal. Misc. Malac. 7, 49–53.
Harvey, P.H., Pagel, M.D., 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford, UK.
Hodell, D.A., Curtis, J.H., Sierro, F.J., Raymo, M.E., 2001. Correlation of late miocene to early pliocene sequences between the mediterranean and North Atlantic. Paleoceanography 16, 164–178.
Hollander, J., Butlin, R.K., 2010. The adaptive value of phenotypic plasticity in two ecotypes of a marine gastropod. BMC Evol. Biol. 10, 333.
Huelsenbeck, J., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Ivanovic, R.F., Valdes, P.J., Flecker, R., Gutjahr, M., 2014. Modelling global-scale climate impacts of the late Miocene Messinian salinity crisis. Clim. Past. 10, 607–622.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Knowlton, N., 1993. Sibling species in the sea. Annu. Rev. Ecol. Syst. 24, 189–216.
Kohn, A.J., 2014. *Conus* of the southeastern United States and caribbean. Princeton University Press, Princeton, NJ, USA.
Kohn, A.J., Nishi, M., Pernet, B., 1999. Snail spears and scimitars: a character analysis of *Conus* radular teeth. J. Molluscan. Stud. 65, 461–481.
Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J., Zardoya, R., 2015. TRUFA: A user-friendly web server for *de novo* RNA-seq analysis using cluster computing. Evol. Bioinform. 11, 97–104.
Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652–655.
Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. Partition finder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20, PA1003.
Marko, P.B., Hoffman, J.M., Emme, S.A., McGovern, T.M., Keever, C.C., Nicole Cox, L., 2010. The 'expansion–contraction' model of pleistocene biogeography: rocky shores suffer a sea change? Mol. Ecol. 19, 146–169.
Monnier, E., Limpalaër, L., 2010. *Conus dorotheae* (Gastropoda, Conidae) A new species of cone from the Cape Verde Peninsula in Senegal. Visaya 2, 73–80.
Monteiro, A., Tenorio, M.J., Poppe, G.T., 2004. The family conidae: the west african and mediterranean species of *Conus*. ConchBooks, Hackenheim, Germany.
Navarro-Pérez, E., Barton, E.D., 2001. Seasonal and interannual variability of the Canary current. Sci. Mar. 65 (Suppl. 1), 205–213.
Nolf, F., Verstraeten, J., 2008. *Conus trencarti* (Mollusca: Gastropoda: Conidae): a new cone from Senegal. Neptunea 74, 1–12.
Olivera, B.M., Watkins, M., Bandyopadhyay, P., Imperial, J.S., de la Cotera, E.P.H., Aguilar, M.B., Vera, E.L., Concepcion, G.P., Lluisma, A., 2012. Adaptive radiation of venomous marine snail lineages and the accelerated evolution of venom peptide genes. Ann. N.Y. Acad. Sci. 1267, 61–70.
Osca, D., Templado, J., Zardoya, R., 2015. Caenogastropod mitogenomics. Mol. Phylogenet. Evol. 93, 118–128.
Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. Annu. Rev. Ecol. Syst. 25, 547–572.
Pereira, C.M., Rosado, J., Seabra, S.G., Pina-Martins, F., Paulo, O.S., Fonseca, P.J., 2010. *Conus pennaceus*: a phylogenetic analysis of the Mozambican molluscan complex. Afr. J. Mar. Sci. 32, 591–599.
Peters, H., O'Leary, B.C., Hawkins, J.P., Carpenter, K.E., Roberts, C.M., 2013. *Conus*: first comprehensive conservation Red List assessment of a marine gastropod mollusc genus. PLoS ONE 8, e83353.
Petuch, E.J., Berschauer, D.P., 2016. A new species of cone shell (Gastropoda: Conidae) from the Saharan coast of northwestern Africa. The Festivus 48, 93–99.
Pin, M., Leung Tack, K.D., 1995. The cones of Senegal. Conchiglia 277, 1–56.
Puillandre, N., Bouchet, P., Duda Jr, T.F., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Mol. Phylogenet. Evol. 78, 290–303.
Puillandre, N., Duda, T.F., Meyer, C., Olivera, B.M., Bouchet, P., 2015. One, four or 100 genera? A new classification of the cone snails. J. Molluscan. Stud. 81, 1–23.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.
Röckel, D., Korn, W., Kohn, A.J., 1995. Manual of the living conidae, vol. I. Christa Hemmen Verlag, Wiesbaden, Indo-Pacific.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., Bierne, N., 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. PLoS. Biol. 14, e2000234.
Rüber, L., Van Tassell, J.L., Zardoya, R., 2003. Rapid speciation and ecological divergence in the American seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. Evolution 57, 1584–1598.
Sacco, F., 1893. I Molluschi Dei Terreni Terziarii Del Piemonte E Della Liguria. Conidae e Conorbidae. Stamperia Reale, Torino. Italy.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.
Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864.
Shen, K.-N., Jamandre, B.W., Hsu, C.-C., Tzeng, W.-N., Durand, J.-D., 2011. Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*. BMC Evol. Biol. 11, 83.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
StJohn, J., 2011. SeqPrep. https://github.com/jstjohn/SeqPrep.
Stramma, L., Siedler, G., 1988. Seasonal changes in the North Atlantic subtropical gyre. J. Geophys. Res. Oceans 93, 8111–8118.
Streelman, J.T., Alfaro, M., Westneat, M.W., Bellwood, D.R., Karl, S.A., 2002. Evolutionary history of the parrotfishes: biogeography, ecomorphology, and comparative diversity. Evolution 56, 961–971.
Tenorio, M.J., Lorenz, F., Dominguez, M., 2016. New insights into *Conus jourdani* da Motta, 1984 (Gastropoda, Conidae), an endemic species from Saint Helena Island. Xenophora Taxonomy 11, 32–42.
Tucker, J.K., Tenorio, M.J., 2009. Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. Conchbooks.
Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MDM Publishing, Wellington, FL, USA.
Uribe, J.E., Puillandre, N., Zardoya, R., 2017. Beyond *Conus*: phylogenetic relationships of Conidae based on complete mitochondrial genomes. Mol. Phylogenet. Evol. 107, 142–151.
Williams, S.T., Reid, D.G., 2004. Speciation and diversity on tropical rocky shores: a global phylogeny of snails of the genus *Echinolittorina*. Evolution 58, 2227–2251.

3.3. – Chapter III: “Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitocondrial genomes”

3.3. - Capítulo III: “Relaciones filogenéticas de los conos endémicos de Cabo Verde basadas en genomas mitocondriales”

BMC Evolutionary Biology, 17: 231

Antecedentes: Debido a su gran diversidad de especies y ecológica, así como a su capacidad para producir cientos de toxinas diferentes, lo conos son de interés para biólogos evolutivos, farmacólogos y naturalistas aficionados por igual. La identificación taxonómica de los conos todavía se basa principalmente en la forma, el color y los patrones de bandas de la concha. Sin embargo, estos rasgos fenotípicos son propensos a la homoplasia. Por lo tanto, se necesita urgentemente usar consistentemente datos genéticos para la delimitación de especies y la inferencia filogenética en este grupo aparentemente hiperdiverso. Aquí, nosotros reconstruimos la filogenia de los conos endémicos del archipiélago de Cabo Verde, una conocida radiación del grupo, utilizando genomas mitocondriales (mt).

Resultados: La filogenia reconstruida agrupó las especies analizadas en dos clados principales, uno de ellos incluyendo *Kalloconus* de África occidental, hermano de *Trovaoconus* de Cabo Verde y el otro con un *Lautoconus* parafilético debido a la relación de grupo hermano de *Africonus* de Cabo Verde y *Lautoconus ventricosus* del Mar Mediterráneo y regiones próximas, excluyendo a los *Lautoconus* endémicos de Senegal (más *Lautoconus guanche* de Mauritania, Marruecos y Canarias). Dentro de *Trovaoconus* se pueden distinguir hasta tres clados principales. El clado de *Africonus* incluye cuatro clados principales (llamados I a IV), cada uno subdividido en dos grupos monofiléticos. La filogenia reconstruida nos permitió inferir la evolución de la rádula en los linajes estudiados, así como los patrones biogeográficos. Se revisó el número de especies de conos endémicas de Cabo Verde a la luz de los datos de divergencia de secuencia y las relaciones filogenéticas inferidas.

Conclusiones: El nivel de divergencia entre los miembros continentales del género *Kalloconus* y las especies endémicas de las islas atribuidas al género *Trovaoconus* son

bajas, lo que lleva a la sinonimización de este último. El género *Lautoconus* es parafilético. *Lautoconus ventricosus* es el grupo hermano vivo más cercano del género *Africonus*. La diversificación de *Africonus* sucedió en alopatría debido al desarrollo directo de sus larvas y principalmente desencadenado por cambios eustáticos en el nivel del mar durante el Mioceno-Plioceno. Nuestro estudio confirma la diversidad de conos endémicos de Cabo Verde, pero reduce significativamente el número de especies válidas. Aplicando un umbral de divergencia, el número de especies válidas dentro de los *Africonus* muestreados se reduce a la mitad.

Abalde et al. BMC Evolutionary Biology (2017) 17:231
DOI 10.1186/s12862-017-1069-x

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitochondrial genomes

Samuel Abalde[1], Manuel J. Tenorio[2], Carlos M. L. Afonso[3], Juan E. Uribe[1], Ana M. Echeverry[1] and Rafael Zardoya[1*]

Abstract

Background: Due to their great species and ecological diversity as well as their capacity to produce hundreds of different toxins, cone snails are of interest to evolutionary biologists, pharmacologists and amateur naturalists alike. Taxonomic identification of cone snails still relies mostly on the shape, color, and banding patterns of the shell. However, these phenotypic traits are prone to homoplasy. Therefore, the consistent use of genetic data for species delimitation and phylogenetic inference in this apparently hyperdiverse group is largely wanting. Here, we reconstruct the phylogeny of the cones endemic to Cabo Verde archipelago, a well-known radiation of the group, using mitochondrial (mt) genomes.

Results: The reconstructed phylogeny grouped the analyzed species into two main clades, one including *Kalloconus* from West Africa sister to *Trovaoconus* from Cabo Verde and the other with a paraphyletic *Lautoconus* due to the sister group relationship of *Africonus* from Cabo Verde and *Lautoconus ventricosus* from Mediterranean Sea and neighboring Atlantic Ocean to the exclusion of *Lautoconus* endemic to Senegal (plus *Lautoconus guanche* from Mauritania, Morocco, and Canary Islands). Within *Trovaoconus*, up to three main lineages could be distinguished. The clade of *Africonus* included four main lineages (named I to IV), each further subdivided into two monophyletic groups. The reconstructed phylogeny allowed inferring the evolution of the radula in the studied lineages as well as biogeographic patterns. The number of cone species endemic to Cabo Verde was revised under the light of sequence divergence data and the inferred phylogenetic relationships.

Conclusions: The sequence divergence between continental members of the genus *Kalloconus* and island endemics ascribed to the genus *Trovaoconus* is low, prompting for synonymization of the latter. The genus *Lautoconus* is paraphyletic. *Lautoconus ventricosus* is the closest living sister group of genus *Africonus*. Diversification of *Africonus* was in allopatry due to the direct development nature of their larvae and mainly triggered by eustatic sea level changes during the Miocene-Pliocene. Our study confirms the diversity of cone endemic to Cabo Verde but significantly reduces the number of valid species. Applying a sequence divergence threshold, the number of valid species within the sampled *Africonus* is reduced to half.

Keywords: Mitochondrial genomes, *Africonus*, *Trovaoconus*, *Kalloconus*

* Correspondence: rafaz@mncn.csic.es
[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain
Full list of author information is available at the end of the article

Background

The cone snails (Conidae, Gastropoda) endemic to the archipelago of Cabo Verde in West Africa represent one of the few textbook examples of a well-documented insular species radiation involving marine organisms [1–3]. Cone snails, which are found in tropical and subtropical marine waters throughout the world, show a hotspot of species diversity in the Cabo Verde archipelago with up to 95 endemic species (roughly 10% of cone species diversity worldwide) narrowly confined to about 4000 km^2 [4]. As in other parts of the world, cone snails endemic to Cabo Verde constitute a key component of the intertidal and subtidal ecosystems associated to rocky shores, coral reefs, and sandy bottoms. All cones endemic to Cabo Verde feed on marine annelid worms [1] and use a sophisticated venom apparatus (including a venom gland that produces conotoxins and a specialized harpoon-like radular tooth) to capture their preys [5]. Another interesting biological feature common to all these endemic species is that they have direct development. Their larvae lack a pelagic stage, and thus show a considerably reduced dispersal capacity [1]. Survival rate is higher for this type of larvae since they are less likely to be eaten by predators and are not dependent on plankton for feeding (i.e, non-planktotrophic).

The origin and evolutionary history of cones endemic to Cabo Verde has been the subject of several recent phylogenetic studies [1, 2, 6, 7]. Molecular phylogenies demonstrated that two different ancestors reached the archipelago independently and subsequently diversified following recurrent biogeographic patterns [1, 2, 7]. The existence of two clades led to the classification of cone species endemic to Cabo Verde into two genera, *Africonus* and *Trovaoconus* [8]. The question of which species are the closest living sister groups to *Africonus* and *Trovaoconus* remains open [1, 2]. According to a previous study, the ancestor of *Africonus* colonized the archipelago in the Miocene, about 16.5 million years ago (mya; [1]), and spread to all islands (except Fogo, the youngest, with steep slopes in the coast and ongoing volcanic activity). Most (95%) of the currently described species endemic to Cabo Verde belong to *Africonus*, and are normally referred to as restricted to a single island and in some cases even to single bays within an island [3]. The ancestor of *Trovaoconus* arrived at Cabo Verde archipelago in the Pliocene, about 4.6 mya, and diversified only in four islands (Sal, Boa Vista, Maio, and possibly Santiago), which are the closest to the continent [1]. These cones are significantly larger in size than those belonging to *Africonus* and show wider distributions extending in some cases to more than one island. It has been hypothesized that diversification within each genus was in allopatry and followed recurrent eustatic sea level changes during the Neogene that intermittently connected and disconnected the islands [1, 7]. However, sea level fluctuations alone do not fully explain the extraordinary diversity of cones in Cape Verde since nearby archipelagos in the Macaronesia biogeographic region such as the Canary Islands subjected to similar trends since the Miocene do not have endemic cone species [6]. A larger distance to the mainland, which enhances isolation and restricts gene flow combined with a higher mean sea surface temperature and the presence of more suitable habitats may have promoted a significant increase in diversification rates in the Cabo Verde archipelago [6].

The rate of description of new cone species endemic to Cabo Verde has accelerated more than expected during the last years (Fig. 1). After the early descriptions in the eighteenth and nineteenth centuries based on samples brought to Europe by naturalists [9], the main contribution to the cataloguing of cone species endemic to Cabo Verde was due to the work of Emilio Rolán [10], who drew attention to this singular radiation. Hence, around year 2000, there were about 50 species recognized [11] and remarkably this number has almost doubled in the last 2-3 years [12–19]. However, it is important to note that many of the recent species diagnoses in cones are mainly based on the shape, color, and banding patterns of the shell. These phenotypic characters are highly variable at the population level and prone to local adaptation and convergence, making species assignment problematic and sometimes, misleading [7]. In many cases, distinguishing whether different shell morphotypes of cone snails represent valid species or ecotypes of the same species is challenging [20]. Therefore, determination of genetic variation and inference of phylogenetic relationships based on DNA sequence data are timely as part of a multidisciplinary approach [21] to identify and delimit species and to understand evolutionary processes underlying diversification within cones, in general, and within those endemic to Cabo Verde, in particular.

Here, we used nearly complete mitochondrial (mt) genomes, which have proven to successfully reconstruct robust phylogenies of Conidae [22] and of particular groups such as the cones endemic to Senegal [23]. In this study, we sequenced the nearly complete mt genomes of 88 individuals representing different populations and species of *Africonus* and *Trovaoconus* endemic to Cabo Verde. We aimed to: (1) reconstruct a highly resolved phylogeny of cones endemic to Cabo Verde; (2) determine the closest living sister groups of *Africonus* and *Trovaoconus*; (3) date major cladogenetic events and analyze biogeographical patterns; (4) study radular tooth evolution within the two genera; and (5) provide a first genetic hypothesis of species delimitation in the radiation of Cabo Verde endemic cones.

Results

Sequencing, assembly, and genome organization

The nucleotide sequences of the near-complete mt genomes of 75 specimens of *Africonus*, 13 specimens of

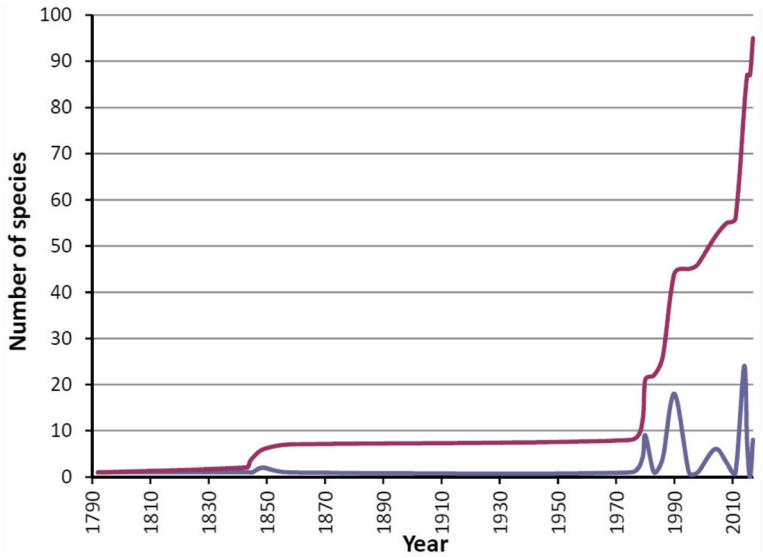


Fig. 1 Number (blue) and accumulated number (red) of cone species described for the Cabo Verde archipelago per year

Trovaoconus, and one specimen of *Lautoconus ventricosus* were determined (Table 1). These mt genomes lacked the *trnF* gene, the control region, and the start of the *cox3* gene because the corresponding fragment was not PCR amplified. The number of reads, mean coverage, and length of each mt genome are provided in Table 1. The mt genomes of *Africonus boavistensis* and *Africonus denizi* received the minimum (42,021) and maximum (906,765) number of reads, respectively. The same samples received the minimum (412×) and maximum (8,885×) mean coverage, respectively (Table 1). All sequenced mt genomes encode for 13 protein-coding, 2 rRNA and 21 tRNA genes (but note that the *trnF* gene could not be determined; see above). They all share the same genome organization: the major strand encodes all genes, except those forming the cluster MYCWQGE (*trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, *trnE*) and the *trnT* gene.

Phylogenetic relationships and sequence divergences between clades

Phylogenetic relationships of cones endemic to Cabo Verde were reconstructed based on the nucleotide sequences of the concatenated 13 mt protein-coding and two rRNA genes using probabilistic methods and *Chelyconus ermineus* as outgroup. The final matrix was 13,572 positions in length. According to the AIC, the best partition scheme for the protein-coding genes was the one combining all these genes but analyzing each codon position separately. The best substitution model for each of the three codon positions was GTR + I + G. For the rRNA genes, the best scheme had both genes combined under the GTR + I + G model. Both, ML (−*lnL* = 75,600.18) and BI (−*lnL* = 76,002.71 for run 1; −*lnL* = 76,288.44 for run 2) arrived at almost identical topology (Figs. 2 and 3). Most nodes received high statistical support and differences in topology were restricted exclusively to three relatively shallow nodes that had low support in ML and were unresolved in BI. Two of these nodes involved almost identical sequences and corresponded to *Africonus bernardinoi*/ *Africonus pseudocuneolus* and *Africonus teodorae*/ *Africonus fiadeiroi*, respectively. The third unresolved node corresponded to a trichotomy involving *Africonus felitae*, *Africonus regonae* and *Africonus longilineus*/ *Africonus cagarralensis*/ *Africonus melissae*.

The reconstructed phylogeny (Fig. 2) grouped the analyzed species into two main clades, one including *Kalloconus* from mainland West Africa sister to *Trovaoconus* from Cabo Verde and the other having paraphyletic *Lautoconus* due to the sister group relationship of *Africonus* from Cabo Verde and *Lautoconus ventricosus* from Mediterranean Sea and neighboring Atlantic Ocean to the exclusion of *Lautoconus* endemic to Senegal (plus *Lautoconus guanche* from Mauritania, Morocco, and Canary Islands). Within *Trovaoconus*, up to three main lineages could be distinguished (Fig. 2). The first one included two specimens from Sal initially identified as *Trovaoconus ateralbus*, which were sister to a clade including one lineage with specimens from Maio and Boa Vista identified as *Trovaoconus venulatus* and another lineage having mostly specimens of *Trovaoconus pseudonivifer* from Maio and Boa Vista but also one specimen of *Trovaoconus trochulus* from Boa Vista and one of *Trovaoconus atlanticoselvagem* from Baixo João Valente (Fig. 2).

The clade of *Africonus* from Cabo Verde included four main lineages (named I to IV), each further subdivided

Table 1 Mitochondrial (mt) genomes analyzed in this study

New mt genomes										
ID CV	Initial species identification	Location	Coordinates	Coverage		Length (bp)	GenBank Acc. No	Voucher DNA (MNCN/ADN)	Voucher shell (MNCN 15.05/)	New species proposed[a]
				n° reads	mean depth					
1020	*Africonus antoniaensis*	Água Doce, Boa Vista, Cabo Verde	16°12'29"N, 22°44'7"W	151104	1476.8	15332	MF491587	95072	79889	—
0885	*Africonus antoniomonteiroi*	Pedra Lume, Sal, Cabo Verde	16°45'44"N, 22°53'2"W	232069	2273.4	15328	MF491578	95063	79794	—
0927	*Africonus bernardinoi*	Pedra Lume, Sal, Cabo Verde	16°45'44"N, 22°53'2"W	59799	583.3	15328	MF491582	95067	79835	*Africonus cuneolus*
0520	*Africonus boavistensis*	Baía do Ervatão (North), Boa Vista, Cabo Verde	16°12'3"N, 22°54'43"W	42021	412.8	15217	MF491563	95045	80413	—
1135	*Africonus cabraloi*	Estancinha, Boa Vista, Cabo Verde	16°13'12"N, 22°55'9"W	74446	730.4	15329	MF491598	95083	80004	*Africonus crotchii*
0895	*Africonus cagarralensis*	Pedra Lume, Sal, Cabo Verde	16°45'44"N, 22°53'2"W	161290	1367.2	15320	MF491579	95064	79804	*Africonus longilineus*
0173	*Africonus calhetae*	Praia da Soca, Maio, Cabo Verde	15°15'8"N, 23°13'4"W	55433	544.7	15242	MF491534	95016	78798	—
0920	*Africonus* cf. *anthonyi*	Ilhéus do Chano, Sal, Cabo Verde	16°41'37"N, 22°52'47"W	172336	1678.6	15315	MF491581	95066	79828	*Africonus cuneolus*
0162	*Africonus* cf. *claudiae*	Praia da Soca, Maio, Cabo Verde	15°15'8"N, 23°13'4"W	87407	858.6	15326	MF491533	95015	78787	*Africonus calhetae*
0465	*Africonus* cf. *delanoyae*	Ponta Antónia, Boa Vista, Cabo Verde	16°13'24"N, 22°46'59"W	382817	3736.1	15335	MF491559	95041	80409	*Africonus fuscoflavus*
0207	*Africonus* cf. *galeao*	Ponta Pipa, Maio, Cabo Verde	15°19'30"N, 23° 9'48"W	81447	797.3	15325	MF491536	95018	78832	*Africonus galeao*
0135	*Africonus* cf. *gonsaloi*	Praia Gonçalo, Maio, Cabo Verde	15°16'13"N, 23°6'15"W	148032	1455.5	15250	MF491529	95011	78760	*Africonus gonsaloi*
0380	*Africonus* cf. *miguelfiaderoi*	Jorrita, Baía da Gata, Boa Vista, Cabo Verde	16°12'9"N, 22°42'22"W	358342	3507.9	15328	MF491548	95030	80398	*Africonus vulcanus*
1400	*Africonus* cf. *miruchae*	Calhau, São Vicente, Cabo Verde	16°51'7"N, 24°51'59"W	523002	5104.9	15321	MF491601	95088	78562	*Africonus* sp. nov. 1
0223	*Africonus claudiae*	Ponta Pipa, Maio, Cabo Verde	15°19'30"N, 23° 9'48"W	148508	1434.5	15337	MF491537	95019	78848	*Africonus galeao*
0303	*Africonus condei*	Baía Grande, Derrubado, Boa Vista, Cabo Verde	16°13'31"N, 22°47'17"W	253863	2472.2	15248	MF491542	95024	80392	*Africonus crotchii*
0045	*Africonus crioulus*	Praia Santana, Maio, Cabo Verde	15°18'13"N, 23°11'49"W	255019	2502	15247	MF491521	95003	78670	*Africonus maioensis*
1075	*Africonus crotchii*	Morro de Areia, Boa Vista, Cabo Verde	16°5'24"N, 22°57'7"W	332385	3237.6	15329	MF491591	95076	79944	—
0803	*Africonus cuneolus*	Calheta Funda, Sal, Cabo Verde	16°39'6"N, 22°56'53"W	184181	1791.6	15329	MF491569	95053	79712	—
0936	*Africonus cuneolus*	Santa Maria, Sal, Cabo Verde	16°35'38"N, 22°53'36"W	80472	787.4	15328	MF491583	95068	79844	—
1420	*Africonus curralensis*	Praia de Palmo Tostão, Santa Luzia, Cabo Verde	16°45'19"N, 24°45'24"W	857123	8358.6	15329	MF491602	95089	78581	—
1017	*Africonus damioi*	Água Doce, Boa Vista, Cabo Verde	16°12'29"N, 22°44'7"W	76477	745.5	15326	MF491586	95071	79886	*Africonus roeckeli*
0405	*Africonus damottai*	Baía da Gata (center), Boa Vista, Cabo Verde	16°11'50"N, 22°42'32"W	315488	2914.3	15358	MF491551	95033	80401	—
1428	*Africonus decoratus*	Curral, Santa Luzia, Cabo Verde	16°46'23"N, 24°47'13"W	566822	5540.2	15326	MF491603	95090	78589	—
0370	*Africonus delanoyae*	Jorrita, Baía da Gata, Boa Vista, Cabo Verde	16°12'9"N, 22°42'22"W	158489	1543.7	15323	MF491547	95029	80397	—
1471	*Africonus denizi*	Praia Grande, São Vicente, Cabo Verde	16°51'40"N, 24°52'30"W	906765	8885.2	15326	MF491605	95092	78621	—
0315				214173	2089.2	15243	MF491543	95025	80393	

Table 1 Mitochondrial (mt) genomes analyzed in this study *(Continued)*

	Africonus derrubado	Baía Grande, Derrubado, Boa Vista, Cabo Verde	16°13'31"N, 22°47'17"W							*Africonus damottai*
0565	*Africonus diminutus*	Ilhéu de Sal Rei, Boa Vista, Cabo Verde	16°9'50"N, 22°55'31"W	840424	8204.1	15330	MF491566	95049	80416	—
1025	*Africonus docensis*	Água Doce, Boa Vista, Cabo Verde	16°12'29"N, 22°44'7"W	47313	464.8	15329	MF491588	95073	79894	*Africonus crotchii*
0385	*Africonus evorai*	Zebraca (near Ilhéu do Galeão), Boa Vista, Cabo Verde	16°12'6"N, 22°42'40"W	226416	2218	15243	MF491549	95031	80399	*Africonus crotchii*
0070	*Africonus fantasmalis*	Porto Cais, Maio, Cabo Verde	15°19'15"N, 23°11'10"W	97527	954.7	15330	MF491524	95006	78695	*Africonus fuscoflavus*
0835	*Africonus felitae*	Rabo de Junco, Sal, Cabo Verde	16°41'44"N, 22°58'35"W	344190	3343.9	15404	MF491573	95057	79744	—
1437	*Africonus fernandesi*	Porto Novo, Santo Antão, Cabo Verde	17°1'4"N, 25°3'22"W	742414	7244.2	15324	MF491604	95091	78598	—
0332	*Africonus fiadeiroi*	Derrubado (bay West), Boa Vista, Cabo Verde	16°13'22"N, 22°47'41"W	205910	2016.5	15243	MF491545	95027	80395	*Africonus crotchii*
0855	*Africonus fontonae*	Baía da Fontona, Sal, Cabo Verde	16°44'22"N, 22°58'46"W	156259	1523.9	15328	MF491575	95059	79764	*Africonus cuneolus*
0945	*Africonus fontonae*	Regona, Sal, Cabo Verde	16°48'5"N, 22°59'33"W	56310	549.8	15327	MF491584	95069	79853	*Africonus regonae*
0450	*Africonus fuscoflavus*	Derrubado (bay East), Boa Vista, Cabo Verde	16° 13'331"N, 22°47'3"W	151904	1478.6	15331	MF491557	95039	80407	—
0052	*Africonus galeao*	Navio Quebrado, Terras Salgadas, Maio, Cabo Verde	15°18'54"N, 23°11'2"W	117940	1139.2	15326	MF491522	95004	78677	—
0134	*Africonus gonsaloi*	Praia Gonçalo, Maio, Cabo Verde	15°16'13"N, 23°6'15"W	188174	1835.9	15339	MF491528	95010	78759	—
1390	*Africonus grahami*	Calhau, São Vicente, Cabo Verde	16°51'7"N, 24°51'59"W	464704	4536.1	15325	MF491599	95086	78552	—
0140	*Africonus irregularis*	Porto Cais (North), Maio, Cabo Verde	15°19'45"N, 23°10'57"W	202254	1937.9	15321	MF491530	95012	78765	*Africonus maioensis*
0317	*Africonus irregularis*	Baía Grande, Derrubado, Boa Vista, Cabo Verde	16°13'31"N, 22°47'17"W	170523	1668.2	15331	MF491544	95026	80394	*Africonus maioensis*
0392	*Africonus irregularis*	Baía da Gata, Boa Vista, Cabo Verde	16°11'50"N, 22°42'32"W	252126	2454.3	15324	MF491550	95032	80400	*Africonus crotchii*
1084	*Africonus irregularis*	Morro de Areia, Boa Vista, Cabo Verde	16°5'24"N, 22°57'7"W	125264	1225.1	15330	MF491593	95078	79953	*Africonus crotchii*
1128	*Africonus irregularis*	Estancinha, Ponta do Sol, Boa Vista, Cabo Verde	16°13'12"N, 22°55'9"W	469101	4597.1	15313	MF491597	95082	79997	*Africonus crotchii*
0225	*Africonus isabelarum*	Ponta do Pau Seco, Maio, Cabo Verde	15°15'26"N, 23°13'16"W	247567	2431.7	15244	MF491538	95020	78850	—
0085	*Africonus josephinae*	Lage Branca, Maio, Cabo Verde	15°18'32"N, 23°8'17"W	224495	2204.6	15239	MF491525	95007	78710	*Africonus* sp. nov. 2
0555	*Africonus josephinae*	Ilhéu de Sal Rei, Boa Vista, Cabo Verde	16°9'50"N, 22°55'31"W	169723	1611.8	15330	MF491565	95048	80415	—
0830	*Africonus longilineus*	Serra Negra, Sal, Cabo Verde	16°38'17"N, 22°53'56"W	148726	1453	15316	MF491572	95056	79739	—
0847	*Africonus longilineus*	Rabo de Junco, Sal, Cabo Verde	16°41'44"N, 22°58'35"W	308057	3000.9	15333	MF491574	95058	79756	*Africonus miruchae*
0410	*Africonus luquei*	Praia Canto, Boa Vista, Cabo Verde	16°11'10"N, 22°42'28"W	83198	815.1	15244	MF491552	95034	80402	*Africonus delanoyae*
0064	*Africonus maioensis*	Porto Cais, Maio, Cabo Verde	15°19'15"N, 23°11'10"W	143797	1402.9	15327	MF491523	95005	78689	—
0510	*Africonus marckeppensi*	Ervatao Norte, Boa Vista, Cabo Verde	16°12'3"N, 22°54'43"W	254030	2480.3	15330	MF491562	95044	80412	*Africonus josephinae*
0102				80254	783.1	15326	MF491527	95009	78727	

Table 1 Mitochondrial (mt) genomes analyzed in this study *(Continued)*

	Africonus marcocastellazzii	Lage Branca, Maio, Cabo Verde	15°18'32"N, 23°8'17"W							*Africonus maioensis*
0870	*Africonus melissae*	Baía da Parda, Sal, Cabo Verde	16°45'7"N, 22°53'56"W	195531	1907.9	15328	MF491577	95061	79779	*Africonus longilineus*
0455	*Africonus messiasi*	Derrubado (bay East), Boa Vista, Cabo Verde	16°13'33"N, 22°47'3"W	250171	2447.7	15260	MF491558	95040	80408	*Africonus fuscoflavus*
0426	*Africonus miguelfiaderoi*	Praia Canto, Boa Vista, Cabo Verde	16°11'10"N, 22°42'28"W	130100	1270.5	15328	MF491554	95036	80404	*Africonus vulcanus*
0905	*Africonus mordeirae*	Baía do Roucamento, Sal, Cabo Verde	16°41'20"N, 22°56'24"W	99337	971.4	15241	MF491580	95065	79814	*Africonus cuneolus*
1091	*Africonus morroensis*	Morro de Areia, Boa Vista, Cabo Verde	16°5'24"N, 22°57'7"W	172039	1515.1	15337	MF491594	95079	79960	*Africonus diminutus*
1395	*Africonus navarroi*	Calhau, São Vicente, Cabo Verde	16°51'7"N, 24°51'59"W	665250	6509.2	15331	MF491600	95087	78557	—
0250	*Africonus nelsontiagoi*	Tarrafal, Santiago, Cabo Verde	15°16'50"N, 23°45'15"W	173873	1687.2	15339	MF491541	95023	78875	*Africonus verdensis*
0820	*Africonus pseudocuneolus*	Serra Negra, Sal, Cabo Verde	16°38'17"N, 22°53'56"W	131838	1288.7	15337	MF491571	95055	79729	*Africonus cuneolus*
0036	*Africonus raulsilvai*	Praia da Soca, Maio, Cabo Verde	15°15'8"N, 23°13'4"W	345872	1699.6	15534	MF491520	95002	78661	—
0865	*Africonus regonae*	Baía da Fontona, Sal, Cabo Verde	16°44'22"N, 22°58'46"W	246041	2411	15328	MF491576	95060	79774	—
0950	*Africonus regonae*	Regona, Sal, Cabo Verde	16°48'5"N, 22°59'33"W	88673	864.1	15337	MF491585	95070	79858	—
0586	*Africonus roeckeli*	Praia Canto, Boa Vista, Cabo Verde	16°11'10"N, 22°42'28"W	141600	1385.7	15320	MF491567	95050	80417	—
0549	*Africonus salreiensis*	Ilhéu de Sal Rei, Boa Vista, Cabo Verde	16°9'50"N, 22°55'31"W	349070	3402.9	15331	MF491564	95047	80414	*Africonus crotchii*
0810	*Africonus serranegrae*	Serra Negra, Sal, Cabo Verde	16°38'17"N, 22°53'56"W	182124	1777.6	15335	MF491570	95054	79719	*Africonus cuneolus*
1078	*Africonus silviae*	Morro de Areia, Boa Vista, Cabo Verde	16°5'24"N, 22°57'7"W	293568	2876.1	15336	MF491592	95077	79947	*Africonus fuscoflavus*
0445	*Africonus swinneni*	Porto Ferreira, Boa Vista, Cabo Verde	16°7'45"N, 22°40'17"W	221672	2169.1	15244	MF491556	95038	80406	*Africonus delanoyae*
1125	*Africonus teodorae*	Estancinha, Ponta do Sol, Boa Vista, Cabo Verde	16°13'12"N, 22°55'9"W	326654	3185.5	15334	MF491596	95081	79994	*Africonus crotchii*
1035	*Africonus umbelinae*	Espingueira, Boa Vista, Cabo Verde	16°12'55"N, 22°47'49"W	110526	1085.5	15333	MF491589	95074	79904	*Africonus damottai*
0240	*Africonus verdensis*	Tarrafal, Santiago, Cabo Verde	15°16'50"N, 23°45'15"W	53400	519.8	15339	MF491540	95022	78865	—
0435	*Africonus vulcanus*	Porto Ferreira, Boa Vista, Cabo Verde	16°7'45"N, 22°40'17"W	284766	2787.4	15242	MF491555	95037	80405	—
1110	*Africonus zinhoi*	Curral Velho, Boa Vista, Cabo Verde	15°58'4"N, 22°47'42"W	388828	3785.5	15331	MF491595	95080	79979	*Africonus maioensis*
7036	*Trovaoconus atlanticoselvagem*	Baixo João Valente, Cabo Verde	15°44′27″N, 23°5′26″W	95264	933.8	15352	MF491606	7036	—	*Kalloconus trochulus*
0616	*Trovaoconus* cf. *ateralbus*	Serra Negra, Sal, Cabo Verde	16°38'17"N, 22°53'56"W	87550	853.7	15344	MF491568	95052	79664	*Kalloconus* sp. nov. 1
0010	*Trovaoconus pseudonivifer*	Ponta do Pau Seco, Maio, Cabo Verde	15°15'26"N, 23°13'16"W	56486	555.5	15351	MF491519	95000	78635	*Kalloconus trochulus*
0094	*Trovaoconus pseudonivifer*	Lage Branca, Maio, Cabo Verde	15°18'32"N, 23°8'17"W	454415	4429.8	15351	MF491526	95008	78719	*Kalloconus trochulus*
0154	*Trovaoconus pseudonivifer*	Porto Cais (north), Maio, Cabo Verde	15°19'45"N, 23°10'57"W	199736	1954.2	15352	MF491532	95014	78779	*Kalloconus trochulus*
0420	*Trovaoconus pseudonivifer*	Praia Canto, Boa Vista, Cabo Verde	16°11'10"N, 22°42'28"W	111182	1085	15347	MF491553	95035	80403	*Kalloconus pseudonivifer*
0500				223915	2177.2	15351	MF491561	95043	80411	

Table 1 Mitochondrial (mt) genomes analyzed in this study *(Continued)*

	Trovaoconus trochulus	Baía do Ervatão (North), Boa Vista, Cabo Verde	16°12'3"N, 22°54'43"W							*Kalloconus trochulus*
0149	*Trovaoconus venulatus*	Lage Branca, Maio, Cabo Verde	15°18'32"N, 23°8'17"W	105732	1014.9	15276	MF491531	95013	78774	*Kalloconus venulatus*
0187	*Trovaoconus venulatus*	Praia Real, Maio, Cabo Verde	15°19'45"N, 23°10'40"W	144651	1415	15330	MF491535	95017	78812	*Kalloconus venulatus*
0234	*Trovaoconus venulatus*	Ponta do Pau Seco, Maio, Cabo Verde	15°15'26"N, 23°13'17"W	67867	661.5	15320	MF491539	95021	78859	*Kalloconus venulatus*
0347	*Trovaoconus venulatus*	Derrubado (bay West), Boa Vista, Cabo Verde	16°13'22"N, 22°47'41"W	48636	475.9	15326	MF491546	95028	80396	*Kalloconus venulatus*
0475	*Trovaoconus venulatus*	Ponta Antónia, Boa Vista, Cabo Verde	16°13'24"N, 22°46'59"W	409041	3966.1	15340	MF491560	95042	80410	*Kalloconus venulatus*
1038	*Trovaoconus venulatus*	Praia Canto, Boa Vista, Cabo Verde	16°11'10"N, 22°42'28"W	143644	2403.8	15336	MF491590	95075	79907	*Kalloconus venulatus*
IB001	*Lautoconus ventricosus*	Estani des Peix, Formentera, Balearic Islands, Spain	38°43'49"N, 1°24'42"E	86290	842.5	15341	MF491607	95094	80426	*Lautoconus* sp. nov 1

GenBank mt genomes

ID	Species	Location	Coordinates	Reference	Length (bp)	GenBank Acc. No	Voucher (MNCN/ADN)	Voucher shell (MNCN 15.05/)	New species proposed[a]
6990	*Africonus borgesi*	Porto Ferreira, Boa Vista, Cabo Verde	16°7'45"N, 22° 40'170"W	Cunha et al., (2009)	15536	NC_013243	6990	—	—
0025	*Africonus infinitus*	Ponta do Pau Seco, Maio, Cabo Verde	15°15'26"N, 23°13'17"W	Abalde et al., (in prep.)	15522	KY864967	95001	78650	—
0875	*Africonus miruchae*	Terrinha Fina, Palhona, Sal, Cabo Verde	16°49'12"N, 22°59'12"W	Abalde et al., (in prep.)	15336	KY864971	95062	79784	—
0534	*Trovaoconus pseudonivifer*	Estancinha, Ponta do Sol, Boa Vista, Cabo Verde	16°13'12"N, 22°55'9"W	Abalde et al., (in prep.)	15351	KY864969	95046	80418	*Kalloconus trochulus*
0550	*Trovaoconus venulatus*	Ilhéu de Sal Rei, Boa Vista, Cabo Verde	16°9'56"N, 22°55'23"W	Uribe et al., (2017)	15524	KX263250	86741	80419	*Kalloconus venulatus*
0601	*Trovaoconus ateralbus*	Calheta Funda, Sal, Cabo Verde	16°39'6"N, 22°56'53"W	Abalde et al., (in prep.)	15327	KY864970	95051	79649	*Kalloconus ateralbus*
1375	*Kalloconus* cf. *byssinus*	North Senegal	unknown	Abalde et al., (in prep.)	15348	KY864973	95085	78536	*Kalloconus pulcher*
1253	*Kalloconus pulcher*	Les Almadies, Dakar, Senegal	14°44'40"N, 17° 31'442"W	Abalde et al., (in prep.)	15332	KY864972	95084	78414	*Kalloconus pulcher*
1343	*Lautoconus belairensis*	Terrou-Bi. Dakar, Senegal	14°40'29"N, 17°28'12"W	Abalde et al., (2017)	15321	KY801849	91293	78504	*Gen. nov. belairensis*
1338	*Lautoconus bruguieresi*	Île de Gorée, Dakar, Senegal	14°40'16"N, 17°23'58"W	Abalde et al., (2017)	15340	KY801851	91291	78499	*Gen. nov. bruguieresi*
1296	*Lautoconus cloveri*	Ndayane, Senegal	14°33'45"N, 17°7'34"W	Abalde et al., (2017)	15323	KY801859	91283	78457	*Gen. nov. cloveri*
CG13	*Lautoconus guanche*	Lanzarote, Canary Islands, Spain	28°57'16"N, 13°34'22"W	Abalde et al., (2017)	15506	KY801847	91295	—	*Gen. nov. guanche*
1266	*Lautoconus hybridus*	NGor, Dakar, Senegal	14°45'67"N, 17° 30'36.33"W	Abalde et al., (2017)	15507	KY801863	91279	78427	*Gen. nov. hybridus*
1278	*Lautoconus mercator*	NGor, Dakar, Senegal	14°45'6"N, 17°30'36"W	Abalde et al., (2017)	15329	KY801862	91280	78439	*Gen. nov. mercator*
CV13	*Lautoconus ventricosus*	Ria Formosa, Faro, Portugal	36°58'0"N, 7°53'2"W	Uribe et al., (2017)	15534	KX263251	86742	—	—
CVERM1	*Chelyconus ermineus*	Praia Gonçalo, Maio, Cabo Verde	15°16'13"N, 23°6'15"W	Abalde et al., (in prep.)	15365	KY864977	95095	78876	—

[a]hyphen indicates that original species name is maintained and considered valid

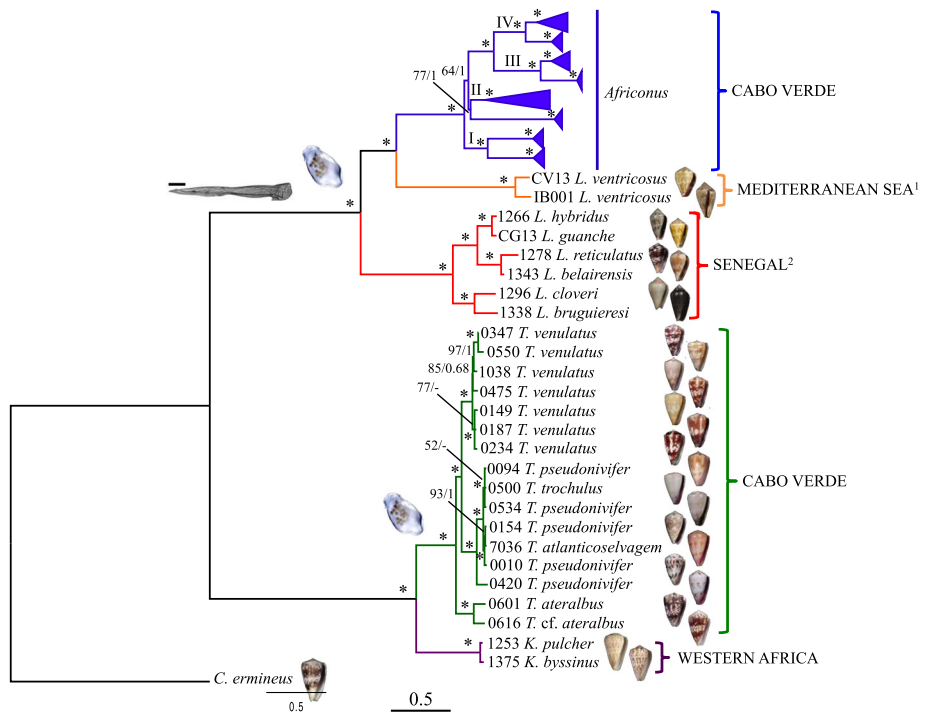


Fig. 2 Closest sister groups of cone snails endemic to Cabo Verde based on mt genomes (concatenated protein coding plus rRNA genes analyzed at the nucleotide level). The reconstructed ML tree using *C. ermineus* as outgroup is shown. Number of specimen, initial species assignment, and a ventral picture of the shell are provided. Numbers at nodes are statistical support values for ML (bootstrap proportions)/BI (posterior probabilities). Scale bar indicates substitutions/site. Two major clades are recovered: the first one includes *Kalloconus* from Western Africa (purple) and *Trovaoconus* from Cabo Verde (green) whereas the second one includes *Lautoconus* from the Mediterranean Sea (and neighboring Atlantic Ocean[1]; orange), *Lautoconus* from Senegal (and *L. guanche* from Canary Islands[2]; red), and *Africonus* (blue) from Cabo Verde. Lineages within *Africonus* are expanded in Fig. 3. The "robust" type of radular tooth is shown as the ancestral character state for the second clade (scale bar equals 0.1 mm; see diversity of radular teeth of the first clade in Additional file 1). The two transitions from planktonic to non-planktonic larvae are indicated by a sac of eggs

into two monophyletic groups (Figs. 2 and 3). Lineage I was the sister group of the remaining *Africonus* and its two lineages had each species from Maio sister to species from Boa Vista (Fig. 3a). Lineage II included species from Santiago and Maio sister to species endemic to the westernmost islands (Santo Antão, São Vicente and Santa Luzia). These latter species could be grouped into three main lineages, one containing species endemic to São Vicente, another containing species distributed both in Santa Luzia and São Vicente, and the third one including species from the three islands (Fig. 3a). Lineage III included species from Maio sister to species from Boa Vista (Fig. 3a). Lineage IV contained specimens representing most of the described species of *Africonus*. One monophyletic group included species endemic to Sal whereas the other clade included *Africonus isabelarum* from Maio as sister to four lineages, two containing exclusively species from Boa Vista, one having species from Maio sister to *Africonus irregularis* from Boa Vista, and one having species from Boa Vista and *Africonus fantasmalis* from Maio (Fig. 3b).

Pairwise uncorrected sequence divergences were estimated based on the alignment including the nucleotide sequences of the 13 mt protein-coding and two rRNA genes. Pairwise uncorrected sequence divergences between *C. ermineus* and ingroup taxa averaged 18%. The average pairwise uncorrected sequence divergence between the two main ingroup clades (genera *Kalloconus* + *Trovaoconus* versus genera *Lautoconus* + *Africonus*) was 16%. Pairwise uncorrected sequence divergences between *Lautoconus* endemic to Senegal (plus *L. guanche*) and *L. ventricosus* plus *Africonus* averaged 11%. The average pairwise uncorrected sequence divergence between the sister groups *L. ventricosus* and *Africonus* was 10% whereas between *Kalloconus* and *Trovaoconus,* it was 5%. The pairwise uncorrected sequence divergences between the four main lineages within *Africonus* averaged 6%. The corresponding values for the pairwise divergences between the two major clades defined within each of the lineages I-IV were 4%, 6%, 3%, and 3%, respectively. Pairwise uncorrected sequence divergence comparisons between sister species level were distributed into two different ranges, one closer to 1% (0.5-

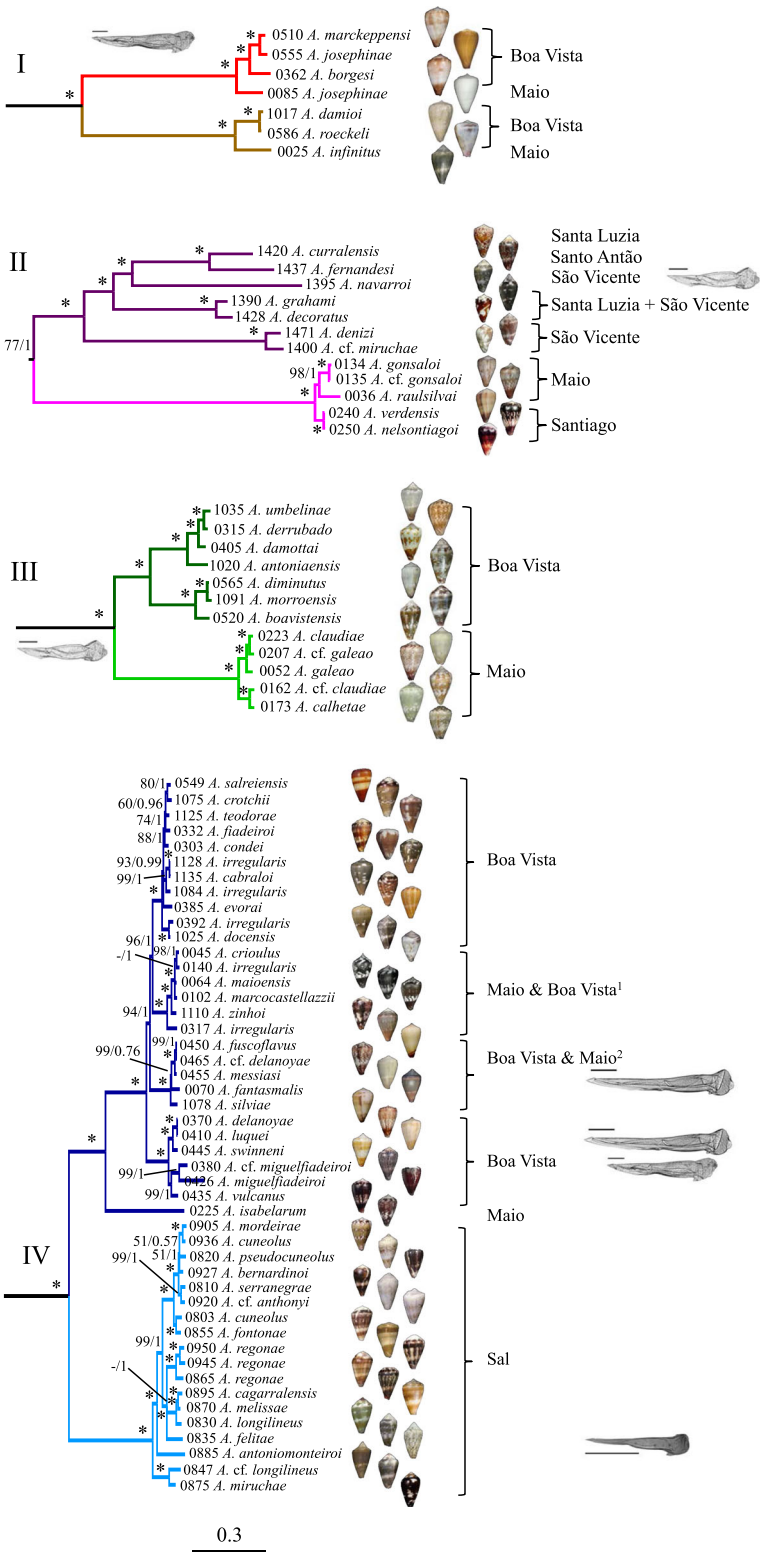


Fig. 3 (See legend on next page.)

(See figure on previous page.)
Fig. 3 Phylogeny of *Africonus* based on mt genomes (concatenated protein coding plus rRNA genes analyzed at the nucleotide level). Number of specimen, initial species assignment, island, and a ventral picture of the shell are provided. Numbers at nodes are statistical support values for ML (bootstrap proportions)/BI (posterior probabilities). Hyphen indicates a bootstrap value below 50%. Scale bar indicates substitutions/site. Four major lineages (I-IV) are recovered and indicated with different colors. All *Africonus* have the "robust" type of radular tooth except when indicated (scale bar equals 0.1 mm). [1]All taxa endemic to Maio except *A. irregularis* endemic to Boa Vista. [2]All taxa endemic to Boa Vista except *A. fantamalis* endemic to Maio

1.5%) and the other closer to 0% (0-0.5%). The latter divergences were particularly abundant among sister species comparisons within Maio, Boa Vista and Sal. Several mt genomes of different species were almost identical (<0.05%) in sequence including (1) *Africonus delanoyae* and *Africonus luquei*, (2) *Africonus fuscoflavus*, *Africonus* cf. *delanoyae*, and *Africonus messiasi*, (3) *Africonus irregularis* (#1128) and *Africonus cabraloi*, (4) *Africonus verdensis* and *Africonus nelsontiagoi*, and (5) *Africonus gonsaloi* and *Africonus* cf. *gonsaloi*.

Evolution of radular types

The different lineages within *Africonus* exhibit distinct radular types (Fig. 3). Most lineages and species showed the "robust" type, which is of medium relative size, with a short, pointed barb and a basal spur (see Additional file 1). The anterior section of the tooth is equal or slightly shorter than the posterior section, and the blade covers most of the anterior section (80% – 85%). There are usually 19 to 30 denticles in the serration, arranged in one row (occasionally two). Several species within lineage IV (*Africonus delanoyae*, *Africonus luquei*, *Africonus swinneni*, *Africonus fuscoflavus*, *Africonus messiasi*, *Africonus silviae* and *Africonus* cf. *delanoyae* from Boa Vista island, and *Africonus fantasmalis* from Maio island) exhibited radular teeth of the "elongated" type, similar to the "robust" type but characterized by an anterior section which is longer than the posterior section, a blade covering 40 to 50% of the anterior section, and more numerous denticles in the serration (usually more than 30) often arranged in two rows. Several species (*Africonus borgesi*, *Africonus josephinae* and *Africonus marckeppensi* in lineage I, *Africonus navarroi* in lineage II), all species in lineage III, plus *Africonus vulcanus*, *Africonus miguelfiadeiroi* and *Africonus* cf. *miguelfiadeiroi* in lineage IV) displayed radular tooth of the "broad" type, which is characterized by a medium-sized (Shell Length/Tooth Length = 32-45) and very broad radular tooth (Shell Length/Anterior section Width = 7-12), with an anterior section which is shorter than the posterior section (Tooth Length/Anterior section Length = 2.1-2.9), a blade covering most of the anterior section, and with a variable number of denticles (8 to 30) in the serration arranged in two or more rows. The radular morphology of *Africonus felitae* may represent a special case with a small relative size (Shell Length/Tooth Length = 63-67), narrow (Tooth Length/Anterior section Width = 20-23), the anterior section shorter than the posterior section (Tooth Length/Anterior section Length = 2.2-2.4), and characterized by the total absence of denticles in the serration. The base of this tooth is relatively large and broad.

The species of *Kalloconus* and *Trovaoconus* exhibit essentially two kinds of radular morphologies (Additional file 1). The teeth in *K. pulcher*, and also in *Trovaoconus trochulus*, *T. pseudonivifer* and *Trovaoconus atlanticoselvagem* are narrow and elongated; the blade is moderately short being about one third to almost one-half the length of the anterior section of the tooth, which is distinctly longer than the posterior section of the tooth. There are many denticles (25 to 45 or more) in the long serration, arranged usually in multiple rows with a major row flanked by numerous smaller serrations. In the case of *T. venulatus*, *Trovaoconus ateralbus*, and *Trovaoconus* cf. *ateralbus* the teeth are broader, and the anterior and posterior sections are almost equal in length. There are 16 to 33 denticles in the serration, often coarse and hook-shaped in the middle portion, arranged initially in one row becoming two rows below.

Dating of major cladogenetic events

Major cladogenetic events within the reconstructed phylogeny were dated using an uncorrelated relaxed molecular clock model, which was calibrated using the age of Sal (28 mya; the oldest island of the archipelago) for the node separating *Africonus* from its sister group, *L. ventricosus*, and the age of the origins of São Vicente, Santo Antão, and Santa Luzia (7.5 mya) for the node splitting the lineage including the endemics to these islands from its sister group lineage including endemics to Maio and Santiago islands [24]. The first divergence event involving *Kalloconus* + *Trovaoconus* versus (paraphyletic) *Lautoconus* + *Africonus* was dated at 34 mya (Fig. 4; note that genera and species labels in the chronogram take into account proposed synonymizations, see Discussion). The divergence between the clade containing cones endemic to Senegal (+ *L. guanche*) and the clade including *L. ventricosus* plus *Africonus* was dated at 26 mya. The split between the latter two lineages was dated at 23 mya. The diversification of the crown group of *Africonus* into its four main lineages (I-IV) was estimated to have occurred between 9.4 - 6.9 mya (Fig. 4). The separation of *Kalloconus* and *Trovaoconus* was dated

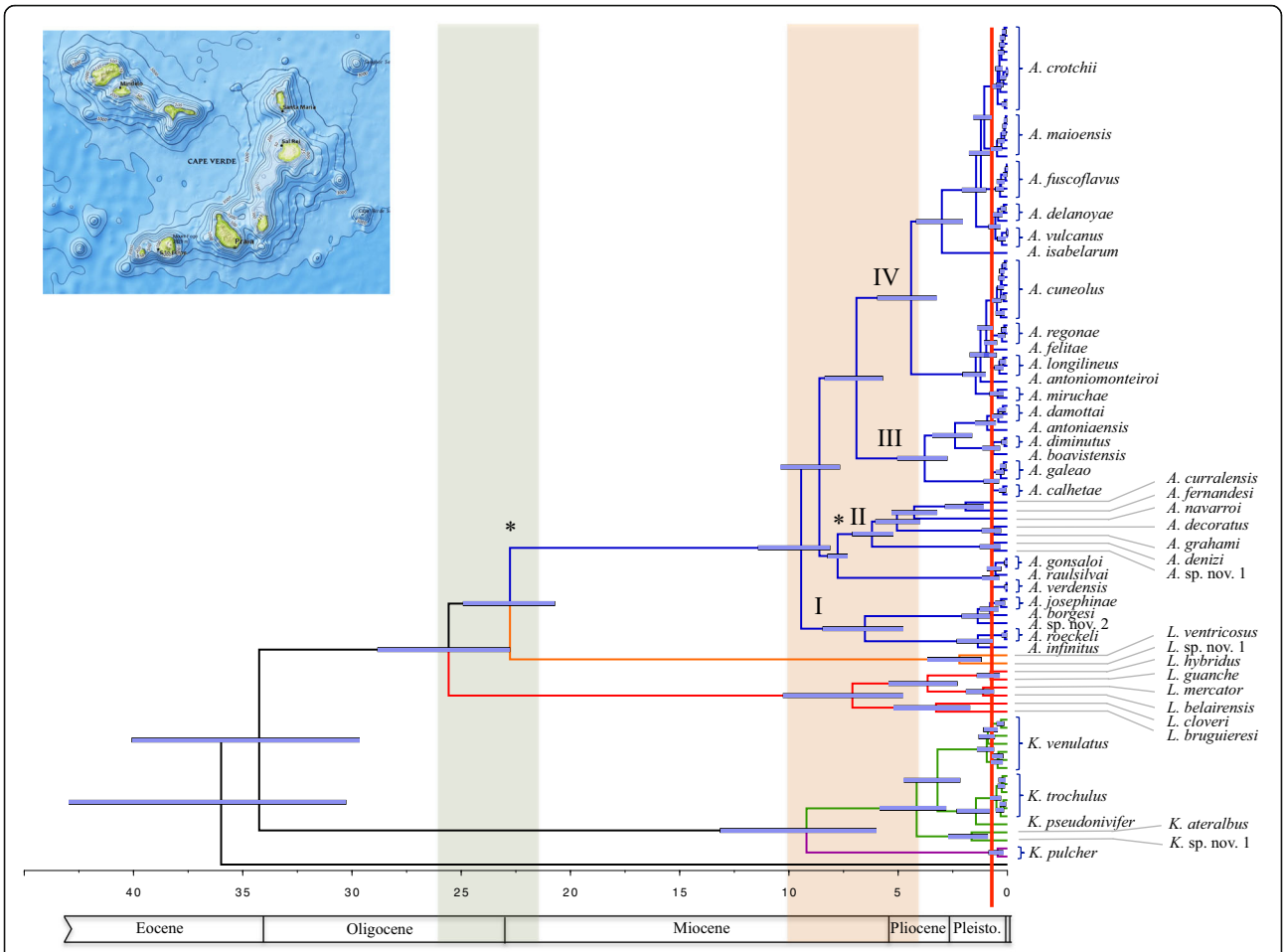


Fig. 4 Chronogram based on mt genomes (concatenated protein coding plus rRNA genes analyzed at the nucleotide level) and using the fixed topology of the ML tree shown in Figs. 2 and 3. Lineages are colored as in Fig. 2. Synonymizations proposed (see text) are taken into account. A Bayesian uncorrelated relaxed lognormal clock with geographic-based calibration priors (denoted by asterisks) was used in BEAST. Horizontal bars represent 95% credible intervals for time estimates; dates are in millions of years. Geological periods are indicated. Brown and orange bands indicate main divergence periods around the Oligocene-Miocene and Miocene-Pliocene transitions, respectively. A red line indicates the threshold for species delimitation. The bathymetry of Cabo Verde archipelago is shown in an inset

9 mya and the diversification of the crown group of *Trovaoconus* was established at 4 mya (Fig. 4).

Diversification rates through time

Variations in the diversification rates through time were estimated for *Africonus* and the hypothesis of a radiation during the evolutionary history of the clade was tested. The gamma-statistic, which measures departures from a constant rate of diversification, had values of 7.19 ($p < 0.05$) and 3.12 ($p < 0.05$) when considering the currently named (based on phenotypic traits) or only the here-proposed (considering genetic evidence) species for the genus, respectively. In both cases, the hypothesis of a radiation is accepted. According to the lineage through time plots (Fig. 5), the initial rate of increase in the number of species slowed down between six and one and a half million years ago regardless of the species delimitation hypothesis tested. Afterwards, the diversification rate accelerated considerably, and the increase in number of species either continued or abandoned a normal Yule process of speciation when considering the species delimitation hypothesis here proposed or the currently number of named species, respectively.

Discussion

Cone snails are marine gastropods well known to evolutionary biologists due to their extraordinary species and ecological diversity [25], but also to molecular biologists and pharmacologists due to their sophisticated venom cocktails [26], as well as to amateur naturalists due to their brightly colored and highly appreciated shells [27]. Therefore, they are the subject of intensive research across disciplines and additionally have received wide attention from the general public. There are more than 800 described species and this number increases steadily

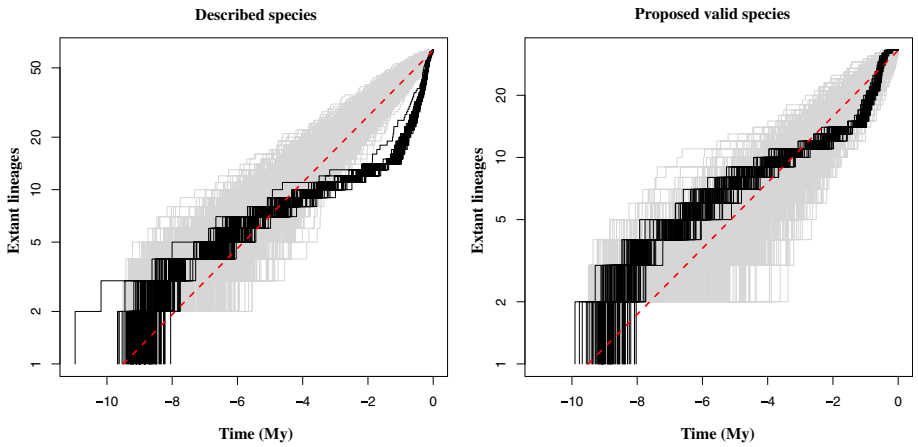


Fig. 5 Logarithmic lineage through time (LTT) plots of described (**a**) and proposed valid (**b**) species. The red bar indicates the pure Yule process of speciation, the grey shadow shows 1000 simulated trees and the black lines represent 100 trees randomly chosen among the trees generated by BEAST

every year [28]. Thus far, species description and identification of cones heavily relies on shell form, color and banding patterns, which may show great variety at local scale leading to important levels of synonymy within the family Conidae [7, 20]. In this regard, species delimitation could greatly improve with the aid of robust molecular phylogenies, which could be used in addition as framework to uncover the evolutionary patterns and processes underlying the diversification of the group. While reconstructing a robust phylogeny for all described cone species worldwide is cumbersome and at present unrealistic within the framework of a single study, it is possible, however, to accomplish a proof-of-concept study in a particular region [23].

We have here reconstructed a molecular phylogeny of cones endemic to Cabo Verde and allied species in the Macaronesian region, continental West Africa, and the Mediterranean region. These cones are particularly interesting from an evolutionary perspective as they have radiated in an oceanic archipelago and constitute a natural experiment to gain insights onto the processes governing diversification and adaptation [29]. Phylogenetic analyses were based on nearly complete mt genomes (only missing the control region and neighboring sequences) and included 105 specimens comprising most of the cone species diversity of the analyzed regions. Probabilistic methods of phylogenetic inference arrived at a robust and highly resolved phylogeny (virtually all nodes received high statistical support, which in most cases was maximal). To our knowledge, this is the first wide application of mt genomes to the resolution of a phylogeny within mollusks (but see [30, 31] for comparable examples in fish and insects, respectively). Previous studies in gastropods were restricted in the number of taxa analyzed (e.g., [22] for the family Conidae) but here we were able to achieve a lineage representation of the reference group (in this case, Cabo Verde cones) only previously attained by studies using few concatenated partial gene sequences (see e.g., [25] for the family Conidae or [1] for the cones of Cabo Verde). Previous phylogenetic studies using complete mt genomes have demonstrated that the level of resolution of these molecular markers is compromised above the superfamily level due to saturation, base composition biases, and among-lineage rate heterogeneity [32, 33]. Here, we show that phylogenetic performance of mt genomes achieves best results when analyzing closely related genera (and their corresponding species). Moreover, results were particularly promising taking into account that *Africonus* diversity in Cabo Verde was originated through radiation processes, which normally lead to relatively short tree nodes (often difficult to disentangle).

A thorough sampling of closely-related outgroup taxa allowed us to tackle key questions on the origin of the cones endemic to Cabo Verde and on their closest living sister groups. As previously reported, there are two independent origins of Cabo Verde cones, leading to the genera *Africonus* and *Trovaoconus*, respectively [1]. The closest living sister group of *Africonus* is *L. ventricosus* from the Mediterranean Sea and neighboring Atlantic Ocean. Therefore, the origin of this clade is clearly Macaronesian/Mediterranean and these cones are only distantly related to the geographically closer cones endemic to Senegal. These latter cones were ascribed to the genus *Lautoconus* (as was the case of *L. guanche* from Canary islands, deeply nested within the clade of Senegal cones; [23]). However, the closest sister group relationship of *Africonus* and *L. ventricosus* requires formal description of a different genus for Senegal cones (plus *L. guanche*), which will be done elsewhere. We could not include any representative of cones endemic to Angola (genus *Varioconus*) but a recent phylogeny based on partial *cox1* gene

sequences recovered all these cones (including *Varioconus jourdani* from Saint Helena Island) as a monophyletic group sister to Senegal cones [34]. Alternatively, all previously mentioned genera could be merged into genus *Lautoconus* [35]. However, the relatively high levels of sequence divergence (using the sequence divergence of genus *Chelyconus* as reference) and the restricted (endemic) distribution of the clades fit better with the former taxonomic proposal. The closest living sister group of *Trovaoconus* is genus *Kalloconus* from West Africa. Therefore, the origin of this clade is clearly related to neighboring regions of the continent. Actually, the sequence divergence between *Trovaoconus* and *Kalloconus* is much lower than that estimated between *Lautoconus* and *Africonus*. This observation argues against maintaining the generic status of *Trovaoconus*, and supports the inclusion of their species within genus *Kalloconus*, as some authors have proposed [35]. Hence, *Kalloconus* would be a genus that is present throughout the coast of West Africa from Morocco to Angola as well as in Canary Islands and Cabo Verde.

Altogether, the two main clades in the reconstructed phylogeny show very distinct patterns of distribution. One clade includes a single genus with widespread distribution in Macaronesia and West Africa whereas the other clade, which occupies the same geographical regions, is divided into several valid genera (*Africonus*, *Lautoconus*, *Varioconus*, Gen. nov. for Senegal endemics). These distinct patterns could be explained partly taking into account differences in larval dispersal capabilities between the two clades [1]. According to the phylogeny, the ancestor of the *Kalloconus* clade was inferred to have planktotrophic larvae, capable of long dispersals whereas the ancestor of the other clade would have non-planktotrophic larvae, and thus a limited dispersal capability leading to restricted gene flow and higher rates of diversification [36, 37]. Interestingly, the ancestor of *Kalloconus* species endemic to Cabo Verde (former *Trovaoconus* species) lost planktotrophy, which is a common evolutionary pattern in insular species [38].

According to the reconstructed phylogeny, cones belonging to genus *Africonus* are divided into four main lineages (I-IV; with each further subdivided into two distinct clades). Species endemic to Maio and Boa Vista are found in all four lineages whereas species endemic to Sal form a clade within lineage IV, species from the westernmost islands (Santo Antão, São Vicente and Santa Luzia) form a clade within lineage II, and the single species from Santiago is recovered within lineage II. Unfortunately, we could not sample specimens of *Africonus furnae* from Brava and *Africonus kersteni* from São Nicolau, and cannot determine whether they could be ascribed to any of the above-mentioned four lineages or form their own independent lineages. The single origin of cones endemic to Sal, Santiago, and westernmost islands could be explained by the deep slopes separating these islands whereas the multiple origins of the cones found in Maio and Boa Vista could be associated to the relatively shallow seamount (Baixo João Valente) connecting both islands [1]. These differences in bathymetry in connection with past eustatic sea level changes could be determinant in preventing or promoting dispersal in *Africonus* species, whose larvae are all non-planktotrophic.

Diversification events among main lineages were concentrated in three major periods. The first one, around the Oligocene-Miocene boundary (23 mya), includes the divergence of cones endemic to Senegal (and Angola) from their sister clade, and the posterior separation within this sister clade of cones endemic to Cabo Verde and those endemic to the Mediterranean Sea and neighboring Atlantic Ocean. During Oligocene-Miocene transition, there was a global cooling event [39, 40], the ice sheet of Antarctica greatly expanded, and a sea level drop of ~50 m occurred [41]. The second period corresponds to a sustained global cooling in the Late Miocene starting 12 mya [42] that produced an eustatic sea level drop between −10 and −30 m from 6.26 to 5.50 Mya [43] and culminated with the Messinian Salinity crisis and the desiccation of the Mediterranean Sea at the end of the Miocene from 5.96 to 5.33 Mya [44]. During this period, the divergence of the main lineages within *Africonus* (I-IV), the cones endemic to Senegal, and *Kalloconus* occurred. Finally, a burst of speciation events is inferred during the Pleistocene when another cooling period characterized by extreme climate oscillations and drastic eustatic sea level changes concurring with glacial-interglacial periods [45]. Global cooling has been recently proposed to be a driver of diversification of marine species [46] in agreement with our results. The reconstructed phylogeny, the chronogram, and the current geographical distribution of the species altogether support that allopatry is the main mode of speciation for cone snails with non-planktotrophic larvae, as previously suggested [1]. The complex geology of the island of Boa Vista with several eruptions at >16, 15-12.5, and 9.5-4.5 mya [47], involving different parts of the island may have also contributed to creating additional niches along the coast and could explain that this island harbors the highest number of endemic cones.

The reconstructed phylogeny also allows inferring the evolution of the radula in the studied lineages [8, 48]. All analyzed ingroup taxa are vermivorous [8]. Studies documenting potential specialization of the vermivore radular type to prey on specific worm species are scarce and restricted thus far to cone species preying on amphinomids [49]. Here, we show that most *Africonus* species show a "robust" radular type, which is shared also with *L. ventricosus* and a lineage of Senegal cones represented by *Lautoconus cloveri* and *Lautoconus bruguieresi* in the phylogeny [23]. Therefore, the common ancestor of cones endemic to Senegal (plus *L. guanche*), *Africonus*, and *L.*

ventricosus was inferred to have a "robust" type radula. The "elongated" type of radular tooth, which was found in several species within lineage IV of *Africonus*, also appears in a lineage of Senegal cones that is represented by *Lautoconus hybridus* and *L. guanche* in the phylogeny [23]. The radular tooth of *A. felitae* resembles the "small" type observed in a lineage of Senegal cones represented by *Lautoconus reticulatus* and *Lautoconus belairensis* in the phylogeny [23]. The "broad" type of radular tooth that appeared independently in several lineages of *Africonus* has not been observed in any cone from Senegal. While shifts in radular type could be correlated with early cladogenesis in cones endemic to Senegal [23], the evolution of different types of radular tooth within *Africonus* was restricted to few specific cases. Thus, future studies are needed to determine whether in such cases there has been a dietary shift to prey on specific worms. The radula teeth identified in *Kalloconus* resemble the types "elongated" and "robust" observed within *Lautoconus* and *Africonus*, although are clearly distinct. This might indicate instances of convergence, and that only a discrete number of different main types of radula could be found in a given clade.

During the last few years, the number of new cone species described from Cabo Verde has increased at an astonishing rate (e.g., [12]). These new species are identified based on differences (often subtle) in shell shape and color, and their status needs to be contrasted with genetic data to uncover cases of local phenotypic variation within species due to either genetic polymorphism or phenotypic plasticity that may be producing overestimations of the number of species in the group [21]. In addition, genetic data could help identify cases of phenotypic convergence due to adaptation of genetically distinct populations (ecotypes) or species (sibling or cryptic) to similar environments [50–52], also affecting the total number of valid species. Comparative analyses of pairwise uncorrected sequence divergences taking into account the reconstructed phylogeny showed that some described species shared almost identical mt genomes with levels of sequence divergence normally considered to be associated to genetic variation at the population level. Clades comprising these sets of closely related sequences indicate that an uncorrected sequence divergence threshold around 1% could be associated to the species status. This threshold lies well within the so-called grey zone of speciation between 0.5-2% [53]. Of course, these results need to be further confirmed with genomic nuclear data that discard potential events of incomplete lineage sorting and hybridization [54]. In addition, the present study could be further improved in the future by increasing the number of individuals analyzed per original species. Importantly, the comparative analyses on variation of diversification rates through time support the here proposed hypothesis of species delimitation as it concurs with a Yule process of speciation whereas the number of currently named species clearly exceed expectations and would imply an extraordinary recent acceleration of speciation rates.

Our study confirms the diversity of cone endemic to Cabo Verde but significantly reduces the number of valid species. Applying the threshold in a conservative manner (i.e., maintaining described species as valid in case of doubt due to closeness to the threshold) to cones endemic to Cabo Verde would reduce the number of valid species within the sampled *Africonus* from 65 to 32 (see Table 1 and Fig. 4). The proposed nomenclatural changes follow standard ICZN recommendations maintaining the most senior (oldest) name. Among the species not sampled, two correspond to São Nicolau and Brava islands, four of them are from the islands of São Vicente and Santa Luzía, and most likely represent valid species (*Africonus bellulus*, *Africonus lugubris*, *Africonus saragasae*, *Africonus santaluziensis*, *A. kersteni* and *A. furnae*) given the relative high sequence divergences found among species endemic to these islands. The 19 remaining ones were recently described, mostly from Boa Vista, and are expected to fall in most cases into some of the clades already discussed in the present work, and therefore may correspond to morphs of other described species. A direct consequence of synonymization is that some previously described species of rather restricted distribution are merged as populations into the new species, which considerably increase their range of distribution (Additional file 1). For instance, *A. crotchii*, which was reported as endemic from Southwest Boa Vista, would be now distributed also in the whole north half of the island. This increase in range of distribution of several species has important effects on their IUCN conservation status [3]. In the case of *Kalloconus*, some morphotypes attributed to *Kalloconus pseudonivifer* are now assigned to *Kalloconus trochulus*, and *Kalloconus atlanticoselvagem* is synonymized with *K. trochulus*. Our specimen of *Kalloconus* cf. *byssinus* is from North Senegal and has little sequence divergence compared to *Kalloconus pulcher*. In this case, it would be important to study *K. byssinus* from Mauritania or Morocco before considering synonymization. In the opposite direction, there are three clear instances of morphological convergence and thus, of the existence of cryptic species. Those are the cases of *Africonus josephinae* from Maio, *Africonus* cf. *miruchae* from São Vicente, and *Kalloconus* cf. *ateralbus* from Sal, which will be described as new species in due course.

Conclusions

We reconstructed a robust phylogeny based on mitochondrial genomes of cone snails endemic to Cabo Verde, which provides the necessary framework for future

evolutionary studies focused on this radiation. The double origin of Cabo Verde endemic cones was supported. The ancestor of *Africonus* separated from L. ventricosus during the Oligocene-Miocene boundary (about 23 mya) and diversified into four main lineages (I to IV) in the Late Miocene (about 9.4-6.9 mya). The divergence of the ancestor of *Kalloconus* endemic to Cabo Verde from those inhabiting mainland occurred also in the Late Miocene whereas its diversification into three main lineages was dated in the Pliocene (4 mya). Main cladogenetic events within cones endemic to Cabo Verde coincide with global cooling periods, which were characterized by radical climate oscillations and eustatic sea level changes. Recurrent cycles of island connection/ disconnection likely favored speciation in allopatry in these cones, which lack a pelagic larval stage, and thus have limited dispersal capacity. Direct development evolved in the ancestor of *Kalloconus* endemic to Cabo Verde, likely associated to the colonization of the archipelago by a cone with a planktotrophic larval stage. However, in the case of *Africonus*, the ancestor that arrived to Cabo Verde was already non-planktotrophic as the corresponding independent evolutionary shift to direct development predated the separation of cones endemic to Senegal (and Canary Islands) from *L. ventricosus* plus *Africonus*. Radular types were modified during the diversification of *Africonus* from an ancestral "robust" type, although correlation with diet specializations await better knowledge of the specific worm species preyed by the different species of cones. Sequence divergence comparisons and reconstructed phylogenies supported the diversity of cone species endemic to Cabo Verde but significantly reduced its number, which was likely overestimated in the past due to important homoplasy in shell morphology, the, thus far, main discriminant character used for species description and identification.

Methods

Samples and DNA extraction

The complete list of specimens analyzed in this study corresponding to different populations and species of *Africonus* and *Trovaoconus* from Cabo Verde is shown in Table 1, as well as details on the respective sampling localities and museum vouchers. As outgroup taxa, we also sampled and analyzed one specimen of *L. ventricosus* from Formentera Island (Spain). Specimens were collected by snorkel at 1-3 m depth, or picked by hand at low tide. All samples were stored in 100% ethanol. The initial species identification (see corresponding column in Table 1) was based on comparison with type material (mostly deposited in the MNCN) or consulting the original publications. Total DNA was isolated from 5 to 10 mg of foot tissue following a standard phenol-chloroform extraction [55].

Radular tooth preparation

The radular sac was dissected from the main body and soft parts were digested in concentrated aqueous potassium hydroxide for 24 h. The resulting mixture was then placed in a petri dish and examined with a binocular microscope. The entire radula was removed with fine tweezers and rinsed with distilled water, then mounted on a slide using Aquatex (Merck, Germany) mounting medium, and observed under a compound microscope. Photographs were taken with a charge-coupled device (CCD) camera attached to the microscope. Terminology for radular morphology follows [8], with abbreviations following [48]. Names of radular types follow [23].

PCR amplification and sequencing

Near-complete (without the control region) mt genomes were amplified through a combination of standard and long PCRs using the primers and following the protocols of [22]. Standard-PCR products were sequenced using Sanger technology. Long-PCR products were subjected to next-generation sequencing. Briefly, PCR amplified fragments from the same mt genome were pooled together in equimolar concentrations. For each cone mt genome a separate indexed library was constructed using the NEXTERA XT DNA library prep kit (Illumina, San Diego, CA, USA). The average size of the Nextera libraries varied between 307 and 345 bp. Libraries were pooled and run in an Illumina MiSeq platform (v.2 chemistry; 2×150 paired-end) at Sistemas Genómicos (Valencia, Spain).

Genome assembly and annotation

The reads corresponding to each mt genome were sorted using the corresponding library indices, and read assembly was performed in the TRUFA webserver [56]. Briefly, adapters were removed using SeqPrep [57], quality of the reads was checked using FastQC v.0.10.1 [58], and raw sequences were trimmed and filtered out according to their quality scores using PRINSEQ v.0.20.3 [59]. Filtered reads were used for de novo assembly of each mt genome using default settings (minimum contig length: 200; sequence identity threshold: 0.95) of Trinity r2012-06-08 [60] in TRUFA, and only retaining contigs with a minimum length of 3 kb. These contigs were used as starting point to assemble the mt genomes using Geneious® 8.0.3. First, the (raw) reads with a minimum identity of 99% were mapped against the contigs to correct possible sequence errors. Then, successive mapping iterations using a 100% identity as threshold were performed to elongate the contigs.

The mt genomes were annotated with the option "Annotate from Database" in Geneious® 8.0.3, using published mt genomes of Conidae as references. Annotations of the 13mt protein-coding genes were refined manually

identifying the corresponding open reading frames using the invertebrate mitochondrial code. The transfer RNA (tRNA) genes were further identified with tRNAscan-SE 1.21 [61], which infer cloverleaf secondary structures (with a few exceptions that were determined manually). The ribosomal RNA (rRNA) genes were identified by sequence comparison with other Conidae mt genomes [22], and assumed to extend to the boundaries of adjacent genes [62]. GenBank accession numbers of each mt genome are provided in Table 1.

Sequence alignment and phylogenetic analyses

The newly sequenced mt genomes were aligned with the mt genomes of *A. borgesi*, *Africonus infinitus*, *Africonus miruchae*, *T. ateralbus*, *T. pseudonivifer*, *T. venulatus*, and *C. ermineus* from Cabo Verde, *L. hybridus, L. mercator, L. belairensis, L. cloveri, L. bruguieresi*, *K. pulcher*, and *K.* cf. *byssinus* from Senegal, *L. guanche* from Canary Islands, and *L. ventricosus* from Portugal, which were downloaded from GenBank (Table 1). A sequence data set was constructed concatenating the nucleotide sequences of the 13 mt protein-coding and two rRNA genes. The deduced amino acid sequences of the 13 mt protein-coding genes were aligned separately and used to guide the alignment of the corresponding nucleotide sequences with Translator X [63]. Nucleotide sequences of the mt rRNA genes were aligned separately using MAFFT v7 [64] with default parameters. Ambiguously aligned positions were removed using Gblocks, v.0.91b [65] with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. Finally, the different single alignments were concatenated using Geneious® 8.0.3. Sequences where format converted for further analyses using the ALTER webserver [66]. The concatenated alignment is available at http://purl.org/phylo/treebase/phylows/study/TB2:S21557.

Phylogenetic relationships were inferred using maximum likelihood (ML, [67]) and Bayesian inference (BI, [68]). For ML, we used RAxML v8.1.16 [69] with the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates (BP). BI analyses were conducted with MrBayes v3.1.2 [70], running four simultaneous Markov chains for 10 million generation, sampling every 1000 generations, and discarding the first 25% generations as burn-in (as judged by plots of ML scores and low SD of split frequencies) to prevent sampling before reaching stationarity. Two independent Bayesian inference runs were performed to increase the chance of adequate mixing of the Markov chains and to increase the chance of detecting failure to converge, as determined using Tracer v1.6 [71]. The effective sample size (ESS) of all parameters was checked to be above 200. Node support was assessed based on Bayesian Posterior Probabilities (BPP). A node was considered highly supported with BP and BPP values above 70% and 0.95, respectively. The ML and BI phylogenetic trees are available at http://purl.org/phylo/treebase/phylows/study/TB2:S21557.

The best partition schemes and best-fit models of substitution for the data set were identified using PartitionFinder2 [72] with the Akaike information criterion [73]. For the protein-coding genes, the partitions tested were: all genes grouped; all genes separated (except *atp6-atp8* and *nad4-nad4L*); and genes grouped by subunits (*atp*, *cob, cox*, and *nad*). In addition, these three partitions schemes were tested taking into account separately the three codon positions. The rRNA genes were tested with two different schemes, genes separated or combined.

Estimation of divergence times

The program BEAST v.1.8.0 [74] was used to perform a Bayesian estimation of divergence times. An uncorrelated relaxed molecular clock was used to infer branch lengths and nodal ages. The tree topology was fixed using the one recovered by the ML analysis. For the clock model, the lognormal relaxed-clock model was selected, which allows rates to vary among branches without any a priori assumption of autocorrelation between adjacent branches. For the tree prior, a Yule process of speciation was employed. Concatenated protein coding plus rRNA genes were analyzed at the nucleotide level. The partitions and models selected by PartitionFinder2 were applied (see results). The final Markov chain was run twice for 100 million generations, sampling every 10,000 generations, and the first 1000 trees were discarded as part of the burn-in process, according to the convergence of chains checked with Tracer v.1.5. [71]. The ESS of all parameters was above 200.

Despite the fact that there are many fossils of Conidae, it is difficult in many instances to be certain about species identifications given the important levels of homoplasy in shell shape [75]. Hence, although there are fossils attributed to *L. ventricosus* [76] and *L. mercator* [77], which could be applied to the reconstructed phylogeny, we opted to calibrate the clock using biogeographical events (i.e., the age of the islands of Cabo Verde). We run a preliminary analysis in which the posterior distribution of the estimated divergence times was obtained by specifying one calibration point as prior for the divergence time of the split between *L. ventricosus* and the genus *Africonus*. This genus is endemic to Cabo Verde, and we used the age of formation of the oldest island, Sal (28 Mya; [24]), as biogeographical calibration point. We applied a log-normal distribution as the prior model for the calibration and enforced the median

divergence time to equal 25 (s.d. = 0.05, offset = 0.7). According to the results of the preliminary analysis, we found that only in the case of São Vicente, Santo Antão, and Santa Luzia, the early divergence of living cone endemic lineages followed the origin of the corresponding island, and therefore, we used a second calibration point corresponding to the origin of these islands about 7.5 mya [24]. We applied a log-normal distribution as the prior model for the calibration and enforced the median divergence time to equal 7.5 (s.d. = 0.03, offset = 0). The BEAST tree is available at http://purl.org/phylo/treebase/phylows/study/TB2:S21557.

Diversification rate through time

The chronogram was used to determine diversification rate through time of genus *Africonus* under alternative (phenotypic versus genetic) species delimitation hypotheses. A lineage through time (LTT) plot analysis was conducted using the APE 4.1 R package [78]. A random sample of 100 trees was selected and mapped over a simulation of 1000 trees following a Yule process of speciation (net diversification rate = 0.4). The phytools R package [79] was used to calculate the Gamma-Statistic [80].

Additional file

Additional file 1: Maps showing sampling localities; diversity of Kalloconus radular teeth. (ZIP 3996 kb)

Acknowledgements

We are indebted to Cabo Verde biology students Paulo Vasconcelos and Stiven Pires, as well as to our colleagues Sara Rocha, David Posada and Julio Rozas for their valuable help during sampling in Cabo Verde. We thank Dr. Rui Freitas from the Universidade de Cabo Verde (UniCV) for his continuous support and for granting us access to the facilities of the UniCV in Mindelo, São Vicente. We also thank Dr. Iderlindo Silva dos Santos and Dra. Sonia Monteiro de Pina Araujo from the Direcçao Nacional do Medio Ambiente of the Ministério do Ambiente, Habitação e Ordenamento do Território (MAHOT) of the Republic of Cabo Verde for their help with collecting permits (Autorizações 07/2013, 26/2013, 01/2104, 04/2015, and 03/2016). The sample of *L. ventricosus* from Formentera Island was obtained by Paula C. Rodríguez Flores. We are grateful to Jesús Marco and Aida Palacios, who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing assembling and phylogenetic analyses.

Funding

This work was supported by the Spanish Ministry of Science and Innovation (CGL2013-45211-C2-2-P and CGL2016-75255-C2-1-P (AEI/FEDER, UE)) to RZ; BES-2011-051469 to JEU; BES-2014-069575 to SA; Doctorado Nacional-567, Colciencias-Universidad Nacional de Colombia to AME.

Availability of data and materials

Sequence data is available in GenBank.

Authors' contributions

MJT, CMLA, and RZ collected the material. MJT prepared the radula. SA, JEU, and AME generated the molecular data. SA analyzed the data. RZ wrote the first draft of the manuscript and all authors contributed to writing the final version. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain. [2]Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cádiz, 11510 Puerto Real, Cádiz, Spain. [3]Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005 - 139 Faro, Portugal.

Received: 15 June 2017 Accepted: 6 November 2017
Published online: 25 November 2017

References

1. Cunha RL, Castilho R, Rüber L, Zardoya R. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde Islands. Syst Biol. 2005;54(4):634–50.
2. Duda TF, Rolán E. Explosive radiation of Cape Verde *Conus*, a marine species flock. Mol Ecol. 2005;14(1):267–72.
3. Peters H, O'Leary BC, Hawkins JP, Roberts CM. The cone snails of Cape Verde: marine endemism at a terrestrial scale. Global Ecology and Conservation. 2016;7:201–13.
4. Tucker JK, Tenorio MJ. Illustrated catalog of the living cone shells. Wellington: MDM Publishing; 2013.
5. Olivera BM, Watkins M, Bandyopadhyay P, Imperial JS, de la Cotera EPH, Aguilar MB, Vera EL, Concepcion GP, Lluisma A. Adaptive radiation of venomous marine snail lineages and the accelerated evolution of venom peptide genes. Ann N Y Acad Sci. 2012;1267(1):61–70.
6. Cunha RL, Lima FP, Tenorio MJ, Ramos AA, Castilho R, Williams ST. Evolution at a different pace: distinctive phylogenetic patterns of cone snails from two ancient oceanic archipelagos. Syst Biol. 2014;63(6):971–87.
7. Cunha RL, Tenorio MJ, Afonso C, Castilho R, Zardoya R. Replaying the tape: recurring biogeographical patterns in Cape Verde *Conus* after 12 million years. Mol Ecol. 2008;17(3):885–901.
8. Tucker JK, Tenorio MJ. Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. Hackenheim: Conchbooks; 2009.
9. Röckel D, Rolán E, Monteiro M. Cone shells from Cape Verde Islands - a difficult puzzle. I. Feito: Vigo, Spain; 1980.
10. Rolán E: Descripción de nuevas especies y subespecies del género *Conus* (Mollusca, Neogastropoda) para el archipiélago de Cabo Verde. Iberus 1990, Suppl. 2:5 - 70.
11. Monteiro A, Tenorio MJ, Poppe GT. The family Conidae. The west African and Mediterranean species of *Conus*. A Conchological Iconography. Hackenheim: ConchBooks; 2004.
12. Cossignani T. Dieci nuovi coni da Capo Verde. Malacol Mostra Mond. 2014;82:18–29.
13. Cossignani T, Fiadeiro R. Cinque nuovi coni da Capo Verde. Malacol Mostra Mond. 2014;84:21–7.
14. Tenorio MJ, Afonso CML, Cunha RL, Rolán E. New species of *Africonus* (Gastropoda, Conidae) from boa vista in the Cape Verde archipelago: molecular and morphological characterization. Xenophora Taxonomy. 2014;2:5–21.
15. Afonso CML, Tenorio MJ. Recent findings from the islands of Maio and boa vista in the Cape Verde archipelago, West Africa: description of three new *Africonus* species (Gastropoda: Conidae). Xenophora Taxonomy. 2014;3:47–60.
16. Cossignani T, Fiadeiro R. Quattro nuovi coni da Capo Verde. Malacol Mostra Mond. 2014;83:14–9.
17. Cossignani T, Fiadeiro R. Otto nuovi coni da Capo Verde. Malacol Mostra Mond. 2017;94:26–36.

18. Cossignani T, Fiadeiro R. Tre nuovi coni da Capo Verde. Malacol Mostra Mond. 2015;86:17–21.
19. Cossignani T, Fiadeiro R. Due nuovi coni da Capo Verde. Malacol Mostra Mond. 2015;87:3–5.
20. Duda TF, Palumbi SR. Developmental shifts and species selection in gastropods. Proc Natl Acad Sci U S A. 1999;96(18):10272–7.
21. Dayrat B. Towards integrative taxonomy. Biol J Linn Soc. 2005;85(3):407–15.
22. Uribe JE, Puillandre N, Zardoya R. Beyond *Conus*: phylogenetic relationships of Conidae based on complete mitochondrial genomes. Mol Phylogenet Evol. 2017;107:142–51.
23. Abalde S, Tenorio MJ, CML A, Zardoya R. Mitogenomic phylogeny of cone snails endemic to Senegal. Mol Phylogenet Evol. 2017;112:79–87.
24. Holm PM, Grandvuinet T, Friis J, Wilson JR, Barker AK, Plesner S. An 40Ar-39Ar study of the Cape Verde hot spot: temporal evolution in a semistationary plate environment. J. Geophys. Res. Solid Earth. 2008;113(B8):n/a-n/a.
25. Puillandre N, Bouchet P, Duda TF Jr, Kauferstein S, Kohn AJ, Olivera BM, Watkins M, Meyer C. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Mol Phylogenet Evol. 2014;78:290–303.
26. Olivera BM. *Conus* peptides: biodiversity-based discovery and exogenomics. J Biol Chem. 2006;281(42):31173–7.
27. Filmer RM. A catalogue of nomenclature and taxonomy in the living Conidae 1758 – 1998. Leiden: Backhuys Publishers; 2001.
28. Bouchet P, Gofas S. Conidae Fleming, 1822. In: MolluscaBase (2017). 2010. Accessed through: World Register of Marine Species at https://marinespeciesorg/aphiaphp?p=taxdetails&id=14107. Accessed 11 Nov 2017.
29. Warren BH, Simberloff D, Ricklefs RE, Aguilée R, Condamine FL, Gravel D, Morlon H, Mouquet N, Rosindell J, Casquet J, et al. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecol Lett. 2015;18(2):200–17.
30. Crampton-Platt A, Timmermans MJTN, Gimmel ML, Kutty SN, Cockerill TD, Vun Khen C, Vogler AP. Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a Bornean rainforest sample. Mol Biol Evol. 2015;32(9):2302–16.
31. Saitoh K, Sado T, Mayden RL, Hanzawa N, Nakamura K, Nishida M, Miya M. Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii: Ostariophysi): the first evidence toward resolution of higher-level relationships of the world's largest freshwater fish clade based on 59 whole mitogenome sequences. J Mol Evol. 2006;63(6):826–41.
32. Stöger I, Schrödl M. Mitogenomics does not resolve deep molluscan relationships (yet?). Mol Phylogenet Evol. 2013;69(2):376–92.
33. Osca D, Irisarri I, Todt C, Grande C, Zardoya R. The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca: Chaetodermomorpha) supports the Aculifera hypothesis. BMC Evol Biol. 2014;14(1):1–10.
34. Tenorio MJ, Lorenz F, Dominguez M. New insights into *Conus jourdani* da Motta, 1984 (Gastropoda, Conidae), an endemic species from Saint Helena Island. Xenophora Taxonomy. 2016;11:32–42.
35. Puillandre N, Duda TF, Meyer C, Olivera BM, Bouchet P. One, four or 100 genera? A new classification of the cone snails. J Molluscan Stud. 2015;81:1–23.
36. Jablonski D. Larval ecology and macroevolution in marine invertebrates. Bull Mar Sci. 1986;39(2):565–87.
37. Scheltema RS. Planktonic and non-planktonic development among prosobranch gastropods and its relationship to the geographic range of species. In: Ryland JS, Tyler RA, editors. Reproduction, genetics and distribution of marine organisms. Fredensborg: Olsen & Olsen; 1989. p. 183–8.
38. Swearer SE, Shima JS, Hellberg ME, Thorrold SR, Jones GP, Robertson DR, Morgan SG, Selkoe KA, Ruiz GM, Warner RR. Evidence of self-recruitment in demersal marine populations. Bull Mar Sci. 2002;70(1):251–71.
39. Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF. The Phanerozoic record of global sea-level change. Science. 2005;310(5752):1293.
40. Zachos JC, Flower BP, Paul H. Orbitally paced climate oscillations across the Oligocene/Miocene boundary. Nature. 1997;388(6642):567–70.
41. Beddow HM, Liebrand D, Sluijs A, Wade BS, Lourens LJ. Global change across the Oligocene-Miocene transition: high-resolution stable isotope records from IODP site U1334 (equatorial Pacific Ocean). Paleoceanography. 2016;31(1):81–97.
42. Herbert TD, Lawrence KT, Tzanova A, Peterson LC, Caballero-Gill R, Kelly CS. Late Miocene global cooling and the rise of modern ecosystems. Nat Geosci. 2016;9(11):843–7.
43. Hodell DA, Curtis JH, Sierro FJ, Raymo ME. Correlation of late Miocene to early Pliocene sequences between the Mediterranean and North Atlantic. Paleoceanography. 2001;16(2):164–78.
44. Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. Chronology, causes and progression of the Messinian salinity crisis. Nature. 1999;400(6745):652–5.
45. Lisiecki LE, Raymo ME. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography. 2005;20:PA1003.
46. Davis KE, Hill J, Astrop TI, Wills MA. Global cooling as a driver of diversification in a major marine clade. Nat Commun. 2016;7:13003.
47. Dyhr CT, Holm PM. A volcanological and geochemical investigation of boa vista, Cape Verde Islands; 40Ar/39Ar geochronology and field constraints. J Volcanol Geotherm Res. 2010;189(1–2):19–32.
48. Kohn AJ, Nishi M, Pernet B. Snail spears and scimitars: a character analysis of *Conus* radular teeth. J Molluscan Stud. 1999;65:461–81.
49. Duda TF, Kohn AJ, Palumbi SR. Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. Biol J Linn Soc. 2001;73(4):391–409.
50. Dowle EJ, Morgan-Richards M, Brescia F, Trewick SA. Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. Mol Ecol. 2015;24(16):4205–21.
51. Hollander J, Butlin RK. The adaptive value of phenotypic plasticity in two ecotypes of a marine gastropod. BMC Evol Biol. 2010;10(1):333.
52. Knowlton N. Sibling species in the sea. Annu Rev Ecol Syst. 1993;24(1):189–216.
53. Roux C, Fraïsse C, Romiguier J, Anciaux Y, Galtier N, Bierne N. Shedding light on the Grey zone of speciation along a continuum of genomic divergence. PLoS Biol. 2016;14(12):e2000234.
54. Gerard D, Gibbs HL, Kubatko L. Estimating hybridization in the presence of coalescence using phylogenetic intraspecific sampling. BMC Evol Biol. 2011;11(1):291.
55. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press; 1989.
56. Kornobis E, Cabellos L, Aguilar F, Frías-López C, Rozas J, Marco J, Zardoya R. TRUFA: a user-friendly web server for de novo RNA-seq analysis using cluster computing. Evol Bioinforma. 2015;11(Supplementary Material 23873):97–104.
57. StJohn J. SeqPrep. 2011. https://githubcom/jstjohn/SeqPrep.
58. Andrews S. FastQC. 2010. http://wwwbioinformaticsbabrahamacuk/projects/fastqc/.
59. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27(6):863–4.
60. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotech. 2011;29(7):644–52.
61. Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 2005;33(suppl 2):W686–9.
62. Boore JL, Macey JR, Medina M. Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol. 2005;395:311–48.
63. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 2010;38(Web Server issue):W7–13.
64. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
65. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17(4):540–52.
66. Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res. 2010;38(suppl 2):W14–8.
67. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol. 1981;17(6):368–76.
68. Huelsenbeck J, Ronquist F. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics. 2001;17:754–5.
69. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22(21):2688–90.
70. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19(12):1572–4.
71. Rambaut A, Drummond AJ. Tracer v1.4. 2007. Available from http://beast.bio.ed.ac.uk/Tracer.

72. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2017;34(3):772–3.
73. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. 2nd international symposium on information theory. Budapest: Akademiai Kiado; 1973. p. 267–81.
74. Drummond A, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214.
75. Duda TF, Bolin MB, Meyer CP, Kohn AJ. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. Mol Phylogenet Evol. 2008;49(3):867–76.
76. Sacco F: I molluschi dei terreni terziarii del Piemonte e della Liguria. Conidae e Conorbidae, vol. XIII. Torino. Italy: Stamperia Reale; 1893.
77. Glibert M. Les Conacea fossiles du Cénozoïque étranger des collections de l'Institut Royal des Sciences Naturelles de Belgique. Institut Royal des Sciences Naturelles de Belgique Mémoire, série 2. 1960;64:1–132.
78. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289–90.
79. Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3(2):217–23.
80. Pybus OG, Harvey PH. Testing macro–evolutionary models using incomplete molecular phylogenies. Proc R Soc Lond B Biol Sci. 2000;267(1459):2267.

3.4. – Chapter IV: "Conotoxin diversity in *Chelyconus ermineus* (Born, 1778) and the convergent origin of piscivory in the Atlantic and Indo-Pacific cones"

Capítulo IV: "Diversidad de conotoxinas en *Chelyconus ermineus* (Born, 1778) y el origen convergente de la piscivoría en los conos atlánticos e indo-pacíficos"

Genome Biology and Evolution, 10 (10): 2643-2662

Se determinó el transcriptoma del conducto de veneno del cono piscívoro del Atlántico, *Chelyconus ermineus* (Born, 1778). El repertorio del veneno de esta especie incluye al menos 378 precursores de conotoxinas, que podrían clasificarse en 33 superfamilias de proteínas conocidas y 22 nuevas (sin asignar), respectivamente. Las superfamilias más abundantes fueron T, W, O1, M, O2 y Z, que representan el 57% de toda la diversidad detectada. Se secuenciaron un total de tres individuos. mostrando una considerable variación intraespecífica: cada individuo tenía muchos precursores de conotoxinas exclusivos, y solo el 20% de todos los péptidos maduros eran comunes a todos los individuos. Tres regiones diferentes (distal, media y proximal con respecto al bulbo del veneno) del conducto del veneno se analizaron de forma independiente. La diversidad (en términos de número de precursores distintos) de superfamilias de conotoxinas aumentó hacia la región distal, mientras que los tránscritos detectados hacia la región proximal mostraron niveles de expresión más altos. Solo las superfamilias A y I3 mostraron una expresión diferencial estadísticamente significativa entre todas las regiones del conducto del veneno. Se detectaron las secuencias pertenecientes a las subfamilias alfa (del "cabal motor") y kappa (del "cabal ightning-strike") de la superfamilia A principalmente en la región proximal del conducto del veneno. Los péptidos maduros de la subfamilia alfa presentan el patrón de espaciado de las cisteínas α4/4, que según se ha demostrado bloquea selectivamente los receptores nicotínicos de acetilcolina del músculo, en última estancia produciendo parálisis. Esta función la realizan los péptidos maduros que tienen un patrón de espaciado de las cisteínas α3/5 en las especies piscívoras de conos de la región del Indo-Pacífico, apoyando así una evolución convergente de la piscivoría en conos.

Conotoxin Diversity in *Chelyconus ermineus* (Born, 1778) and the Convergent Origin of Piscivory in the Atlantic and Indo-Pacific Cones

Samuel Abalde[1], Manuel J. Tenorio[2], Carlos M.L. Afonso[3], and Rafael Zardoya[1,*]

[1]Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, Spain

[2]Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cadiz, Puerto Real, Spain

[3]Fisheries, Biodiversity and Conervation Group, Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, Faro, Portugal

*Corresponding author: E-mail: rafaz@mncn.csic.es.

Accepted: July 28, 2018

Data deposition: Raw RNA seq data:

SRA database: project number SRP139515 (SRR6983161-SRR6983169)

Abstract

The transcriptome of the venom duct of the Atlantic piscivorous cone species *Chelyconus ermineus* (Born, 1778) was determined. The venom repertoire of this species includes at least 378 conotoxin precursors, which could be ascribed to 33 known and 22 new (unassigned) protein superfamilies, respectively. Most abundant superfamilies were T, W, O1, M, O2, and Z, accounting for 57% of all detected diversity. A total of three individuals were sequenced showing considerable intraspecific variation: each individual had many exclusive conotoxin precursors, and only 20% of all inferred mature peptides were common to all individuals. Three different regions (distal, medium, and proximal with respect to the venom bulb) of the venom duct were analyzed independently. Diversity (in terms of number of distinct members) of conotoxin precursor superfamilies increased toward the distal region whereas transcripts detected toward the proximal region showed higher expression levels. Only the superfamilies A and I3 showed statistically significant differential expression across regions of the venom duct. Sequences belonging to the alpha (motor cabal) and kappa (lightning-strike cabal) subfamilies of the superfamily A were mainly detected in the proximal region of the venom duct. The mature peptides of the alpha subfamily had the α4/4 cysteine spacing pattern, which has been shown to selectively target muscle nicotinic-acetylcholine receptors, ultimately producing paralysis. This function is performed by mature peptides having a α3/5 cysteine spacing pattern in piscivorous cone species from the Indo-Pacific region, thereby supporting a convergent evolution of piscivory in cones.

Key words: conotoxin, conopeptide, convergence, transcriptome, Conidae, expression.

Introduction

The family of Conidae (Fleming, 1822 sensu lato) that includes cone snails is well known for their astonishing species diversity (> 800 species; Tucker and Tenorio 2013) as well as for their sophisticated feeding behavior, which includes the production and injection of venom in preys through a specialized harpoon-like radular tooth (Salisbury et al. 2010; Dutertre et al. 2014; Olivera et al. 2015). Although all cone snails were traditionally classified into the single genus *Conus*, recent phylogenetic studies based on morphological (Tucker and Tenorio 2009) and molecular (Puillandre, Bouchet, et al. 2014; Uribe et al. 2017) data supported the split of *Conus* into several lineages, which are ranked either at the family or genus levels, respectively. According to Puillandre, Duda, et al. (2014) and Uribe et al. (2017), the following six genera are recognized: *Profundiconus*, *Californiconus*, *Lilliconus*, *Pygmaeconus*, *Conasprella*, and *Conus*. The latter genus holds most of the species diversity with up to 60 monophyletic groups, either recognized as subgenera (Puillandre, Duda, et al. 2014) or genera (Tucker and Tenorio 2009) depending on the author (herein we will use the taxonomy of Tucker and Tenorio 2009).

The last common ancestor of Conidae likely fed on worms, as most of all living species have been suggested to do (Duda et al. 2001; Puillandre, Bouchet, et al. 2014). During the evolution and diversification of the group, there was suggested to be one diet shift to prey on other snails in the last common ancestor of genera *Calibanus*, *Cylinder*, *Conus*, *Darioconus, Eugeniconus*, and *Leptoconus* (Puillandre, Bouchet, et al. 2014). Instead, phylogenetic analyses suggested at least two diet shifts to prey on fishes as the Atlantic/Eastern Pacific genus *Chelyconus* did not share a most recent common ancestor with Indo-Pacific piscivorous genera: *Phasmoconus*, *Gastridium, Pionoconus*, *Textilia*, *Afonsoconus*, *Embrikena*, and *Asprella* (for the latter three there is no direct observation of prey capture; Duda et al. 2001; Duda and Palumbi 2004; Puillandre, Bouchet, et al. 2014; Olivera et al. 2015). Here, we reconstructed a simplified maximum likelihood (ML) phylogeny of cone snails based on complete mitochondrial (mt) genomes showing the same evolutionary trends in feeding behavior (fig. 1). A remarkable singularity within the group is *Californiconus californicus*, which has a diverse diet including fish, snails, worms, and shrimps (Biggs et al. 2010). The shape and number of barbs of the hollow radular tooth as well as the feeding behavior of cone snails appear to be, at least in some cases, adapted to capturing most efficiently the different types of prey. For instance, some molluscivorous cones make successive injections of radular teeth into the prey (Prator et al. 2014) whereas piscivorous cones show up to three different hunting modes including electrical stunning and tethering of single preys using the proboscis, engulfing of several prey fish at once by the rostrum, and flailing the proboscis around the fish without tethering (Olivera et al. 2015).

The success of a strike relies on the high efficacy of the injected venoms, which readily elicit sedation, paralysis or sensory overload in the prey (Robinson, Li, Bandyopadhyay, et al. 2017). Cone venoms are for the most part complex mixtures of short bioactive peptides termed conotoxins or conopeptides (Lavergne et al. 2013; Robinson and Norton 2014). It is possible to distinguish at least two different components based on their ultimate target: 1) specific peptides, which target voltage-gated and ligand-gated ion channels, neurotransmitter transporters, and receptors in the central and peripheral nervous system of the preys (Olivera 2002; Terlau and Olivera 2004; Lewis et al. 2012; Lavergne et al. 2013) and 2) hormone/neuropeptide-like components, which target neuroendocrine processes in the prey (Safavi-Hemami et al. 2015; Robinson, Li, Bandyopadhyay, et al. 2017). In addition, the venom duct also produces several proteins that are involved in the processing of conotoxins or in enhancing venom activity (Safavi-Hemami et al. 2010, 2014; Hu et al. 2011; Terrat et al. 2012; Barghi et al. 2015b). Each cone species biosynthesizes in the apical secretory cells of the venom duct (Endean and Duchemin 1967) its own venom profile, which shows remarkable intraspecific variability (Davis et al. 2009; Rivera-Ortiz et al. 2011; Rodriguez et al. 2015; Chang and Duda 2016; Peng et al. 2016) as well as striking changes in composition over time within single individuals (Dutertre et al. 2010; Prator et al. 2014). Moreover, it has been shown that conotoxin expression is regionalized along the venom duct (Garrett et al. 2005; Tayo et al. 2010; Hu et al. 2012; Jin et al. 2015; Prashanth et al. 2016) and that the proximal and distal regions (with respect to the bulb) of the venom duct of several cones produce distinct defense- and predation-evoked conotoxin cocktails, respectively (Dutertre et al. 2014).

Conotoxins are generally synthesized as precursors with a typical three domain structure including: 1) a highly conserved hydrophobic N-terminal signal region, which guides the conotoxin precursor to the endoplasmic reticulum and the cellular secretory pathway; 2) an intervening, moderately conserved propeptide region, which for some conotoxins participates in secretion, post-translational modification and folding (Bandyopadhyay et al. 1998; Conticello et al. 2003; Buczek et al. 2004); and 3) a C-terminal region, which constitutes the mature functional peptide (Woodward et al. 1990; Kaas et al. 2010). The conserved sequence profiles of the signal region have been used to classify conotoxin precursors into 40–50 protein superfamilies, originally named with alphabet letters (Terlau and Olivera 2004; Corpuz et al. 2005; Kaas et al. 2010; Puillandre et al. 2012; Lavergne et al. 2013; Robinson and Norton 2014; Li et al. 2017). Furthermore, sequence comparison of mature peptides revealed up to 26 conserved cysteine frameworks named with roman numbers, which generally correlate with conotoxin precursor superfamilies (Lavergne et al. 2015), although most of these protein superfamilies have been shown to have more than one cysteine framework (Terlau and Olivera 2004; Corpuz et al. 2005; Luo et al. 2006; Kaas et al. 2010; Puillandre et al. 2010, 2012; Robinson and Norton 2014; Lavergne et al. 2015). Notably, some mature peptides are cysteine-poor or completely lack cysteines (Puillandre et al. 2012).

The renowned conotoxin hyperdiversity, which is typical for gene families that mediate interactions between organisms (Conticello et al. 2001; Barghi et al. 2015a), can be explained by a combination of several (evolutionary) processes including: 1) extensive gene duplication (Duda and Palumbi 1999; Espiritu et al. 2001; Puillandre et al. 2010; Chang and Duda 2012); 2) high mutation rates and diversifying selection of the mature domain (Conticello et al. 2001); 3) recombination events (Espiritu et al. 2001; Wu et al. 2013); and 4) variable peptide processing (Dutertre et al. 2013) and post-translational modifications (Lu et al. 2014; Peng et al. 2016).

The advent of high-throughput sequencing techniques, and in particular of RNA sequencing (first using the 454 GS FLX Titanium and currently the Illumina HiSeq platforms) has produced a quantum leap in the characterization of whole conotoxin precursor repertoires (Prashanth et al. 2012; Barghi et al. 2015b) compared with the more traditional sequencing

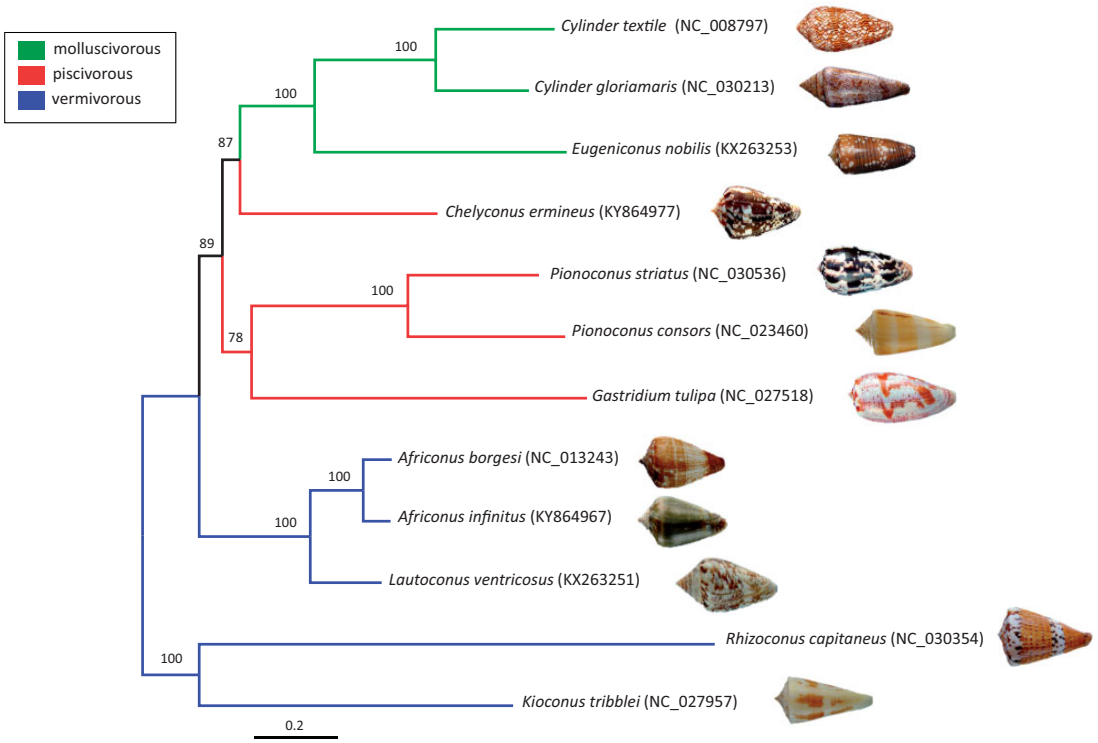


Fig. 1.—Simplified ML phylogeny of cone snails based on complete mitochondrial genomes (concatenated 13 protein-coding genes plus two rRNA genes analyzed at the nucleotide level). The evolution of diet is mapped onto the phylogeny. Bootstrap values are indicated above each node. Scale bar indicates substitutions/site. GenBank accession numbers are indicated for each species mt genome.

of individual cDNA clones (Pi, Liu, Peng, Jiang, et al. 2006; Pi, Liu, Peng, Liu, et al. 2006; Liu et al. 2012; Lu et al. 2014). RNA sequencing is highly sensitive and even rare transcripts with low expression levels can be identified (Barghi et al. 2015b). Therefore, several recent studies have determined the complete conotoxin diversity of 1) piscivorous species such as *Textilia bullatus* (Hu et al. 2011), *Pionoconus consors* (Terrat et al. 2012; Violette et al. 2012), *Gastridium geographus* (Hu et al. 2012; Dutertre et al. 2014; Safavi-Hemami et al. 2014), and *Pionoconus catus* (Himaya et al. 2015); 2) molluscivorous species such as *Conus marmoreus* (Dutertre et al. 2013; Lavergne et al. 2013), *Darioconus episcopatus* (Lavergne et al. 2015), and *Cylinder gloriamaris* (Robinson, Li, Lu, et al. 2017); and 3) vermivorous species such as *Puncticulis pulicarius* (Lluisma et al. 2012), *Rhizoconus miles* (Jin et al. 2013), *Kioconus tribblei* (Barghi et al. 2015b), *Dendroconus betulinus* (Peng et al. 2016), and the congeneric species *Turriconus andremenezi* and *Turriconus praecellens* (Li et al. 2017), among others. Altogether, these transcriptome studies show that each cone species produces at least 100–400 different conotoxin precursors, and have been very successful in discovering new superfamilies and cysteine frameworks (Barghi et al. 2015b; Lavergne et al. 2015; Peng et al. 2016; Li et al. 2017). The emerging general pattern is that several gene superfamilies are widespread among cone species, although with different degrees of expansion, whereas others are restricted to a few lineages (Duda and Remigio 2008; Puillandre et al. 2012).

To further contribute to the cataloguing of conotoxin diversity in the main lineages of cone snails, we characterized the full transcriptome of *Chelyconus ermineus* (Born 1778). This is a cone species, which can be found on both shores of the Atlantic Ocean and feeds on fishes. Together with *Chelyconus purpurascens* from the Eastern Pacific region, it forms a clade, which according to the phylogeny of Conidae (Duda and Palumbi 2004; Puillandre, Bouchet, et al. 2014; fig. 1) may underwent a shift to piscivory independent to the one occurred in the ancestor of Indo-Pacific piscivorous genera. Thus far, the study of conotoxins in *C. ermineus* has been limited to the identification of few mature peptides (Martinez et al. 1995; Jacobsen et al. 1997; Barbier et al. 2004; Rivera-Ortiz et al. 2011; Echterbille et al. 2017) and some conotoxin precursors belonging to the A, B1, and O1 superfamilies (Duda and Palumbi 2004; Gowd et al. 2008; Puillandre et al. 2010; 18 entries in ConoServer, Kaas et al. 2012). Notably, however, some of the mature conotoxins of *C. ermineus* such as EVIA are among the few peptides whose functional activity (Barbier et al. 2004) and tertiary structure (Volpon et al. 2004) have been determined experimentally. There are a few more conotoxin precursors identified in the closely related *C. purpurascens* belonging to the A, B1, M, O1, and T superfamilies (e.g., Shon et al. 1995, 1998; Duda and

Palumbi 2004; 45 entries in ConoServer, Kaas et al. 2012). By sequencing the transcriptome of *C. ermineus*, the first one of an Atlantic cone species, we aimed: 1) to catalogue the diversity of conotoxin precursors in this species and classify them into superfamilies; 2) to identify other proteins that are transcribed in the venom duct and are potentially involved in the processing of conotoxins or in enhancing venom activity; 3) to estimate intraspecific variation of conotoxin precursors; 4) to determine differences in the spatial distribution of conotoxin precursors along the distal, medium, and proximal regions of the venom duct; 5) to quantify the expression levels of conotoxin genes in the different individuals and along the venom duct; and 6) to compare the venom composition of Atlantic and Indo-Pacific piscivorous cones, and to identify putative differences with the venoms of cones preying on snails and worms.

Materials and Methods

Sampling and RNA Extraction

Three adult specimens of *C. ermineus* were captured, respectively, in Boa Vista (CVERM3; hereafter ERM1), Sal (CVERM13; hereafter ERM2), and Santa Luzia (CV1446; hereafter ERM3) islands in Cabo Verde with corresponding permits (table 1). Each individual, in a resting stage, was extracted from the shell and dissected to remove the venom duct, which was excised into three equal parts: proximal, medium, and distal with respect to the venom bulb (following Tayo et al. 2010). These fragments were stored in 1 ml RNA*later* (Invitrogen, Life Technologies), first at 4°C and for the long term at −20°C.

For RNA extraction, each venom duct portion was incubated independently in a 2 ml eppendorf with 500 µl of TRIzol LS Reagent (Invitrogen, Life Technologies) and grinded with ceramic beads in a Precellys Evolution tissue homogenizer. The solution was mixed with 100 µl of chloroform. After centrifugation (12,000 x g for 15 minutes at 4°C), the aqueous phase was recovered and RNA precipitated with 250 µl of isopropanol and stored overnight at −80°C. The Direct-zol RNA miniprep kit (Zymo Research, Irvine) was used to purify total RNA (5–15 µg) following manufacturer's instructions.

Library Preparation and Sequencing

Dual-indexed cDNA libraries (307–345 bp insert average size) for each sample were constructed using the TruSeq RNA Library Prep Kit v2 (Illumina, San Diego) and following manufacturer's instructions at Sistemas Genómicos (Valencia, Spain). Briefly, the poly(A)+ mRNA fraction was isolated using oligo-$(dT)_{25}$ magnetic beads. Subsequently purified mRNA was chemically fragmented prior to reverse transcription and the construction of the cDNA library. The quality of the libraries was analyzed with the TapeStation 4200, High Sensitivity assay; the quantity of the libraries was determined by real-time PCR in LightCycler 480 (Roche). The pool of libraries (including other cone species for different projects) was split into different lanes and sequenced by paired-end sequencing (100×2) in an Illumina HiSeq2500 (two flowcells) following standard procedures at Sistemas Genómicos (Valencia, Spain).

Assembly

The reads corresponding to the different regions of the venom duct and individuals were sorted using the corresponding library indices. Adapter sequences were removed using SeqPrep (St John 2011). Assembly was performed using the TRUFA webserver (Kornobis et al. 2015). Briefly, the quality of the sequencing was checked using FastQC v.0.10.1 (Andrews 2010). Ends of reads were trimmed (PHRED < 30) and resulting trimmed reads were filtered out according to their mean quality scores (PHRED < 20) using PRINSEQ v.0.20.3 (Schmieder and Edwards 2011). This step also ensured minimizing cross-contamination resulting from potential index misassignment, as this tends to be associated to low quality scores (Wright and Vetsigian 2016). Filtered reads were used for de novo assembly of trancriptomes with Trinity r2012-06-08 (Grabherr et al. 2011) with default settings (minimum contig length: 200; sequence identity threshold: 0.95). The transcriptome raw reads produced in this project have been deposited at the NCBI SRA database under accession SRP139515 (see also table 1).

Prediction and Annotation of Conotoxin Precursors and Associated Proteins

The sequences of all conotoxin precursors and associated proteins of cone venoms available in GenBank release 217 (Benson et al. 2005), Uniprot release 2016_11 (Uniprot Consortium 2017), and ConoServer release 12-26-2016 (Kaas et al. 2012) were downloaded in December 26, 2016 to construct a local reference database. Redundant entries from the three databases were removed. Subsequently, BLASTX was used to identify those sequences encoding putative conotoxin precursors and associated proteins (with an E-value of 1e-5) among the assembled contigs by similarity searches against the reference database. These sequences were translated into amino acids using the universal genetic code and manually inspected, in order to discard false positives (hits not corresponding to canonical conotoxins) or assembly artifacts (due to indels that interrupt open reading frames). Duplicate and highly truncated (>55% of the estimated total length of a precursor) sequences were removed to produce the final working list of conotoxin precursors and associated proteins of *C. ermineus* (provided in supplementary table 1, Supplementary Material online). The three domains of the predicted conotoxin precursors and the cysteine frameworks of the mature peptides were identified using the Conoprec tool (Kaas et al. 2012). Assignment of amino acid

Table 1
Specimens of *Chelyconus ermineus* Analyzed in This Study and Main Statistics of Illumina Sequencing and Assembly

Specimen	Voucher ID MNCN	Island	Segment	SRA Accesion No.	# Raw Reads	# Clean Reads	# Contigs	% Mapping[a]	# Conotoxins	% Mapping[b]
ERM1	15.05/80980	Boa Vista	Proximal	SRR6983168	13,023,114	12,882,970	64,233	92	59	61
ERM1	15.05/80980	Boa Vista	Medium	SRR6983169	25,823,481	25,541,087	69,836	83	75	70
ERM1	15.05/80980	Boa Vista	Distal	SRR6983166	27,702,513	27,160,103	119,384	88	117	17
ERM2	15.05/80013	Sal	Proximal	SRR6983167	26,754,509	26,754,509	52,506	69	75	69
ERM2	15.05/80013	Sal	Medium	SRR6983164	26,986,678	26,986,220	57,887	76	89	63
ERM2	15.05/80013	Sal	Distal	SRR6983165	26,107,666	26,107,195	73,809	91	109	40
ERM3	15.05/78606	Santa Luzia	Proximal	SRR6983162	27,163,849	27,163,368	49,195	76	71	78
ERM3	15.05/78606	Santa Luzia	Medium	SRR6983163	31,223,312	31,222,733	68,103	92	90	58
ERM3	15.05/78606	Santa Luzia	Distal	SRR6983161	31,717,505	31,716,948	71,785	56	83	58

[a]Percentage of clean reads that map onto assembled contigs.
[b]Percentage of clean reads that map onto assembled conotoxin precursors.

sequences to different superfamilies was based on the two highest scoring full-length conotoxin precursor hits in the BLAST results and taking into account the percentage of sequence identity (>90%) to the highly conserved signal region (Robinson and Norton 2014; Barghi et al. 2015b). Further refinement of the superfamily assignment was achieved by aligning conotoxin precursor amino acid sequences of *C. ermineus* to selected canonical representatives of each superfamily using Mafft v7 (Katoh and Standley 2013) with default parameters (see supplementary file 1, Supplementary Material online). This step revealed important diversity (i.e., presence of potential paralogs) at the propeptide domain within several superfamilies, which was further analyzed. All *C. ermineus* conotoxin precursor amino acid sequences are deposited (as nucleotide sequences) in GenBank under accession numbers MH360289–MH360712.

Phylogenetic Analyses of the M and T Superfamilies

In order to infer the evolutionary origin of cysteine-poor conotoxins, we performed phylogenetic analyses of the M and T conotoxin precursor superfamilies, which have both cysteine-rich and cysteine poor members. Concatenated amino acid alignments of the signal and propeptide domains of the M and T superfamilies, respectively, were constructed using Mafft v7 (Katoh and Standley 2013) with default parameters. Phylogenetic relationships were inferred using ML (Felsenstein 1981) with PhyML v3.0 (Guindon et al. 2010) with default settings in the ATGC platform (http://www.atgc-montpellier.fr/phyml/; last accessed September 05, 2018) and using the smart model selection option. Statistical support was assessed with 1,000 bootstrap pseudoreplicates (BP).

Expression Analyses

Approximate relative expression levels were estimated by mapping only clean reads, back to all assembled contigs of *C. ermineus*. TPM (transcripts per kilobase million), which normalize for gene length and sequencing depth, were estimated with the RSEM package (which uses the mapper Bowtie 2; Langmead and Salzberg 2012) included in Trinity r2012-06-08 (Grabherr et al. 2011). In addition, we run the EBSeq software (Leng et al. 2013) as implemented in Trinity to estimate the posterior probability of being differentially expressed (PPDE), setting the False Discovery Rate (FDR) at 0.95, of conotoxins as a whole and of each of the different superfamilies along the different regions of the venom duct using the three individuals as biological replicates.

Reconstruction of Cone Snail Phylogeny

In order to determine diet shifts during the evolutionary history of cone snails, a simplified phylogeny was reconstructed using ML based on complete mt genomes (13 protein-coding and two rRNA genes) available in GenBank. Protein-coding genes were individually aligned using TranslatorX (Abascal et al. 2010), which generates a nucleotide alignment based on corresponding deduced amino acid alignments. The rRNA genes were aligned using Mafft v7 (Katoh and Standley 2013). All ambiguously aligned positions were removed using GBlocks v.0.9.1b (Castresana 2000) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous nonconserved positions: 8; minimum block length: 10; gaps in final blocks: no. Finally, the different single alignments were concatenated using Geneious 8.1.8.

The best-fit partition scheme and models of substitution for the data set were identified using PartitionFinder (Lanfear et al. 2012) with the Bayesian Information Criterion (Schwarz 1978). The following partitions were tested: all genes together, all genes arranged in subunits (*atp*, *cob*, *cox*, *nad*, and *rrn*), and all genes separated (except atp6-atp8 and nad4-nad4L). In addition, we also tested separately the three codon positions in the protein-coding genes. The best partition scheme was the one considering each codon position separately, all protein-coding genes concatenated, and

rRNA genes concatenated. For each partition, the selected best-fit model was GTR +I + G.

In order to reconstruct the ML tree, we used RAxML-HPC2 on XSEDE 8.2.10 (Stamatakis 2014) as implemented in the CIPRES Science Gateway v 3.3 (http://www.phylo.org/; last accessed September 05, 2018) with the rapid hill-climbing algorithm and 1,000 BP. The outgroups were *Rhizoconus capitaneus* and *K. tribblei* based on Puillandre et al. (2014).

Results

Sequencing and Assembly

A total of nine samples were sequenced corresponding to three regions (distal, medium, and proximal) of the venom duct of three individuals (ERM1-3) of *C. ermineus*. The main statistics associated to the sequencing and assembly procedures are summarized in table 1. Sequencing generated between 13 and 32 million raw reads per sample. Most (99–100%) of the reads were kept as clean after adapter and quality trimming. The number of assembled contigs varied between 49,195 and 119,384 with a mean of 69,637.6 per sample. Mapping of clean reads onto assembled contigs indicated that on an average 80% of the reads were used for further analyses (table 1). After BLASTX searches against a local reference database, the number of distinct (with at least one amino acid difference) putative conotoxin precursor sequences per sample (i.e., venom duct portion of an individual) varied between 59 and 117 (table 1). The majority of these sequences were full-length but a few were slightly truncated at the N- or C-terminus (see supplementary table 1, Supplementary Material online). Mapping of clean reads onto assembled transcripts indicated that conotoxin production made up on an average 57% of the transcriptome in the venom duct, although there was important variability among venom duct regions within the same individual (table 1).

Intraspecific Variation in Venom Composition

The sequences of the conotoxin precursor transcripts expressed in the venom duct of individuals ERM1-3 of *C. ermineus* inhabiting three different islands (Boa Vista, Sal, and Santa Luzia) of the archipelago of Cabo Verde were determined (fig. 2). The total numbers of inferred conotoxin precursors in each specimen were 161, 176, and 144, respectively. Of these, 25 were found in all specimens, nine were common to ERM1 and ERM2 (but not found in ERM3), 18 were shared by ERM1 and ERM3 (not present in ERM2), and 26 were common to ERM2 and ERM3 (and not found in ERM1). The number of peptides common to all three specimens rose up to 33 when only differences in the functional mature peptide were taken into consideration (fig. 2). In such analyses, the numbers of sequences exclusive to each specimen were 96, 104, and 56, respectively (fig. 2). We also estimated intraspecific diversity taking into account putative allele variation by clustering together sequences, which diverged in one or less, two or less, and three or less amino acids, respectively. The numbers of conotoxin precursor sequences common to all three specimens rose up to 44, 48, and 53, respectively (supplementary fig. 1, Supplementary Material online).

Diversity of Conotoxin Precursor Sequences in *C. ermineus*

Of the 422 distinct transcripts related to venom activity identified as produced in the venom duct of the three individuals of *C. ermineus*, a total of 296 could be assigned (based on the signal region sequence) to 33 known conotoxin precursor superfamilies already described in other cone venom ducts (fig. 3). In addition, 82 conotoxin precursor sequences were grouped, using reciprocal BLASTs and taking into account a 90% identity threshold per superfamily, into 22 unassigned conotoxin superfamilies, not formally described in other cone species but also present in some of them (see supplementary file 1, Supplementary Material online). Finally, 44 peptides corresponded to six associated protein families (see supplementary table 1 and file 1, Supplementary Material online). All but one (alpha conotoxin EI; P50982) previously reported conotoxin precursors and mature peptides from *C. ermineus* (Martinez et al. 1995; Jacobsen et al. 1997; Barbier et al. 2004; Duda and Palumbi 2004; Gowd et al. 2008; Puillandre et al. 2010; Rivera-Ortiz et al. 2011) were detected (see supplementary file 2, Supplementary Material online). Homologs to conotoxin precursors and mature peptides from *C. purpurascens* were identified as well (Shon et al. 1995, 1998; Duda and Palumbi 2004). The six most diverse conotoxin precursor superfamilies (in terms of the number of distinct members) were T, W, O1, M, O2, and Z, accounting for 57% of all observed diversity (fig. 3). Conversely, several of the conotoxin precursor superfamilies were restricted to only one or two representatives (e.g., A2, E, J, K, P, R, and several unassigned superfamilies).

Some of the inferred mature domains in the conotoxin precursors showed no cysteine framework. In some cases, these mature peptides belonged to superfamilies exclusively formed by members without cysteines, such as the W and Z superfamilies. In other cases, mature conotoxins with and without framework were grouped together within the same superfamily (fig. 3). Notably, the M and T superfamilies had both types of mature conotoxins. Phylogenetic trees of both superfamilies were reconstructed based on the amino acid sequences of the signal and propeptide regions, and allowed distinguishing several paralog groups within each conotoxin precursor superfamily (fig. 4). While mature conotoxins without cysteine framework form a distinct paralog group within the M superfamily, they seem to have originated independently and recurrently in the different paralogs within the T superfamily (fig. 4).

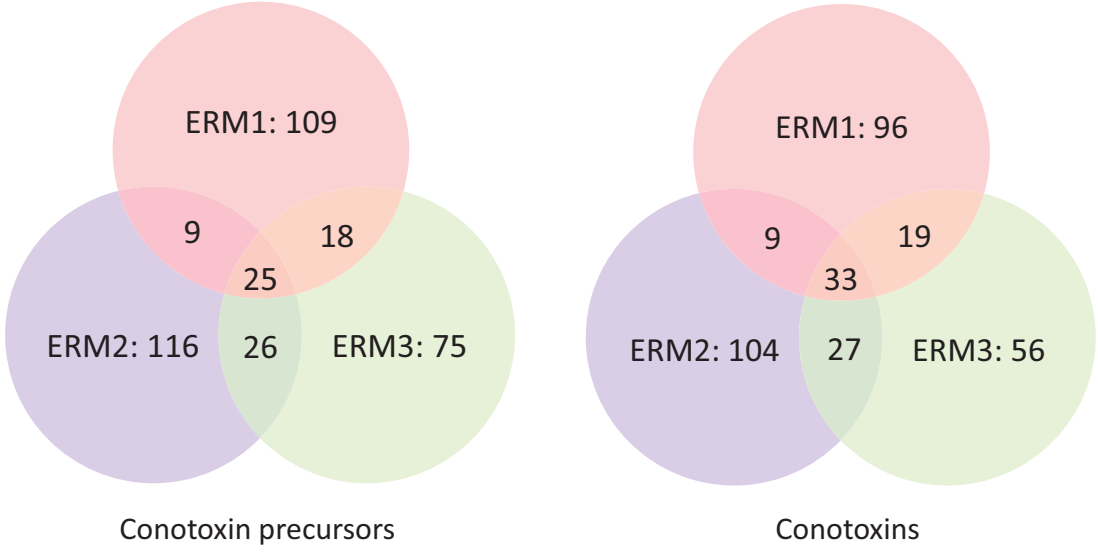


Fig. 2.—Distinct conotoxin precursors (left) and mature peptides (right) identified in the three analyzed individuals of *Chelyconus ermineus*.

Within unassigned superfamilies, most precursors have the canonical three-domain structure (see signal sequences in table 2). However, an interesting case was that of several precursors (unassigned superfamilies 16–22), which could not be assigned to any known superfamily because they lacked a signal region, but had a mature peptide, which could be confidently aligned to the mature peptides normally associated to the O3 and T superfamilies in other cone species (see supplementary file 1, Supplementary Material online). In trying to assign these sequences to known superfamilies, we used hidden Markov model searches as implemented in Conodictor (Koua et al. 2012) but retrieved no significant hit.

Differential Spatial Distribution of Conotoxin Precursor Transcripts along the Venom Duct

Changes in the diversity of conotoxin precursors were analyzed along the three regions (distal, medium, and proximal) of the venom duct (fig. 5). The numbers of distinct conotoxin precursor sequences found in the distal, medium, and proximal regions were 190, 162, and 130, respectively. Up to 30–33 (depending on the individual) precursor transcripts were expressed throughout the venom duct (fig. 5). Of these, seven were common to the three individuals (not shown). A total of 33–64, 24–34, and 15–25 conotoxin precursors were found exclusively in the distal, medium, and proximal portions, respectively (fig. 5). Most conotoxin precursor superfamilies were identified in the three regions of the venom duct with the exception of E, K, and Thyrostimulin β, which were only detected in the distal portion; L and Q, which were missing in the proximal portion; A2, which was absent in the medium portion; and J, which was missing in the distal portion (fig. 5). In general, the diversity of members of the different conotoxin precursor superfamilies was relatively uniform across venom duct segments. However, the H, I2, I3, L, M, O1, Q, T, and W superfamilies, and the hormone conophysin showed more diversity toward the distal portion; A2 and B2 superfamilies had more diversity toward the proximal portion; and most of the members of Z superfamily were detected in the medium portion (fig. 5).

Expression of Conotoxin Precursor Transcripts along the Venom Duct

The relative expression levels of the different conotoxin precursor superfamilies along the distal, medium and proximal regions of the venom duct in the three individuals were estimated as TPMs, thus normalizing for gene length and sequencing depth (fig. 6). Expression levels varied extensively among individuals hindering the inference of expression patterns. TPM values were declared reliable when they were of similar level in at least two out of the three individuals. Taking this into consideration, the most expressed conotoxin precursor superfamilies were: A, which accounted for much of the conotoxin precursor expression in the medium and proximal fractions, O2, which was expressed abundantly in the distal region; and O1, which showed expression throughout the venom duct but particularly in the distal and medium portions (fig. 6*A*). A second batch of midexpressed superfamilies included: T, which showed higher values in the distal region; M, which was mostly expressed in the proximal region; and S with expression levels higher in the medium and distal regions (fig. 6*A*). The remaining conotoxin precursor superfamilies had much lower expression levels, mostly concentrated in the medium and distal regions of the venom duct with the exception of O3 members, which were higher expressed in the medium and proximal regions (fig. 6*B*). We tested, within a Bayesian framework, whether any of the conotoxin

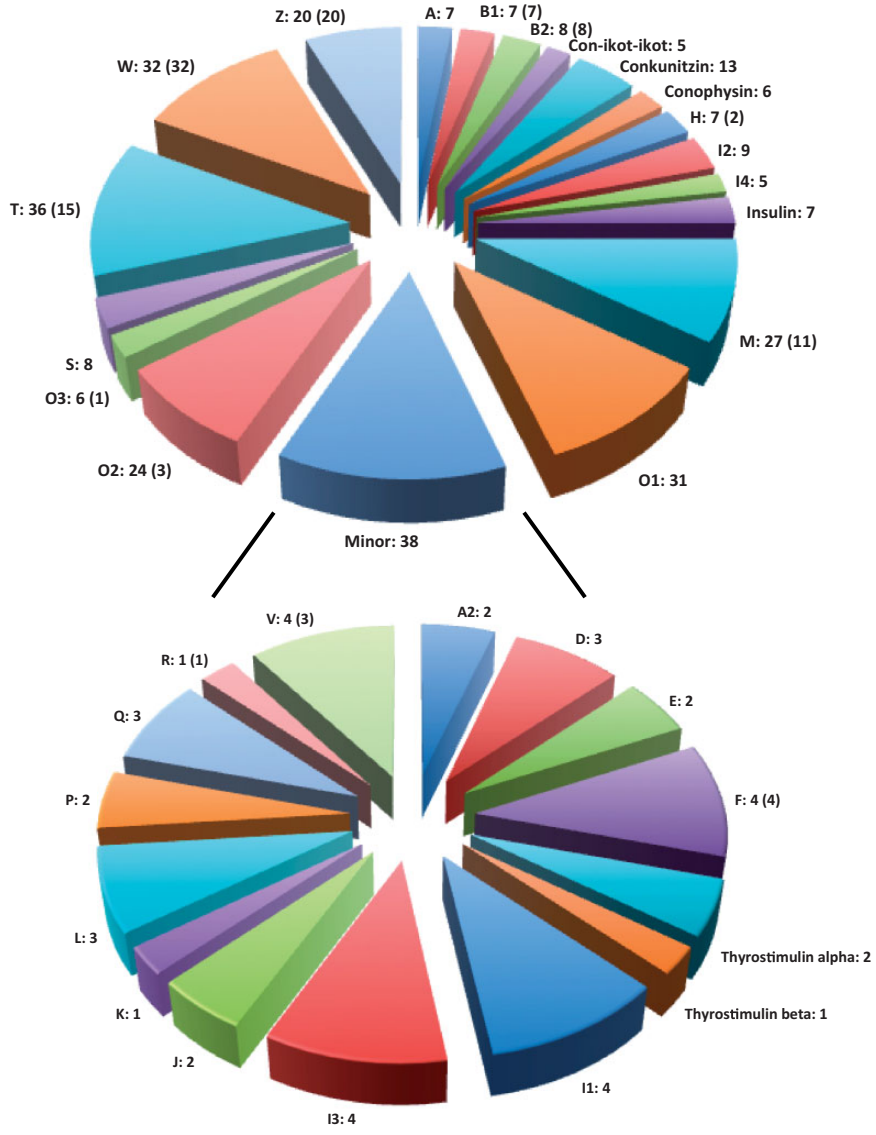


Fig. 3.—Distribution in superfamilies of the 296 identified conotoxin precursors. Numbers in parentheses indicate the number of mature peptides without cysteine framework.

precursor superfamilies had differential expression among regions of the venom duct using the individuals as biological replicates (table 3). Only the A and I3 superfamilies showed significant posterior probabilities. Within the A superfamily, paralogs A-1 (named alpha 4/4; see below) and A-2 (name kappa; see below) had PPDE values of 0.94 and 1, respectively (table 3). Within each paralog, only conotoxin precursors Cerm_405 (A-1) and Cerm_342 (A-2) both detected in ERM3 showed significant differential expression in the proximal and distal portions, respectively (table 3).

Altogether, conotoxin precursor transcripts showed differential expression along the venom duct (table 3), and accounted for ~60% and 70% of the overall expression in the medium and proximal regions, respectively (fig. 6*C*). In contrast, conotoxin precursor expression in the distal region was restricted to 30% whereas other (house-keeping) transcripts dominated (fig. 6*C*). Ferritin showed important levels of expression, being preferentially expressed in the distal region (fig. 6*C*).

Member Diversity of the a Superfamily across Cone Species

Given the importance of the A superfamily in the overall expression of the venom duct of *C. ermineus*, we performed a more detailed analysis of the diversity of its members across species (table 4; see also classification of the 288 A superfamily conotoxin precursors available in ConoServer in supplementary file 3, Supplementary Material online). Within the A superfamily (Santos et al. 2004), there are two main groups of conotoxins with very distinct structure and function (Azam

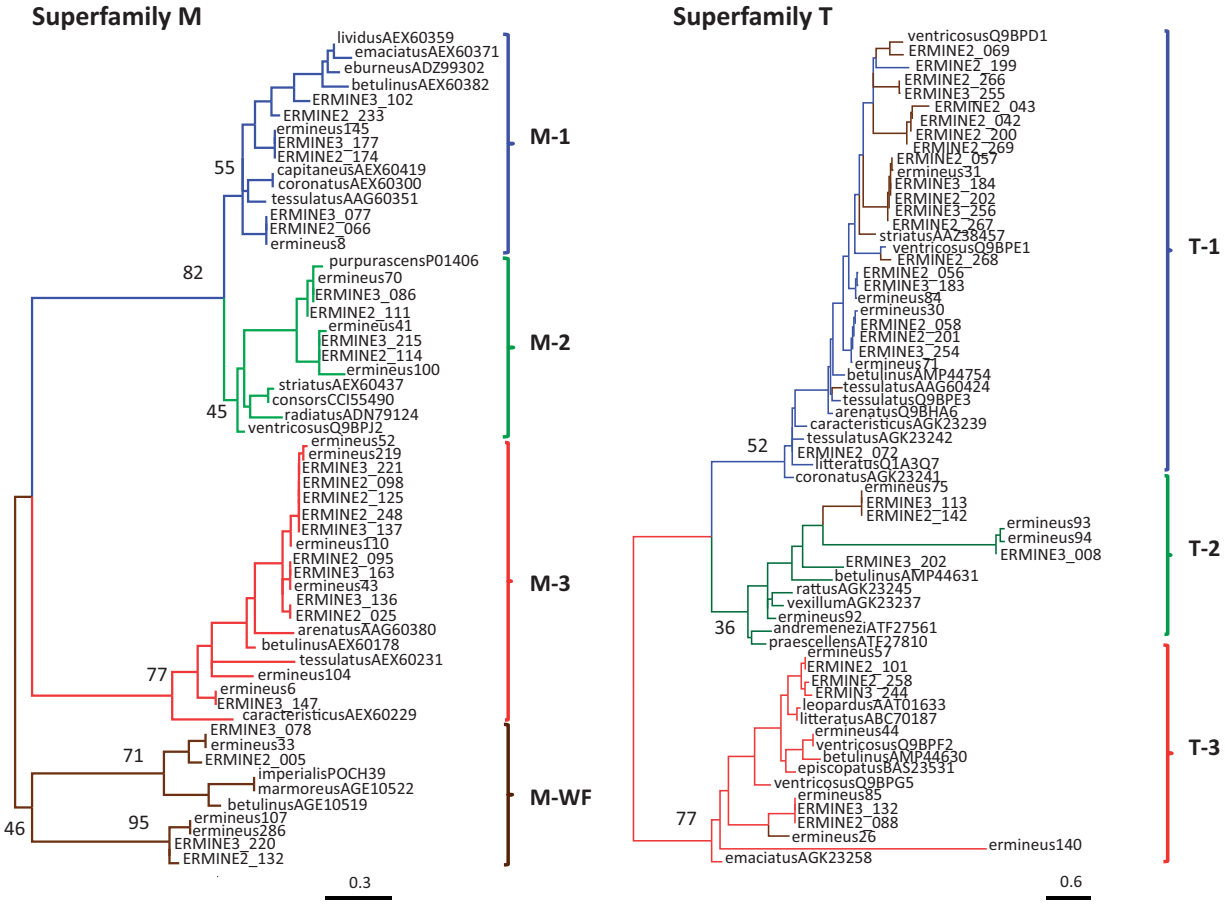


Fig. 4.—Reconstructed ML phylogenies of the M and T superfamilies, recovering several clades (paralogs; in different colors) and indicating the differential evolutionary origin of cysteine-poor (in brown; WF meaning without framework) mature peptides. Bootstrap values of main clades are indicated. Scale bar indicates substitutions/site. GenBank accession numbers are indicated after each species except for *C. ermineus*.

and McIntosh 2009; Puillandre et al. 2012; Robinson and Norton 2014). One group (alpha, α) has cysteine framework I and selectively target nicotinic-acetylcholine receptors (nAChRs), ultimately inhibiting neuromuscular transmission (Azam and McIntosh 2009). The other group (Kappa, κ) has cysteine framework IV and their members target preferentially K^+ channels, producing an excitatory effect (Robinson and Norton 2014). The Kappa subfamily is, thus far, only found in piscivorous cones (table 4; Santos et al. 2004). Within framework IV, the most frequent cysteine spacing pattern is cc7c2c1c3c (Puillandre et al. 2012), which is shared by *C. ermineus*, several *Pionoconus* species, and *Embrikena* (table 4). In contrast, specific cysteine spacing variations have been reported for *C. purpurascens*, *Gastridium*, and *Textilia* (table 4). Within the alpha subfamily, the most frequent cysteine spacing pattern is cc4c7c (named α4/7; Puillandre et al. 2012). Vermivorous (including *Rhombiconus imperialis*, which has a strict diet on amphinomids) and molluscivorous species show important diversity of α4/7 conotoxins whereas this subfamily is represented by only 1–2 members in most piscivorous species but *G. geographus* (table 4; a striking exception is *Asprella*, which has been proposed to be piscivorous, although not based on direct evidence, and has seven α4/7 conotoxin precursors, a pattern typical of vermivorous or molluscivorous cones). A second frequent cysteine spacing pattern is cc3c5c (named α3/5; Puillandre et al. 2012), which is almost exclusive of piscivorous species, and particularly diverse in *G. geographus*, *P. consors*, and *P. striatus* but not found in *Chelyconus* (table 4). Finally, a third cysteine spacing pattern, which is also relatively frequent, is cc4c4c (named α4/4; Puillandre et al. 2012). It is particularly diverse in the piscivorous genera *Chelyconus* and *Textilia*, and in the vermivorous genera *Virgiconus emaciatus* and *Calamiconus quercinus* (table 4). Interestingly, it has been also reported with lower diversity in the piscivorous genus *Pionoconus* and in the putative piscivorous genus *Afonsoconus* (Puillandre et al. 2012).

Discussion

The cocktail of bioactive peptides produced in the venom duct of a cone snail is a complex mixture aimed at paralyzing specific preys and deterring predators (Dutertre et al. 2014).

Table 2

New Signal Sequences in Conotoxin Precursors of *Chelyconus ermineus*

Unassigned Superfamily	Signal	Cysteine Framework	Also Found In:	Best-Hit Known Superfamily		
				% Coverage	% Identity	Superfamily
1	MRFYMLLAVALLLTSVMS	VI/VII	–	66	75	O2 (Q9NDA7)
2	MRFLLFLCIAVLLTSFRETEA	VI/VII	*betulinus*	85	35	T (BAS25421)
3	MKLSMMFILSLVLTLSMTDG	XIV	*praecellens*	90	67	L (ABC74975)
4	MKLSVMVIVLVLAMAFTPGLL	XIV	*betulinus*	80	67	L (ABC74975)
5	MNFSVMFILALVLTLSMTDA	XIV	*betulinus, praecellens*	90	61	L (ABC74975)
6	MKVVVVLLAVLVAASA	XIV	*betulinus*	100	56	Hyaluronidase (C0HKM3)
7	MCLSTMPSVILMMVLMFAFDNVDG	IX	*betulinus, imperialis*	58	57	P (ATF27727)
8	MKLFMFTAIIFTMASTTVT	VIII	*andremenezi, caracteristicus*	78	53	O1 (Q5K0B8)
9	MSKTGLVLVVLYLLSSPVNL	XIII	*miles, praecellens, andremenezi*	85	60	M (ACV87169)
10	MKFTTFVMVLMAAVLLTSILETEA	VI/VII	*betulinus, praecellens*	54	62	Con-ikot-ikot (BAO65537)
11	MEFRRLVTVGLLLTLVMSTDS	IX	*betulinus*	47	88	Insulin (AOF40168)
13	MLSMLAWTLMTAMVVMNAKS	$(C)_{12}$	*praecellens, gloriamaris*	55	73	O1 (BAS22670)
14	MNMRMTIIVFVVVATAATVVGST	CC-C-C-C-C-C-CC-C-C-C-C	*lenavati, tribblei*	100	61	Con-ikot-ikot (P0CB20)
15	MSVVYCKPSVPVDSVSSNFCVRGPDNGHQA	VI/VII	–	40	86	T (Q9BPD9)

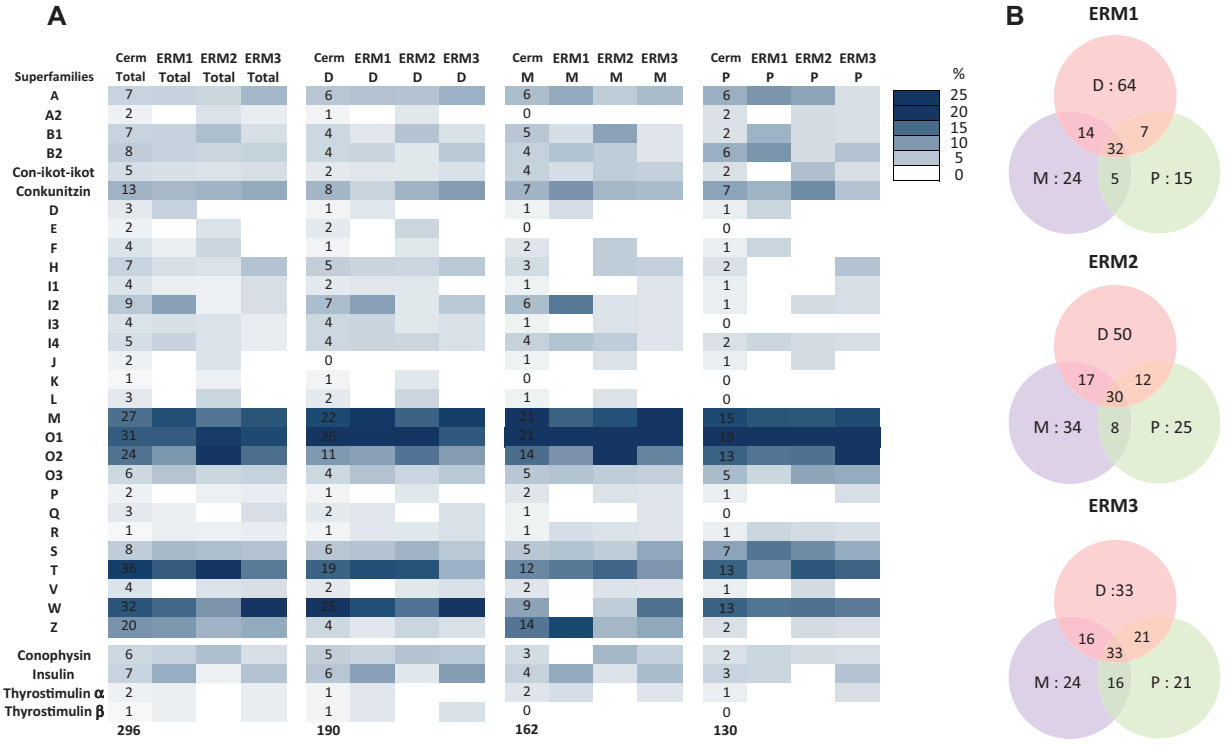
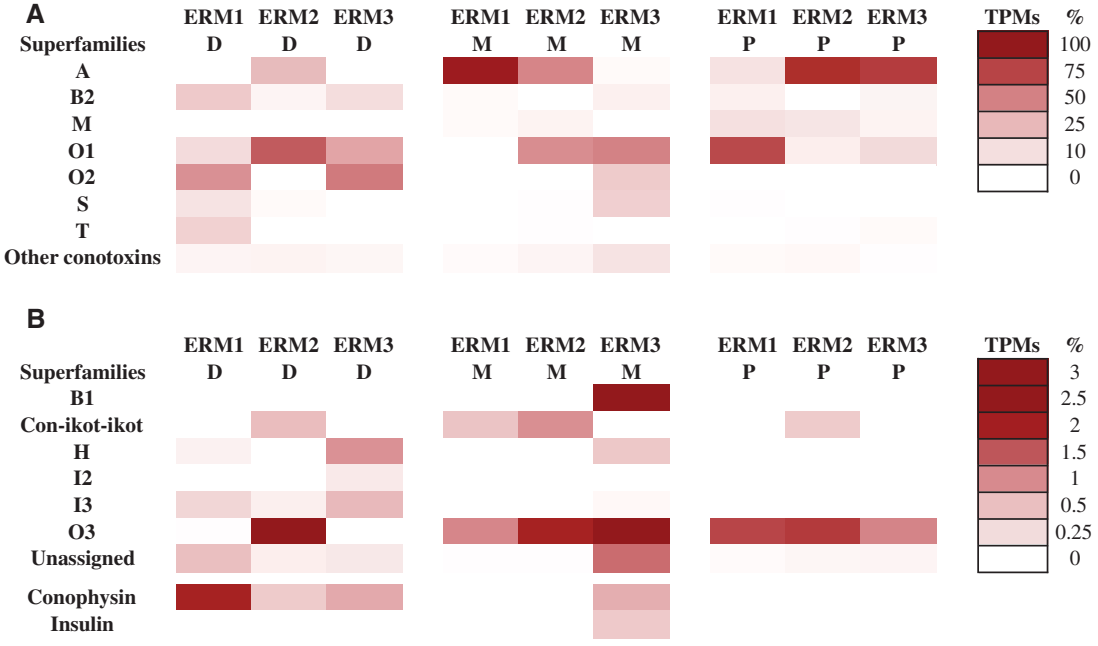


FIG. 5.—Distribution of conotoxin precursor superfamily diversity (number of members) along the distal (D), medium (M) and proximal (P) portion of the venom duct (with respect to the venom bulb) is shown in panel (*A*). Number of common members in the three analyzed individuals per venom duct region is depicted in panel (*B*). Cerm code indicates unique peptide sequences after considering the three analyzed individuals (see supplementary table 1, Supplementary Material online).



C

	ERM1	ERM2	ERM3	ERM1	ERM2	ERM3	ERM1	ERM2	ERM3
	D	D	D	M	M	M	P	P	P
Others	78.356	55.851	69.227	28.144	37.243	51.307	26.676	30.860	29.893
Conodipine	0.028	0.003	0.074	0.344	0.023	0.001	0.000	0.048	0.012
Conohyaluronidase	0.001	0.291	0.000	0.207	0.430	0.523	0.688	0.265	0.000
Ferritin	3.112	0.848	1.774	0.325	0.114	1.115	0.318	0.133	0.497
PDI	0.200	0.082	0.167	0.081	0.125	0.119	0.149	0.083	0.120
Conotoxins	18.3023	42.923	28.758	70.898	62.064	46.934	72.168	68.611	69.479

Fig. 6.—Distribution of conotoxin precursor superfamily transcript relative expression (TPMs) along the distal (D), medium (M) and proximal (P) portion of the venom duct (with respect to the venom bulb). Highly-medium expressed transcripts are shown in panel (*A*). Low expressed components are depicted in panel (*B*). Overall conotoxin expression with respect to other protein expression in the venom duct is shown in panel (*C*).

Table 3

Differential Expression of Conotoxins in the Different Regions of the Venom Gland of *Chelyconus ermineus*

	Region	PPDE
Conotoxins	PM	1
Superfamilies		
A	PM	0.96
I3	D	0.99
Superfamily A		
A-1 (alpha 4/4)	PM	0.94
A-2 (Kappa)	P	1
Superfamily A-1		
Cerm_138	M	0.71
Cerm_255	P	0.70
Cerm_405	P	1
Superfamily A-2		
Cerm_008	PD	0.89
Cerm_145	PM	0.82
Cerm_268	P	0.76
Cerm_342	D	1

The composition of this cocktail greatly varies among species and currently, we are just starting to catalogue the repertoire of conotoxins and associated proteins produced by the more than 800 species of cones (e.g., Peng et al. 2016; Phuong et al. 2016; Li et al. 2017; Robinson, Li, Lu, et al. 2017). The recent advent of next generation RNA sequencing allows a robust approach to cataloguing all the transcripts expressed in the venom duct of a cone snail: not only it is possible to determine the bulk of mRNAs that are synthesized but also the variability of expression at distinct portions of the venom duct (Dutertre et al. 2014), among different individuals (Li et al. 2017), and under variable external conditions (Dutertre et al. 2010). Beyond identifying and describing all the key components of the venom, the ultimate, more ambitious goal of cataloguing studies is to tackle long-standing evolutionary and ecological questions, for example, how the diversity of conotoxins evolved (Duda and Palumbi 1999), how the venom mixtures were adapted to highly distinct diets (Duda et al. 2001) and predation strategies (Olivera et al. 2015), how a cone individual is able to modulate the composition of the

Table 4

Superfamily A Diversity in Several Cone Snails with Different Diets Based on Conoserver Entries

		Framework IV				Framework I									Framework IX	
Genus	**species**	**cc7c2c1c3c**	**cc7c2c1c6c**	**cc6c2c1c3c**	**cc6c2c1c4c**	**cc4c7c**	**cc4c3c**	**cc4c4c**	**cc4c5c**	**cc4c6c**	**cc4c8c**	**cc3c5c**	**cc3c6c**	**cc5c10c**	**c2c4c6c**	**c10c1c3c**
Chelyconus	*ermineus*	4 (3)				2 (2)		5 (3)								
Chelyconus	*purpurascens*		1 (1)			3 (2)		(4)[a]								
Gastridium	*geographus*			3 (2)		5 (5)					1 (1)	6 (3)		1 (1)		
Pionoconus	*achatinus*	4 (4)		1 (1)								5 (2)	1 (1)			
Pionoconus	*catus*	3 (3)										1 (1)				
Pionoconus	*consors*	5 (5)				1 (1)						7 (4)	2 (1)			
Pionoconus	*magus*	3 (3)				1 (1)						1 (1)				
Pionoconus	*monachus*											1 (1)	1 (1)			
Pionoconus	*striatus*	4 (4)						1 (1)				8 (6)				
Textilia	*bullatus*				3 (3)	1 (1)		5 (4)								
Embrikena	*pergrandis*	(2)				1 (1)										
Afonsoconus	*kinoshitai*							2 (2)								
Asprella	*sulcata*					7 (5)										
Conus	*marmoreus*					7 (5)										
Cylinder	*textile*					7 (5)										
Rhombiconus	*imperialis*					1 (1)	4 (4)									
Calamiconus	*quercinus*					12 (11)		3 (2)								
Dendroconus	*betulinus*					8 (8)						1 (1)				
Fraterconus	*distans*					2 (2)				1 (1)						
Lithoconus	*leopardus*					11 (9)				3 (2)						
Lividoconus	*lividus*					6 (6)	3 (2)									
Puncticulis	*pulicarius*					5 (4)			1 (1)	1 (1)					2 (2)	1 (1)
Virgiconus	*emaciatus*					1 (1)		4 (1)	1 (1)	2 (2)						
Virgiconus	*flavidus*					5 (4)										
Vituliconus	*vitulinus*					1 (1)					2 (2)					

NOTE.—Number of precursors and mature proteins (in parentheses) are shown.

Red, piscivorous; green, molluscivorous; brown, vermivorous (fire worms); blue, vermivorous.

[a]Puillandre et al. (2012), Hoggard et al. (2017).

venom both to prey and to deter predators (Dutertre et al. 2014), or what is the role of conotoxin diversification in speciation (Li et al. 2017).

According to reconstructed phylogenies of the family Conidae, the piscivorous diet evolved independently in some Atlantic and Indo-Pacific cone genera (Duda et al. 2001; Duda and Palumbi 2004; Puillandre, Bouchet, et al. 2014; this work). The transcriptomes of several piscivorous genera (*Gastridium, Pionoconus*, and *Textilia*) from the Indo-Pacific have been reported but none from Atlantic piscivorous genera was available yet. Hence, the importance of sequencing the transcriptome of *C. ermineus*: its comparison with those of piscivorous cone snails from the Indo-Pacific, and against those of cones eating snails and worms could provide important clues on which conotoxins are needed specifically for fish hunting and how they evolved (Duda and Palumbi 2004).

The obtained number of assembled contigs per *C. ermineus* venom duct transcriptome is comparable to those typically reported in equivalent studies also based on the Illumina platform (Peng et al. 2016; Li et al. 2017; Robinson, Li, Lu, et al. 2017). Moreover, the total number of clean reads, which mapped onto conotoxin transcripts constituted 57%. This number is somewhat higher than those reported by studies based on direct sequencing of individual cDNA clones such as, for example, the 39–50% of transcripts being conotoxins reported for *Virgiconus virgo*, *Tesselliconus eburneus*, *R. imperialis*, and *C. marmoreus* (Liu et al. 2012) or based on 454 sequencing, for example, the 42.7% of transcripts being conotoxins reported for *P. consors* (Terrat et al. 2012) but lower to the 88% reported for *G. geographus* (Hu et al. 2012). A total of 378 transcripts encoding conotoxin precursors were identified in the venom duct of *C. ermineus*. This number is similar to those reported for the venom duct transcriptomes of *Virroconus coronatus* (331; Phuong et al. 2016), *Puncticulis arenatus* (326; Phuong et al. 2016), or *Harmoniconus sponsalis* (401; Phuong et al. 2016) and larger than others such as those reported in *D. betulinus* (215; Peng et al. 2016), *C. marmoreus* (158; Lavergne et al. 2013), *T. praecellens* (149–155; Li et al. 2017), *T. andremenezi* (107–128; Li et al. 2017), or *C. gloriamaris* (108; Robinson, Li, Lu, et al. 2017). Until recently, most studies were based on single individuals whereas the current trend is to sequence several specimens as here (Barghi et al. 2015a; Peng et al. 2016; Li et al. 2017), which affects the comparison of numbers. In fact, the *C. ermineus* individuals produced each 145–176 conotoxin precursors, which is consistent with most reported single individual transcriptomes (see above). In *C. ermineus*, ~20% of the inferred mature conotoxins were common to the three analyzed individuals and strikingly, each of them showed an important number of sequences not found in the other two. Therefore, it is very likely that even the number here reported underestimates the whole diversity of conotoxins produced by this species. Such interindividual differences are congruent with results reported in the closely related *C. purpurascens* (Rodriguez et al. 2015) and for three specimens of *D. betulinus* with different body sizes (Peng et al. 2016). In this regard, it has been suggested that differences in age/size could be a factor fostering intraspecific variation (Barghi et al. 2015a; Peng et al. 2016). The three specimens of *C. ermineus* here analyzed were presumably adults and they differed in their size: ERM1-3 had shells of 73.1, 55 and 46 mm in length, respectively. In contrast, other studies found little intraspecific variation within species of the genus *Turriconus* (Li et al. 2017) and of the genus *Kioconus* (Barghi et al. 2015a). Here, it is important to note that despite our specimens were from different islands (Boa Vista, Sal, and Santa Luzia), they have not accumulated larger *cox1* sequence divergences than individuals within species of *Turriconus* or *Kioconus* (see additional supplementary fig. 2, Supplementary Material online). Altogether, our results suggest that, in the near future, many more individuals than those currently analyzed would be required to describe the whole richness of the venom of any cone species (Dutertre et al. 2010).

The Illumina-based venom repertoires of *C. ermineus* and of other recently investigated cone species with various diets were compared in table 5. Most conotoxin precursor superfamilies reported in other cone species (Peng et al. 2016; Phuong et al. 2016; Li et al. 2017; Robinson, Li, Lu, et al. 2017) were also identified in *C. ermineus*. The most diverse (in number of members) superfamilies in *C. ermineus* were O1, O2, M, and T, a pattern that is conserved in other cone species (table 5). Interestingly, the W and Z superfamilies, which have a cysteine-poor mature peptide and were originally reported in *C. marmoreus* (Lavergne et al. 2013), were particularly abundant in *C. ermineus*. Different cone species seem to have undergone diversification bursts of particular superfamilies, which are otherwise poorly represented in other species (Duda and Remigio 2008; Puillandre et al. 2012; Barghi et al. 2015a). A paradigmatic case is the A superfamily, which is highly diverse in the Indo-Pacific piscivorous *G. geographus* (Safavi-Hemami et al. 2014), *P. consors* (Terrat et al. 2012), *Pionoconus catus* (Himaya et al. 2015), and *T. bullatus* (Hu et al. 2011) but underrepresented in *C. ermineus*. Other examples are: the superfamilies P and O1d, which are particularly rich in *Turriconus* (Li et al. 2017), the B1 superfamily in *G. geographus*, the I2 superfamily in *V. virgo*, the A, I3, and N superfamilies in *Rolaniconus varius*, the D superfamily in *Rhizoconus vexillum* (Prashanth et al. 2016), and the con-ikot-ikot and B2 superfamilies in *K. tribblei* and *K. lenavati* (Barghi et al. 2015a).

About 20% of the newly identified conotoxin precursors could not be assigned to known superfamilies based on their signal domain, and new unassigned conotoxin precursor superfamilies had to be proposed temporarily as in other studies (e.g., Barghi et al. 2015b; Peng et al. 2016). However, most of these unassigned superfamilies had homologs in other cone species, and thus, it is foreseen that as more cone transcriptomes become available, formal (phylogeny-

Table 5

Comparison of the Diversity of Main Conotoxin Precursor Superfamilies in Cone Snails with Different Diets

Species	*Chelyconus ermineus*	*Gastridium geographus*	*Pionoconus consors*	*Cylinder gloriamaris*	*Conus marmoreus*	*Conus marmoreus*	*Dendroconus betulinus*	*Turriconus andremenezi*	*Turriconus praecellens*	*Rolaniconus varius*	*Virgiconus virgo*	*Rhombiconus imperialis*
												
Reference	**This work**	Hu et al. 2012	Terrat et al. 2012	Robinson et al. 2017	Phuong et al. 2016	Lavergne et al. 2013	Peng et al. 2016	Li et al. 2017	Li et al. 2017	Phuong et al. 2016	Phuong et al. 2016	Phuong et al. 2016
NGS platform	**HiSeq2000**	**FLX Titanium**	**454**	**HiSeq2000**	**HiSeq2000**	**454**	**HiSeq2000**	**HiSeq2000**	**HiSeq2000**	**HiSeq2000**	**HiSeq2000**	**HiSeq2000**
Diet	**fish**	**fish**	**fish**	**snail**	**snail**	**snail**	**worm**	**worm**	**worm**	**worm**	**worm**	**fire worms**
A	8	12	14	3	1	0	9	0	0	16	5	3
A2	2	–[a]	–	–	–	–	–	4	2	–	–	–
B1 (conantokin)	7	2	1	1	0	0	8	0	0	2	1	0
B2	8	0	0	1	1	0	1	2	3	2	0	1
B4	–	–	–	–	0	–	–	2	0	8	0	0
C	0	0	0	0	0	0	1	0	0	0	0	0
Con-ikot-ikot	5	6	0	2	0	0	4	2	5	8	1	0
Conkunitzin	13	1	7	0	1	0	8	0	0	2	1	0
D	3	0	0	0	0	0	0	0	2	0	0	1
E	2	0	0	1	4	0	2	0	0	2	1	1
F	4	0	0	1	4	0	3	0	0	2	1	0
H	7	0	0	3	3	0	3	1	3	0	0	0
I1	4	1	0	2	2	0	5	3	2	2	0	2
I2	9	0	0	3	0	0	8	4	5	3	14	2
I3	4	0	0	0	0	0	2	6	5	16	0	0
I4	5	–	–	4	4	2	0	4	6	0	0	0
J	2	4	0	4	0	0	3	0	9	1	0	0
K	1	0	0	0	0	0	0	0	0	2	0	5
L	3	0	0	0	0	0	0	0	3	3	0	0
M	27	2	8	7	20	55	30	28	33	23	7	8
N	0	0	0	1	2	0	6	0	1	11	6	1
O1	31	18	15	12	13	61	33	22	36	22	26	9
O1d	–	–	–	–	–	–	–	5	12	–	–	–
O2 (contryphan)	24	1	0	18	3	4	16	17	17	5	13	3
O3	6	0	0	1	0	0	4	0	0	2	1	0
P	2	0	1	3	0	0	13	26	32	14	0	12
Q	3	–	–	0	0	1	0	0	0	0	2	0
R	1	–	–	0	–	1	0	0	0	–	–	–
S	8	5	3	1	1	0	1	2	3	5	0	1
T	36	6	5	21	20	7	17	41	39	19	13	8
U	0	–	–	1	0	1	0	0	2	0	1	0

V	4	0	0	0	0	0	3	3	0	3	0	0	0	0	0	0	0	0	0	0
W	32	–	–	–	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
X					2															
Y	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Y3					1															
Z	20	–	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
conoCAP	0	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Conomap	0	–	–	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Conophysin (Conopressin)	6	1	1	1	0	0	1	2	7	0	1	0	0	0	0	0	0	0	0	1
Conorfamide	0	–	–	1	–	0	2	3	–	–	–	–	–	–	–	–	–	–	–	–
Contulakin	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Insulin	7	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Prohormone-4	0	–	–	3	–	0	2	1	–	–	–	–	–	–	–	–	–	–	–	–
Thyrostimulin α	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Thyrostimulin β	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

[a]Hyphen indicates that the study apparently did not search for the corresponding peptide.

based) classification of these unassigned superfamilies will be possible (Puillandre et al. 2012; Lavergne et al. 2013). Remarkably, we found a combination of mature peptides normally associated to the signal and propeptide domains of the O3 and T superfamilies, which in *C. ermineus* were associated to amino terminal sequences lacking signal and/or propeptide domains (unassigned families 16–22). This observation may evoke the possibility of domain shuffling as one of the underlying mechanisms for generating precursor diversity (Pi, Liu, Peng, Liu, et al. 2006). Despite the coverage is high and homogeneous throughout these assembled transcripts (including the boundary of the mature peptide and the rest of the putative precursor), the possibility of an assembly artifact cannot be fully excluded given that most of these sequences do not show a canonical three-domain structure (and should be validated experimentally).

One hot debate in the past has been the relevance of classifying conotoxins into cysteine-rich and cysteine-poor categories (Puillandre et al. 2012; Robinson, Li, Bandyopadhyay, et al. 2017). Our results support that this dichotomy is irrelevant from an evolutionary perspective, although it may have some functional meaning. According to our phylogenetic and sequence comparison analyses, there are at least three different evolutionary origins for cysteine-poor mature peptides: 1) whole conotoxin precursor superfamilies carrying mature peptides with no cysteine framework, such as, for example, R, W, and Z; 2) conotoxin precursor superfamilies in which cysteine-poor mature peptides are associated to a distinct signal and propeptide combination, and thus have a single evolutionary origin, such as, for example, M; and 3) conotoxin precursor superfamilies in which cysteine-poor mature peptides are associated to signal and propeptide combinations that also can be linked to mature peptides with a known cysteine framework, indicating multiple evolutionary origins, such as, for example, T. This third evolutionary pattern supports the importance of modularity as evolutionary mechanism for generating conotoxin diversity (Pi, Liu, Peng, Liu, et al. 2006).

Importantly, the use of the amino acid sequences of the propeptide domain in addition to those of the signal domain for phylogenetic analyses and sequence similarity comparisons proved to be very informative (Lavergne et al. 2013) and led to discrimination of potential distinct paralogs within the different conotoxin precursor superfamilies. For instance, at least four paralogs were identified within the M superfamily (named M1-3, M-WF), three within the T superfamily (T1-3; see Liu et al. 2012), which detected up to four clades in their phylogenetic analysis), six within the O1 superfamily (O1-1-6; see Li et al. 2017), which already distinguish O1d from the remaining members of O1) and six within the O2 superfamily (O2-1-6). This, thus far, mostly overlooked diversity within each superfamily will need to be taken into account in future studies when updating classifications of conotoxin precursor superfamilies (Lavergne et al. 2013) and when summarizing the composition of venoms in the different species as the

different paralogs may well have different functions (Altenhoff et al. 2012).

The current classification system of conotoxin superfamilies originally based on the alphabet can barely integrate the many novel signal sequences (and corresponding superfamilies) found in every new study and would have even more serious problems in dealing with paralog diversity within each superfamily, if this information is also incorporated. Therefore, a radically new evolutionary-based classification (beyond the goals of the present study) is urgently needed, and should come out from the consensus among experts in the field.

The distribution of conotoxin precursor diversity along the venom duct of *C. ermineus* showed some degree of regionalization. The A2 and B2 superfamilies had more diversity toward the proximal portion; Z was more diverse in the medium portion; and O1, M, and T superfamilies had more diversity toward the distal portion. The corresponding pattern in *G. geographus* showed the A and M superfamilies to be more abundant in the proximal portion, and O1 and T more diverse in the distal portion (Hu et al. 2012). The distribution pattern found in the molluscivorous *Cylinder textile*, showed the M and T superfamilies as more abundant in the proximal region whereas the O superfamily was more diverse in the distal region (Garrett et al. 2005). Overall, these comparisons show two discriminant patterns: 1) the T superfamily is more diverse in the distal portion of piscivorous cones and in the proximal region of the molluscivorous cone; and 2) the M superfamily is more diverse in the distal region of *C. ermineus* and in the proximal portion of *G. geographus*.

The venom duct of a cone snail is a specialized convoluted duct mostly devoted to the biosynthesis of conotoxins (Safavi-Hemami et al. 2014), as further demonstrated here by the elevated proportion of conotoxin transcripts detected in the transcriptome of the *C. ermineus* venom duct. Moreover, our results support that conotoxin expression is localized preferentially in the medium and proximal regions of the venom duct of *C. ermineus* whereas the distal region is mostly devoted to the expression of house-keeping genes. The inferred expression patterns of conotoxin precursor superfamilies in *C. ermineus* showed drastic variations among individuals. This may reflect natural conditions or potential methodological biases (despite the use of common sample handling, laboratory, sequencing, and analytical procedures), although we cannot discern between both possibilities. In any case, these results highlight the need of sequencing a fair amount of specimens to generate statistically robust quantitative comparisons and conclusions, as it is becoming the rule for model system species (Schurch et al. 2016). Being cautious in the interpretation of the results of our expression analyses (i.e., considering reliable only those TPM values, which are similar in at least two individuals), we observed that those conotoxin precursor superfamilies showing higher levels of expression are also those having more member diversity (O1, O2, M, and T). A striking exception to this pattern is the A superfamily, which has few distinct members in *C. ermineus*, but the highest levels of expression (see below). Most superfamilies showed low expression levels, suggesting a subtle contribution of them to the final venom composition. The expression of most superfamilies appears to reflect some degree of compartmentalization. In particular, the A superfamily is preferentially expressed in the medium and proximal regions of the venom duct whereas most other superfamilies tend to be expressed toward the distal region (as occurs in Indo-Pacific cones such as *G. geographus*; Hu et al. 2012). Moreover, we found that this compartmentalization of the expression of the A superfamily is n statistically significant, further indicating its functional importance in the venom of *C. ermineus*.

The different strategies of prey capture among piscivorous cone species determine the exact mixture (termed "cabal") of venom components, which will act coordinately to produce a specific physiological response (Olivera et al. 2016). While the Indo-Pacific species *G. geographus* has a "net engulfment" strategy, the Indo-Pacific species of the genus *Pionoconus* and *Textilia* as well as the Atlantic and Eastern Pacific species of the genus *Chelyconus* have a "taser and tether" or "hook and line" strategy (Olivera et al. 2015, 2016). The "net engulfment" strategy requires disorienting the fish by the release of the "nirvana cabal" into the water. This is a mixture, among others, of B1 superfamily (Hu et al. 2012) and insulin-like peptides (Safavi-Hemami et al. 2015; Robinson, Li, Bandyopadhyay, et al. 2017). Once fishes are disoriented and engulfed, the cone injects into each captured fish, a group of paralytic conotoxins, the "motor cabal," which includes the αA conotoxins, and the M superfamily μ- and ψ-conotoxins. In the case of the "taser and tether" strategy, the capture of the prey occurs through direct injection of two different mixtures, the "lightning-strike" and the motor cabals (Olivera 2002). The lightning-strike cabal induces an excitatory response, which ultimately causes tetanic paralysis. This cabal includes δ-conotoxins and κ-conotoxins from the O1 superfamily, conkunitzins, and κA conotoxins (Himaya et al. 2018).

Therefore, while the nirvana cabal is exclusive of some of species of the genus *Gastridium*, the lightning-strike cabal is found in genera such as *Pionoconus*, *Textilia*, and *Chelyconus*, and the motor cabal is found in all four above-mentioned genera (Olivera et al. 2016). The κA conotoxins of the lightning-strike cabal have generally the same cysteine spacing pattern in *Pionoconus*, *Textilia*, and *Chelyconus* (except *C. purpurascens*). However, the main blockers of K^+ channels belong to the O superfamily in *Chelyconus*, whereas this physiological role is accomplished by conkunitzins in *Pionoconus* (Olivera et al. 2016). With regards to the motor cabal, there are striking instances of differential recruitment of the αA conotoxins. All piscivorous genera have generally at least one member of the α4/7 subfamily, which blocks neuronal nAChRs (Azam and McIntosh 2009). In addition, *Gastridium*

and *Pionoconus* have members of the α3/5 subfamily, which inhibit muscle nAChRs. However, *Chelyconus* and *Textilia* have instead members of the α4/4 subfamily. The α4/4 conotoxins of *C. ermineus* and *C. purpurascens* have been shown to selectively bind muscle nAChRs (López-Vera et al. 2007; Quinton et al. 2013) whereas those of *Textilia* block neuronal nAChRs (Chi et al. 2006). The presence of α4/4 conotoxins has been detected also in *Afonsoconus*, and three species of *Pionoconus*, but their functions have not been determined (Puillandre et al. 2012). The α4/4 sequences of *Pionoconus* and *Textilia* (and probably *Afonsoconus*) are more closely related phylogenetically than those of *Chelyconus* (Puillandre et al. 2012). Altogether, our results suggest independent genetic and biochemical pathways to evolve the same diet adaptation, and thus, favor the hypothesis of a convergent origin of piscivory in cones from the Indo-Pacific and Atlantic oceans (Duda et al. 2001; Puillandre, Bouchet, et al. 2014).

Finally, it is interesting to note that recent analysis of conotoxin envenomation in *C. purpurascens* showed that different individuals could include alternatively either the lightning-strike, the motor or both cabals in the composition of their venoms when preying (Himaya et al. 2018). In our case (see supplementary table 1, Supplementary Material online), the three individuals of *C. ermineus* produced conotoxins belonging to the motor cabal including α4/4 conotoxins (the α4/7 conotoxin was not detected), and M superfamily μ- and ψ-conotoxins as well as had conotoxins belonging to the lightning-strike cabal including the κA conotoxins, and the O1 superfamily δ- (Aman et al. 2015) and κ-conotoxins. Moreover, the differential high levels of A superfamily transcripts in the proximal region of the venom duct in *G. geographus* as part of the motor cabal have been associated to defense-evoked responses (Dutertre et al. 2014). In *C. ermineus* (see supplementary table 1, Supplementary Material online), for the A superfamily, the members involved in the motor cabal (α4/4 conotoxins) and the lightning-strike cabal (κA conotoxins) showing differential expression are located in the proximal and distal regions of the venom duct, respectively, supporting the regionalization of the cabals as reported in *Protostrioconus obscurus* (Dutertre et al. 2014). Nevertheless, all the above-mentioned inferences need to be interpreted with caution and as tentative until further comparative analyses based on more individuals are carried out.

Conclusions

The venom duct of *C. ermineus* produces a great diversity of conotoxin precursors, most corresponding to known superfamilies and several showing novel signal domains. Comparison of these data to the venom repertoires reported from different cone species with various diets supports that some superfamilies (O1, O2, T, M) are widespread among cone species, making the basic venom toolkit, whereas others are restricted to fewer lineages. The different superfamilies show various degrees of expansion depending on the species. In the case of *C. ermineus*, the cysteine-poor superfamilies W and Z are particularly diverse. In this regard, the wide distribution of cysteine-poor mature peptides among superfamilies indicate multiple and diverse origins. Both, diversity and expression of conotoxins are regionalized along the venom duct. Diversity in the number of members of a superfamily increases toward the distal region whereas the less diverse superfamilies in the proximal region show higher expression levels. In particular, the A superfamily, which is highly diverse in piscivorous cones from the Indo-Pacific Ocean, consists of rather few and distinct (α4/4) members in the *C. ermineus* venom, but these show differentially and significantly high expression levels toward the proximal region. These contrasting patterns support convergent strategies to produce the motor cabal, which targets nicotinic acetylcholine receptors, and seems essential for deterring/preying fishes.

Our results show that each newly analyzed cone species uncovers additional conotoxin diversity and thus, that we are still far from covering the whole repertoire of conotoxins, as the venom duct transcriptome of the majority of cone species awaits sequencing and analysis. Moreover, the numerous unassigned superfamilies, which are discovered in every new cone transcriptome together with the emerging evidence of the existence of distinct paralogs within each superfamily prompt for a revision and an update of the nomenclature of conotoxins as the use of the alphabet-based classification seems to be too constrained and obsolete in evolutionary terms.

Supplementary Material

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

Acknowledgments

We thank four anonymous reviewers, Sébastien Dutertre, and Christoph Bleidorn for insightful comments on a previous version of the manuscript. We are indebted to Cabo Verde biology students Paulo Vasconcelos and Stiven Delgado for their valuable help during sampling in Santa Luzia (Cabo Verde). We thank Dr Rui Freitas from the Universidade de Cabo Verde (UniCV) for his continuous support of our research in Cabo Verde. We also thank Dr Iderlindo Silva dos Santos and Dra. Sonia Monteiro de Pina Araujo from the Direccçao Nacional do Medio Ambiente of the Ministério do Ambiente, Habitação e Ordenamento do Território (MAHOT) of the Republic of Cabo Verde for their help with collecting permits (Autorizações 07/2013, 26/2013, 01/2104, 04/2015, and 03/2016). We are grateful to Jesús Marco and Aida Palacios, who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing assembling and

phylogenetic analyses. We are grateful to Etienne Kornobis for help with early analyses. Thanks to Alexander Medvedev, who shares free the pictures of the different species of cones from his collection (www.coneshells-am.ru; last accessed September 05, 2018). This work was funded by the Spanish Ministry of Economy, Industry and Competitiveness (CGL2013-45211-C2-2-P and CGL2016-75255-C2-1-P [AEI/FEDER, UE]) to R.Z.; BES-2014-069575 to S.A.

Literature Cited

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38(Web Server issue):W7–13.

Altenhoff AM, Studer RA, Robinson-Rechavi M, Dessimoz C. 2012. Resolving the ortholog conjecture: orthologs tend to be weakly, but significantly, more similar in function than paralogs. PLoS Comput Biol. 8(5):e1002514.

Aman JW, et al. 2015. Insights into the origins of fish hunting in venomous cone snails from studies of *Conus tessulatus*. Proc Natl Acad Sci U S A. 112(16):5087.

Andrews S. 2010. FastQC. available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/; last accessed September 05, 2018.

Azam L, McIntosh JM. 2009. Alpha-conotoxins as pharmacological probes of nicotinic acetylcholine receptors. Acta Pharmacol Sin. 30(6):771.

Bandyopadhyay PK, et al. 1998. Conantokin-G precursor and its role in γ-carboxylation by a vitamin K-dependent carboxylase from a conus-snail. J Biol Chem. 273(10):5447–5450.

Barbier J, et al. 2004. A δ-conotoxin from *Conus ermineus* venom inhibits inactivation in vertebrate neuronal Na+ channels but not in skeletal and cardiac muscles. J Biol Chem. 279(6):4680–4685.

Barghi N, Concepcion GP, Olivera BM, Lluisma AO. 2015a. Comparison of the venom peptides and their expression in closely related *Conus* species: insights into adaptive post-speciation evolution of *Conus* exogenomes. Genome Biol Evol. 7(6):1797–1814.

Barghi N, Concepcion GP, Olivera BM, Lluisma AO. 2015b. High conopeptide diversity in *Conus tribblei* revealed through analysis of venom duct transcriptome using two high-throughput sequencing platforms. Mar Biotechnol 17(1):81–98.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2005. GenBank. Nucleic Acids Res. 33:D34–D38.

Biggs JS, et al. 2010. Evolution of *Conus* peptide toxins: analysis of *Conus californicus* Reeve, 1844. Mol Phylogenet Evol. 56(1):1–12.

Born IEv. 1778. Index rerum naturalium Musei Cæsarei Vindobonensis. Pars I. Testacea. Verzeichniss der natürlichen Seltenheiten des k. k. Naturalien Cabinets zu Wien. Erster Theil. Schalthiere. Vindobonae: Vindobonæ ex officina Krausiana.

Buczek O, Olivera BM, Bulaj G. 2004. Propeptide does not act as an intramolecular chaperone but facilitates protein disulfide isomerase-assisted folding of a conotoxin precursor. Biochemistry 43(4):1093–1101.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17(4):540–552.

Chang D, Duda TF. 2012. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Mol Biol Evol. 29(8):2019–2029.

Chang D, Duda TF. 2016. Age-related association of venom gene expression and diet of predatory gastropods. BMC Evol Biol. 16(1):1–12.

Chi S-W, Kim D-H, Olivera BM, McIntosh JM, Han K-H. 2006. NMR structure determination of α-conotoxin BulA, a novel neuronal nicotinic acetylcholine receptor antagonist with an unusual 4/4 disulfide scaffold. Biochem Biophys Res Commun. 349(4):1228–1234.

Conticello SG, et al. 2001. Mechanisms for evolving hypervariability: the case of conopeptides. Mol Biol Evol. 18(2):120–131.

Conticello SG, et al. 2003. The prodomain of a secreted hydrophobic miniprotein facilitates its export from the endoplasmic reticulum by hitchhiking on sorting receptors. J Biol Chem. 278(29):26311–26314.

Corpuz GP, et al. 2005. Definition of the M-Conotoxin superfamily: characterization of novel peptides from molluscivorous *Conus* venoms. Biochemistry 44(22):8176–8186.

Davis J, Jones A, Lewis RJ. 2009. Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. Peptides 30(7):1222–1227.

Duda TF, Kohn AJ, Palumbi SR. 2001. Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. Biol J Linn Soc. 73(4):391–409.

Duda TF, Palumbi SR. 1999. Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. Proc Natl Acad Sci U S A. 96(12):6820–6823.

Duda TF, Palumbi SR. 2004. Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*. Proc R Soc B Biol Sci. 271(1544):1165–1174.

Duda TF, Remigio EA. 2008. Variation and evolution of toxin gene expression patterns of six closely related venomous marine snails. Mol Ecol. 17(12):3018–3032.

Dutertre S, Biass D, Stöcklin R, Favreau P. 2010. Dramatic intraspecimen variations within the injected venom of *Conus consors:* an unsuspected contribution to venom diversity. Toxicon 55(8):1453–1462.

Dutertre S, et al. 2013. Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. Mol Cell Proteomics 12(2):312–329.

Dutertre S, et al. 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nat Commun. 5:3521.

Echterbille J, et al. 2017. Discovery and characterization of EIIB, a new α-conotoxin from *Conus ermineus* venom by nAChRs affinity capture monitored by MALDI-TOF/TOF mass spectrometry. Toxicon 130:1–10.

Endean R, Duchemin C. 1967. The venom apparatus of *Conus magus*. Toxicon 4(4):275–284.

Espiritu DJD, et al. 2001. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon 39(12):1899–1916.

Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol. 17(6):368–376.

Garrett JE, Buczek O, Watkins M, Olivera BM, Bulaj G. 2005. Biochemical and gene expression analyses of conotoxins in *Conus textile* venom ducts. Biochem Biophys Res. Communications 328(1):362–367.

Gowd KH, et al. 2008. Conantokin-P, an unusual conantokin with a long disulfide loop. Toxicon 52(2):203–213.

Grabherr MG, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644–652.

Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3):307–321.

Himaya SWA, et al. 2015. Comparative venomics reveals the complex prey capture strategy of the piscivorous cone snail *Conus catus*. J Proteome Res. 14(10):4372–4381.

Himaya SWA, Marí F, Lewis RJ. 2018. Accelerated proteomic visualization of individual predatory venoms of *Conus purpurascens* reveals separately evolved predation-evoked venom cabals. Sci Rep. 8(1):330.

Hoggard MF, et al. 2017. *In vivo* and *in vitro* testing of native α-conotoxins from the injected venom of *Conus purpurascens*. Neuropharmacology 127:253–259.

Hu H, Bandyopadhyay PK, Olivera BM, Yandell M. 2011. Characterization of the *Conus bullatus* genome and its venom-duct transcriptome. BMC Genomics 12:60.

Hu H, Bandyopadhyay PK, Olivera BM, Yandell M. 2012. Elucidation of the molecular envenomation strategy of the cone snail *Conus geographus* through transcriptome sequencing of its venom duct. BMC Genomics 13:284.
Jacobsen R, et al. 1997. Differential targeting of nicotinic acetylcholine receptors by novel αA-conotoxins. J Biol Chem. 272(36):22531–22537.
Jin A-h, et al. 2013. Transcriptomic messiness in the venom duct of *Conus miles* contributes to conotoxin diversity. Mol Cell Proteomics 12(12):3824–3833.
Jin A-H, et al. 2015. δ-Conotoxin SuVIA suggests an evolutionary link between ancestral predator defence and the origin of fish-hunting behaviour in carnivorous cone snails. Proc R Soc B Biol Sci. 282(1811):20150817.
Kaas Q, Westermann J-C, Craik DJ. 2010. Conopeptide characterization and classifications: an analysis using ConoServer. Toxicon 55(8):1491–1509.
Kaas Q, Yu R, Jin A-H, Dutertre S, Craik DJ. 2012. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. Nucleic Acids Res. 40(D1):D325–D330.
Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
Kornobis E, et al. 2015. TRUFA: a user-friendly web server for de novo RNA-seq analysis using cluster computing. Evol Bioinformatics 11:EBO.S23873–104.
Koua D, et al. 2012. ConoDictor: a tool for prediction of conopeptide superfamilies. Nucleic Acids Res. 40(Web Server issue):W238–W241.
Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 29(6):1695–1701.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357.
Lavergne V, et al. 2013. Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. BMC Genomics 14:708.
Lavergne V, et al. 2015. Optimized deep-targeted proteotranscriptomic profiling reveals unexplored Conus toxin diversity and novel cysteine frameworks. Proc Natl Acad Sci U S A. 112(29):E3782–E3791.
Leng N, et al. 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics 29(8):1035–1043.
Lewis RJ, Dutertre S, Vetter I, Christie MJ. 2012. *Conus* venom peptide pharmacology. Pharmacol Rev. 64(2):259–298.
Li Q, et al. 2017. Divergence of the venom exogene repertoire in two sister species of *Turriconus*. Genome Biol Evol. 9(9):2211–2225.
Liu Z, et al. 2012. Diversity and evolution of conotoxins in *Conus virgo*, *Conus eburneus*, *Conus imperialis* and *Conus marmoreus* from the South China Sea. Toxicon 60(6):982–989.
Lluisma AO, Milash BA, Moore B, Olivera BM, Bandyopadhyay PK. 2012. Novel venom peptides from the cone snail *Conus pulicarius* discovered through next-generation sequencing of its venom duct transcriptome. Marine Genomics 5:43–51.
López-Vera E, Jacobsen RB, Ellison M, Olivera BM, Teichert RW. 2007. A novel alpha conotoxin (α-PIB) isolated from C. purpurascens is selective for skeletal muscle nicotinic acetylcholine receptors. Toxicon 49(8):1193–1199.
Lu A, Yang L, Xu S, Wang C. 2014. Various conotoxin diversifications revealed by a venomic study of *Conus flavidus*. Mol Cell Proteomics 13(1):105–118.
Luo S, et al. 2006. Identification and molecular diversity of T-superfamily conotoxins from *Conus lividus* and *Conus litteratus*. Chem Biol Drug Des. 68(2):97–106.
Martinez JS, et al. 1995. alpha-conotoxin EI, a new nicotinic acetylcholine receptor antagonist with novel selectivity. Biochemistry 34(44):14519–14526.
Olivera BM. 2002. *Conus* venom peptides: reflections from the biology of clades and species. Annu Rev Ecol Syst. 33(1):25–47.
Olivera BM, Fedosov A, Imperial JS, Kantor K. 2016. Physiology of envenomation by conoidean gastropods. In: Saleuddin S, Mukai S, editors. Physiology of molluscs. New York: Apple Acdemic Press. p. 154–188.
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.
Peng C, et al. 2016. High-throughput identification of novel conotoxins from the Chinese tubular cone snail (*Conus betulinus*) by multi-transcriptome sequencing. GigaScience 5:17.
Phuong MA, Mahardika GN, Alfaro ME. 2016. Dietary breadth is positively correlated with venom complexity in cone snails. BMC Genomics 17:401.
Pi C, Liu Y, Peng C, Jiang X, et al. 2006. Analysis of expressed sequence tags from the venom ducts of *Conus striatus*: focusing on the expression profile of conotoxins. Biochimie 88(2):131–140.
Pi C, Liu J, Peng C, Liu Y, et al. 2006. Diversity and evolution of conotoxins based on gene expression profiling of *Conus litteratus*. Genomics 88(6):809–819.
Prashanth JR, et al. 2016. The role of defensive ecological interactions in the evolution of conotoxins. Mol Ecol. 25(2):598–615.
Prashanth JR, Lewis RJ, Dutertre S. 2012. Towards an integrated venomics approach for accelerated conopeptide discovery. Toxicon 60(4):470–477.
Prator CA, Murayama KM, Schulz JR. 2014. Venom variation during prey capture by the cone snail, *Conus textile*. PLoS One 9(6):e98991.
Puillandre N, Duda TF, Meyer C, Olivera BM, Bouchet P. 2014. One, four or 100 genera? A new classification of the cone snails. J Mollus Stud. 81(1):1.
Puillandre N, Bouchet P, et al. 2014. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Mol Phylogenet Evol. 78:290–303.
Puillandre N, Koua D, Favreau P, Olivera BM, Stöcklin R. 2012. Molecular phylogeny, classification and evolution of conopeptides. J Mol Evol. 74(5–6):297–309.
Puillandre N, Watkins M, Olivera BM. 2010. Evolution of *Conus* peptide genes: duplication and positive selection in the A-superfamily. J Mol Evol. 70(2):190–202.
Quinton L, et al. 2013. Identification and functional characterization of a novel α-conotoxin (EIIA) from *Conus ermineus*. Anal Bioanal Chem. 405(15):5341–5351.
Rivera-Ortiz JA, Cano H, Marí F. 2011. Intraspecies variability and conopeptide profiling of the injected venom of *Conus ermineus*. Peptides 32(2):306–316.
Robinson SD, Li Q, Bandyopadhyay PK, et al. 2017. Hormone-like peptides in the venoms of marine cone snails. Gen Comp Endocrinol. 244:11–18.
Robinson SD, Li Q, Lu A, et al. 2017. The venom repertoire of *Conus gloriamaris* (Chemnitz, 1777), the glory of the sea. Mar Drugs 15:145.
Robinson SD, Norton RS. 2014. Conotoxin gene superfamilies. Mar Drugs 12(12):6058–6101.
Rodriguez A, Dutertre S, Lewis R, Marí F. 2015. Intraspecific variations in *Conus purpurascens* injected venom using LC/MALDI-TOF-MS and LC-ESI-TripleTOF-MS. Anal Bioanal Chem. 407(20):6105–6116.
Safavi-Hemami H, Bulaj G, Olivera BM, Williamson NA, Purcell AW. 2010. Identification of *Conus* Peptidylprolyl Cis-Trans Isomerases (PPIases) and assessment of their role in the oxidative folding of conotoxins. J Biol Chem. 285(17):12735–12746.

Safavi-Hemami H, et al. 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.
Safavi-Hemami H, et al. 2015. Specialized insulin is used for chemical warfare by fish-hunting cone snails. Proc Natl Acad Sci U S A. 112(6):1743–1748.
Salisbury SM, Martin GG, Kier WM, Schulz JR. 2010. Venom kinematics during prey capture in *Conus*: the biomechanics of a rapid injection system. J Exp Biol. 213(5):673.
Santos AD, McIntosh JM, Hillyard DR, Cruz LJ, Olivera BM. 2004. The A-superfamily of Conotoxins: structural and functional divergence. J Biol Chem. 279(17):17596–17606.
Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27(6):863–864.
Schurch NJ, et al. 2016. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? RNA 22(10):1641–1851.
Schwarz G. 1978. Estimating the dimension of a model. Ann Stat. 6(2):461–464.
Shon K-J, et al. 1995. Purification, characterization, synthesis, and cloning of the Lockjaw peptide from *Conus purpurascens* venom. Biochemistry 34(15):4913–4918.
Shon K-J, et al. 1998. μ-Conotoxin PIIIA, a new peptide for discriminating among tetrodotoxin-sensitive Na channel subtypes. J Neurosci. 18(12):4473.
St John J. 2011. SeqPrep. Available from: https://github.com/jstjohn/SeqPrep; last accessed September 05, 2018.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.
Tayo LL, Lu B, Cruz LJ, Yates JR. 2010. Proteomic analysis provides insights on venom processing in *Conus textile*. J Proteome Res. 9(5):2292–2301.
Terlau H, Olivera BM. 2004. Conus venoms: a rich source of novel ion channel-targeted peptides. Physiol Rev. 84(1):41–68.
Terrat Y, et al. 2012. High-resolution picture of a venom gland transcriptome: case study with the marine snail *Conus consors*. Toxicon 59(1):34–46.
Tucker JK, Tenorio MJ. 2009. Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. Harxheim: Conchbooks.
Tucker JK, Tenorio MJ. 2013. Illustrated catalog of the living cone shells. Wellington (FL): MDM Publishing.
Uniprot Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 45(D1):D158–D169.
Uribe JE, Puillandre N, Zardoya R. 2017. Beyond *Conus*: phylogenetic relationships of Conidae based on complete mitochondrial genomes. Mol Phylogenet Evol. 107:142–151.
Violette A, et al. 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.
Volpon L, et al. 2004. NMR solution structures of δ-conotoxin EVIA from *Conus ermineus* that selectively acts on vertebrate neuronal Na+ channels. J Biol Chem. 279(20):21356–21366.
Woodward SR, Cruz LJ, Olivera BM, Hillyard DR. 1990. Constant and hypervariable regions in conotoxin propeptides. EMBO J. 9(4):1015–1020.
Wright ES, Vetsigian KH. 2016. Quality filtering of Illumina index reads mitigates sample cross-talk. BMC Genomics 17(1):876.
Wu Y, et al. 2013. Molecular evolution and diversity of *Conus* peptide toxins, as revealed by gene structure and intron sequence analyses. PLoS One 8(12):e82495.

Associate editor: Mandë Holford

3.5. – Chapter V: “Comparative transcriptomics of the venoms of continental and insular radiations of West African cone snails”

3.5. - Capítulo V: “Transcriptómica comparada de los venenos de las radiaciones continental e insular de los conos de África occidental”

Pendiente de envío

Se secuenciaron los transcriptomas de las glándulas de veneno de 12 especies estrechamente relacionadas de conos vermívoros endémicos de África Occidenta de los géneros *Africonus* y *Lautoconus*. Estos conos pertenecen, respectivamente, a radiaciones insulares y continentales, para las que había disponibles filogenias robustas que permiten estudios evolutivos comparativos. El número total de precursores de conotoxinas, hormonas y proteínas asociadas del veneno varió entre 95 y 210, y los mayores repertorios podrían indicar dietas más variadas. Realizamos análisis de reconstrucción ancestral con parsimonia y encontramos péptidos compartidos a nivel individual, de especie y de género, así como casos de evolución convergente. Los individuos de la misma especie compartían de la mitad hasta un tercio del total de los precursores de conotoxinas. Debido a la alta variabilidad de estos péptidos secretados, el número de secuencias comunes se redujo drásticamente en las comparaciones por pares entre especies estrechamente relacionadas y prácticamente casi ninguna secuencia era compartida a nivel de género. Los dos géneros mostraban distintos catálogos de precursores de conotoxinas en términos de tipo de superfamilias, abundancia de miembros por superfamilia, y niveles de expresión relativos. . Sin embargo, se encontró un conjunto común de seis superfamilias (T, O1, O2, M, Cerm_03 y conkunitzin) expandidas en todas las especies de conos estudiadas. Detectamos una sobreexpresión significativa de la superfamilia B1 en especies de *Africonus* con respecto a las especies de *Lautoconus*, y de la superfamilia A en el piscívoro *Chelyconus ermineus* con respecto a las especies vermívoras.

REGULAR ARTICLE (to be submitted)

Comparative transcriptomics of the venoms of continental and insular radiations of West African cone snails

Samuel Abalde[1] | Manuel J. Tenorio[2] | Carlos M. L. Afonso[3] | Rafael Zardoya[1]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain;
[2]Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cadiz; 11510 Puerto Real, Cádiz, Spain;
[3]Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

Correspondence

Rafael Zardoya, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain
Email: rafaz@mncn.csic.es

Abstract

The transcriptomes of the venom glands of 12 closely related species of vermivorous cones endemic to West Africa from genera *Africonus* and *Lautoconus* were sequenced. These cones belong, respectively, to insular and continental radiations, for which robust phylogenies were available, allowing comparative evolutionary studies. The total number of conotoxin precursors, hormones and associated venom proteins per species varied between 95 and 210, and larger repertoires could indicate broader diets. We were able to perform parsimony ancestral reconstructions and find shared peptides at the individual, species and genus levels, as well as instances of convergent evolution. Individuals of the same species shared half to one third of the total conotoxin precursors. Due to the high variability of these secreted peptides, the number of common sequences was drastically reduced in the pairwise comparisons between closely related species and virtually almost no sequence was shared at the genus level. The two genera showed distinct catalogues of conotoxins precursors in terms of type of superfamilies, abundance of members per superfamily, and relative expression levels. Yet, a common set of six superfamilies (T, O1, O2, M, Cerm_03, and conkunitzin) was found to be expanded in all studied cone species. We detected significant overexpression of B1 superfamily in *Africonu*s species with respect to *Lautoconus* species, and of A superfamily in the piscivorous *Chelyconus ermineus* with respect to the vermivorous species.

KEYWORDS
conotoxin precursors, transcriptomes, *Africonus*, *Lautoconus*

1 | INTRODUCTION

Cone snails (Gastropoda: Conidae) are key predators in marine ecosystems that actively hunt on worms, snails, and fish (Kohn, 1959). Cones present a sophisticated venom system: a radular sac produces hollow radular teeth, which are loaded with venom generated in a convoluted venom duct (Tucker and Tenorio, 2009). The venom of cone snails is a cocktail constituted by hundreds of peptides named conotoxins, together with hormones and other proteins that participate in the synthesis or enhance the activity of the venom (Olivera, 2006; Barghi et al., 2015; Safavi-Hemami et al., 2015). Once secreted and inoculated into the prey, conotoxins can interact with different targets such as ionic channels and neurotransmitter receptors, triggering different physiological responses: from sedation to tetanic paralysis by muscle hyperactivity (Olivera et al., 1990; Lopez-

Vera et al., 2007; Robinson and Norton, 2014; Olivera et al., 2015; Ahorukomeye et al., 2019). Conotoxin precursors typically present a three domain structure, consisting of signal, pro-peptide, and mature (which constitutes, after cleavage of the other domains, the functional toxin) regions (Kaas et al., 2010). Sometimes a post-peptide region is also found. The signal region is highly conserved, and it has been used to classify these peptides into different toxin "superfamilies" (Robinson and Norton, 2014).

The composition of the venom secreted by cone snails is highly variable among species, specimens, and even within the same individual depending on its physiological status (Prator et al., 2014; Chang and Duda, 2016; Peng et al., 2016; Li et al., 2017; Abalde et al., 2018). This striking variability has been proposed to be generated through different mechanisms, including gene duplication (Duda and Palumbi, 1999; Espiritu et al., 2001), accelerated substitution rates (Conticello et al., 2001), recombination (Espiritu et al., 2001), differential expression (Duda and Palumbi, 2004), and/or post-translational modifications (Bergeron et al., 2013; Dutertre et al., 2013).

Thus far, cone snail venomics has been driven preferentially by the pharmacological potential of conotoxins, and the different studies were mostly limited to the purification of mature peptides and the identification of their function, thus lacking the wider phylogenetic perspective that is already being applied in the study of other venomous animals (Binford, 2001; Gibbs et al., 2013; Lomonte et al., 2014). Comparing venom cocktails in different cone species within a phylogenetic framework should provide insights on venom evolution, including how the rich diversity of conotoxins was generated and is maintained (Chang and Duda, 2012; Dutertre et al., 2014), to what extent the distinct repertoires are adapted to different diet specializations (Remigio and Duda, 2008; Chang and Duda, 2016; Phuong et al., 2016), which are the functional constraints and levels of convergence imposed by this coevolutionary arms race system (Conticello et al., 2001; Abalde et al., 2018), and which is the ultimate influence (if any) of conotoxin diversity in the extraordinary rates of species diversification of the group (Phuong et al., 2019).

Several recent studies have started exploring the evolutionary processes underlying the adaptive nature of venom composition in cones (Aman et al., 2014; Phuong et al., 2016; Abalde et al., 2018; Jin et al., 2019). As in other venomous animals (Pahari et al., 2007; Pekar et al., 2018), dietary breadth has been proposed to be a main factor triggering venom evolution in cones (Remigio and Duda, 2008; Elliger et al., 2011; Phuong et al., 2016). Since cone snail hunting performance relies on the specificity of their venom, shifts in diet can trigger changes in venom composition (Duda, 2008; Duda et al., 2009; Chang and Duda, 2016) and in general, species with more generalized diets tend to have more complex venoms (Phuong et al., 2016). Moreover, instances of functional convergence have been shown in the cocktails of Atlantic versus Indo-Pacific piscivorous cones (Abalde et al., 2018). In addition, another level of complexity comes from the capacity of cone snails to modulate the composition of their venom depending on its final use, whether to subdue preys or defend themselves against predators (Dutertre et al., 2014; Prashanth et al., 2016; Prashanth et al., 2017; Jin et al., 2019). Despite the extraordinary variability of the venom cocktails, and thus the great potential for ecological adaptation and species diversification, a recent study found no significant correlation between conotoxin gene diversity and speciation rates (Phuong et al., 2019), suggesting that other traits hampering gene flow may have been more critical in promoting the astonishing species diversity of cones (Cunha et al., 2005).

All the above-mentioned studies explored general venom evolutionary trends at the family (Conidae) level, comparing distantly related lineages. Here, we propose to analyze venom evolution within two radiations of closely related cone species inhabiting West Africa. This region is a hotspot of cone

diversity, including approximately 10% of all described species thus far (Tucker and Tenorio, 2013). This species diversity was generated through independent radiation events, leading to high rates of endemicity (Pin and Leung Tack, 1995; Cunha et al., 2005; Duda and Rolán, 2005). In particular, we focused on two species-rich lineages; one comprising cones endemic to the Cabo Verde archipelago and the other including cones endemic to Senegal (plus one closely related species inhabiting Canary Islands). Recently, robust phylogenies based on mitogenomes were reconstructed for both clades (ascribed to the genera *Africonus* and *Lautoconus*, respectively), providing the necessary framework for evolutionary studies (Abalde et al., 2017a; Abalde et al., 2017b). The ancestor of the genus *Africonus* arrived at the archipelago of Cabo Verde about 23 mya and diversified about nine mya into four main clades and at least 40 endemic species (Abalde et al., 2017a). The lineage of *Lautoconus* endemic to Senegal and Canary Islands diversified about six mya into three main clades and at least 15 endemic species (Abalde et al., 2017b). All of the cones in both clades are vermivorous. However, while cones endemic to Cabo Verde show no apparent differences in radular tooth morphology, the three clades of *Lautoconus* have each distinct radular teeth (Abalde et al., 2017b). All cones have non-planktotrophic larvae with restricted dispersal capacities. Therefore, it has been proposed that diversification of West African cones was in allopatry and mainly triggered by eustatic sea level changes during the Miocene-Pliocene (Cunha et al., 2005; Abalde et al., 2017a). Since *Africonus* species are normally restricted to single islands, the difference in number of species between the archipelago and the continent would be explained in terms of more chances to restrict gene flow in the former. However, the genus *Lautoconus* has only one species in Canary Islands, *Lautoconus guanche*, contradicting the pattern found in Cabo Verde. In this case, differences in the mean temperature of the water (in the limits of tolerance for cones in Canary Islands) and the proximity of Canary Islands to the continent (Fuerteventura Island is 120 km away from the coast of Morocco) could explain the lack of diversification (Cunha et al., 2014). No study, to our knowledge, has analysed the composition of the venoms of the vermivorous cones endemic to Cabo Verde, Canary Islands, and Senegal, nor their potential contribution to the observed enhanced rates of speciation in these areas.

Here, we sequenced the transcriptomes of the venom glands from ten individuals representing nine species of *Africonus* and four specimens of *Lautoconus* representing two species from Senegal and one from the Canary Islands. By sequencing these transcriptomes, we aimed to: 1) describe the venom composition for all these species in terms of presence and diversity of the conotoxin precursor superfamilies as well as relative abundance of the transcripts as proxy of expression levels; 2) assess the levels of divergence in venom composition at different hierarchical levels (within species, between species from the same clade within *Africonus*, between species from different clades within a genus, and between genera) and discern between shared derived peptides and cases of functional convergence; 3) determine whether there could be instances of differential expression between the two genera as footprint of adaptation; 4) compare venom compositions of these vermivorous species to that of the Atlantic piscivorous *Chelyconus ermineus* (Abalde et al., 2018) to further understand the connections between venom evolution and diet specialization; and 5) to determine whether there is any influence of venom variability in enhancing rates of diversification.

2 | Materials and Methods

2.1. Taxon sampling

A total of 14 specimens of cone snails were included in this study, four belonging to *Lautoconus* and ten to *Africonus*. Taxon selection was aimed at maximizing lineage representation and based on reconstructed phylogenies of the two genera (Abalde et al.,

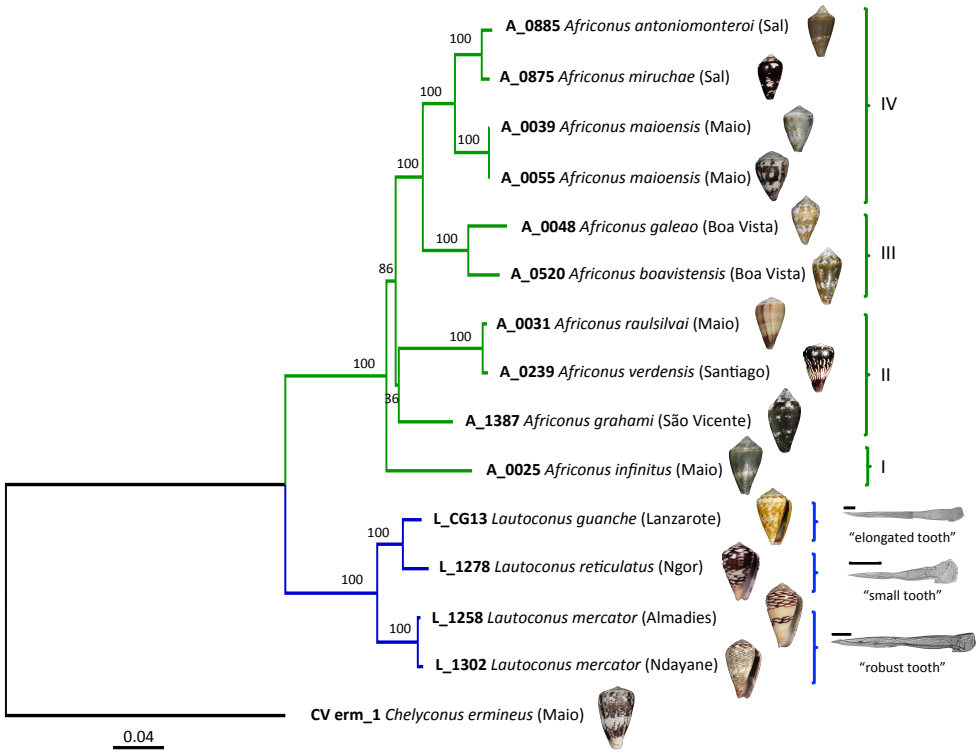


Figure 1. Phylogenetic relationships among the species included in this study (*Africonus* in green; *Lautoconus* in blue), including the sampling location. The four clades of the genus *Africonus* are identified in roman numbers (Abalde et al., 2017a) whereas the three clades of Lautoconus are identified by their typical radular tooth (Abalde et al., 2017b).

2017a; Abalde et al., 2017b). Thus, the three main clades in the genus *Lautoconus* and the four main clades recovered in *Africonus* were represented. For the genus *Lautoconus*, we studied one specimen of *Lautoconus guanche* (L_CG13) and *Lautoconus reticulatus* (L_1278) and two of *Lautoconus mercator* (L_1258 and L_1302; representing two different shell phenotype formerly classified as distinct species). In the case of *Africonus*, we included one specimen of *Africonus infinitus* (A_0025), *Africonus raulsilvai* (A_0031), *Africonus galeao* (A_0048), *Africonus verdensis* (A_0239), *Africonus boavistiensis* (A_0520), *Africonus miruchae* (A_0875), *Africonus antoniomonteroi* (A_0885) and *Africonus grahami* (A_1387), and two of *Africonus maioensis* (A_0055 and A_0039; representing two different shell phenotype formerly classified as distinct species). Sampling localities and voucher numbers for shells are shown in Table 1. Phylogenetic relationships including sequence divergences (branch lengths) are depicted in Fig. 1. All the specimens were adults and were dissected in a resting stage to remove the venom duct, which was preserved in RNALater (Invitrogen, Life technologies) at 4ºC during the sampling and -20ºC for the long term.

2.2. RNA extraction and sequencing

RNA extraction and sequencing were essentially performed as in (Abalde et al., 2018). First, each venom duct was incubated in a 2 ml eppendorf with 500 µl of TRIzol LS Reagent (Invitrogen, Life Technologies) and grinded with ceramic beads in a Praecellys Evolution tissue homogeneizer. Then, the solution was mixed with 100 µl of chloroform and centrifuged at 12000xg for 15 min at 4ºC. The supernatant was collected, mixed with 250 µL of isopropanol and left for precipitation overnight at -80ºC. Total RNA

Table 1. Specimens of *Africonus* and *Lautoconus* here analysed and main statistics of Illumina sequencing and assembly.

ID	Species	Country	Locality/Island	Voucher MNCN	SRA_accesion	Sequencing_date	Number_reads	%_clean_reads	Number_contigs	Number_blast_hits	Number_proteins	Number_conotoxins
A_0025	Africonus infinitus	Cabo Verde	Ponta do Pau Seco, Maio	15.05/78650	to be provided	13/3/14	35854397	98.52	76339	1095	206	175
A_0031	Africonus raulsilvai	Cabo Verde	Praia da Soca, Maio	15.05/78656	to be provided	28/10/13	56718528	100	99699	1249	223	189
A_0039	Africonus maioensis	Cabo Verde	Praia Santana, Maio	15.05/78664	to be provided	28/10/13	52523501	100	77336	783	169	140
A_0048	Africonus galeao	Cabo Verde	Navío Quebrado, Maio	15.05/78673	to be provided	21/12/16	28109709	100	50811	803	177	154
A_0055	Africonus maioensis	Cabo Verde	Navío Quebrado, Maio	15.05/78680	to be provided	28/10/13	44748977	100	102227	850	173	143
A_0239	Africonus verdensis	Cabo Verde	Tarrafal, Santiago	15.05/78864	to be provided	28/10/13	40237424	100	77906	1266	239	205
A_0520	Africonus boavistensis	Cabo Verde	Ervatao, Boa Vista	15.05/80413	to be provided	21/12/16	26715260	100	39935	797	199	175
A_0875	Africonus miruchae	Cabo Verde	Terrinha Fina, Sal	15.05/79784	to be provided	21/12/16	24097307	100	62006	615	131	108
A_0885	Africonus antoniomonteroi	Cabo Verde	Pedra Lume, Sal	15.05/79794	to be provided	21/12/16	26026957	100	84001	731	157	122
A_1387	Africonus grahami	Cabo Verde	Calhau, São Vicente	15.05/78549	to be provided	21/12/16	22718525	100	51601	850	180	159
L_1258	Lautoconus mercator	Senegal	Almadies	15.05/78419	to be provided	8/3/16	28883175	100	69501	631	143	126
L_1278	Lautoconus reticulatus	Senegal	Ngor	15.05/78439	to be provided	21/12/16	24263358	100	52496	479	109	89
L_1302	Lautoconus mercator	Senegal	Ndayane	15.05/78463	to be provided	8/3/16	28392465	100	78580	783	176	157
L_CG13	Lautoconus guanche	Spain	Playa del Cable, Lanzarote	–	to be provided	8/3/16	29973740	100	87516	815	175	150

was purified using the Direct-Zol RNA miniprep kit (Zymo Research, Irvine) following manufacturer's instructions.

Dual-indexed cDNA libraries were constructed for each sample using the TruSeq RNA library Prep kit v2 (Illumina, San Diego) at Sistemas Genómicos (Valencia, Spain) following manufacturer's instructions. The quality of the libraries was checked with the TapeStation 4200, High Sensitivity Assay, and the quantity determined by real-time PCR in LightCycler 480 (Roche). The pool of libraries (including samples of other cone snail species) was split into several runs of paired-end sequencing (2x100bp) in an Illumina HiSeq2500 (two flowcells per run) following the standard procedures at Sistemas Genómicos (Valencia, Spain).

2.3. RNA assembly and conotoxin identification

The raw reads corresponding to the different individuals were sorted using the corresponding library indices, which were removed using Cutadapt v.1.3 (Martin, 2011). Raw read quality was checked using FastQC v.0.10.1 (Andrews, 2010), and the assembly was performed using Trinity v.2.6.6 (Grabherr et al., 2011) with default settings (minimum contig length = 200bp, sequence identity threshold = 0.95) and the --trimmomatic option active with default parameters. The raw reads of all transcriptomes are available at the SRA database (Table 1).

All conotoxin precursors, hormones, and associated proteins publicly available in different databases (GenBank release 222 (Benson et al., 2013), Uniprot release 2017_09 (UniProt, 2015) and ConoServer release 30/10/2017 (Kaas et al., 2012) were downloaded in October 30th, 2017 and concatenated into a single fasta file. Duplicated sequences were removed, and the resulting file was formatted as a Blast database using Blast+ (Camacho et al., 2009) to create the custom reference database.

All putative conotoxin precursor, hormones, and associated protein sequences in the assembled transcriptomes were identified using BLASTX over the reference database (e-value: 1e-5). The selected sequences were manually inspected and compared against the most similar sequences in the reference database, and translated into the appropriate open reading frame (ORF). All the sequences considered as false positives or assembly artifacts (showing internal stop codons and chimeras), those that were duplicated or highly truncated (missing >55% of the estimated length of the reference protein), and those showing low coverage values were discarded. We implemented an extra curation step consisting on TBLASTX searches over the nr database in GenBank to discard wrong ORF assignations.

The remaining sequences constituted our working list of conotoxin precursors, hormones, and associated proteins (see Suppl. Mat. File 1 and Suppl. Mat. Table 1). The three domain structure and cysteine frameworks of conotoxin precursor alignments were inferred using Conoprec (Kaas et al., 2010). The different proteins were assigned to a given superfamily by comparison with best-hit results using BLASTP searches against GenBank, and in the case of the conotoxin precursors, taking into consideration the percentage of identity in the signal region using a general threshold

Table 2. New signal sequences of conotoxin precursors here described, including the cysteine framework and main Blast resul

				Best-hit known superfamily		
Unassigned superfamily	Signal	Cysteine framework	Also found in:	Superfamily	% coverage	% Identity
1	MNCLQPLLVLLLISTITA	XIII	betulinus	M	65	28.81
2	MSGTMIVLLAVLLLVDLSTS	VI/VII	*betulinus*	O3	82	37.29
3	MPGSRVALLAFLLLLSLVTNLQG	VI/VII	*betulinus, leopardus*	O3	25	62.5
4	MTMDMKMTFSGFVLVVLVTTVVG	VIII	*betulinus, praecellens, andremenezi*	—	—	—
5	MMTLRHVLLFTLLLLPLATIR	XXII	*betulinus*	A	68	31.37
6	MLSVFTVVWVLTTAMMMTDVTFQ	C7C6C8C3C1C5C23C8C18C1C10C	*praecellens, andremenezi*	I2	45	24.59
7	MWSGKDQAAFLALVLMVVGAHSTTA	IX	*praecellens, andremenezi*	—	—	—

of 70% (Robinson and Norton, 2014). We further checked the correct identification of all conotoxin precursor superfamilies by aligning all the signal regions and building with neighbor-joining a guide tree (Supp. Mat. Fig. 1) based on uncorrected *p* distances on ClustalW (Thompson et al., 1994). Within each superfamily, sequences were assigned to different groups of paralogy based on the sequence divergence at the pro-peptide region, different cysteine frameworks in the mature peptide, and the presence of clades in the reconstructed guide tree (Supp. Mat. Fig. 1). Those sequences that did not match any previously reported conotoxin precursor superfamily were considered unassigned superfamilies and described here. The nucleotide sequences of all venom proteins here identified are available at Genbank under accession numbers XXXX-XXXX.

2.*4. Comparative analyses of venom composition*

The final list of conotoxin precursors for each species was pairwise compared and common sequences were detected using the ClustalW algorithm as implemented in Geneious® 8.0.3. All sequences that were common to two or more species were mapped onto the phylogeny using parsimony ancestral character reconstruction as implemented in Mesquite v 3.6 (Maddison and Maddison, 2018).

In order to infer venom composition similarities between species and genera, we run in R (R Core Team, 2013), six principal component analyses (PCAs) comparing at the paralog group and superfamily levels: 1) the presence or absence of conotoxin precursor superfamilies; 2) the relative abundance of each superfamily in terms of number of different sequences; and 3) the relative expression level of each superfamily measured as transcript per million (TPMs) estimates (see below).

2.5. Expression analyses

Relative expression levels for each individual were calculated by mapping the raw reads to the nucleotide sequence of each contoxin precursor using Bowtie 2 (Langmead and Salzberg, 2012), and the values transformed to TPM estimates using RSEM (Li and Dewey, 2011) as implemented in Trinity v.2.6.6 (Grabherr et al., 2011). In order to identify those conotoxin superfamilies that could be differentially expressed between *Lautoconus* and *Africonus*, we run the EBSeq software (Leng et al., 2013), that estimates the posterior probability of being differentially expressed (PPDE), using all the specimens of each genera as biological replicates. We considered as differentially expressed all those conotoxin precursor superfamilies with a PPDE > 0.95 and with a fold change above 32 (calculated as $\log_2$ RealFC $\geq$ 5). The same type of analysis was performed to identify those superfamilies significantly expressed in the comparison between vermivory, (using the 14 specimens of West Africa as replicates) and piscivory (using the three individuals of *C. ermineus* from (Abalde et al., 2018). In both comparisons, we run an ANOVA test in R (R Core Team, 2013) over those superfamilies identified as differentially expressed to take into consideration variance among replicates and further confirm the statistical significance of the results.

3 | RESULTS

3.1. Sequencing and assembly

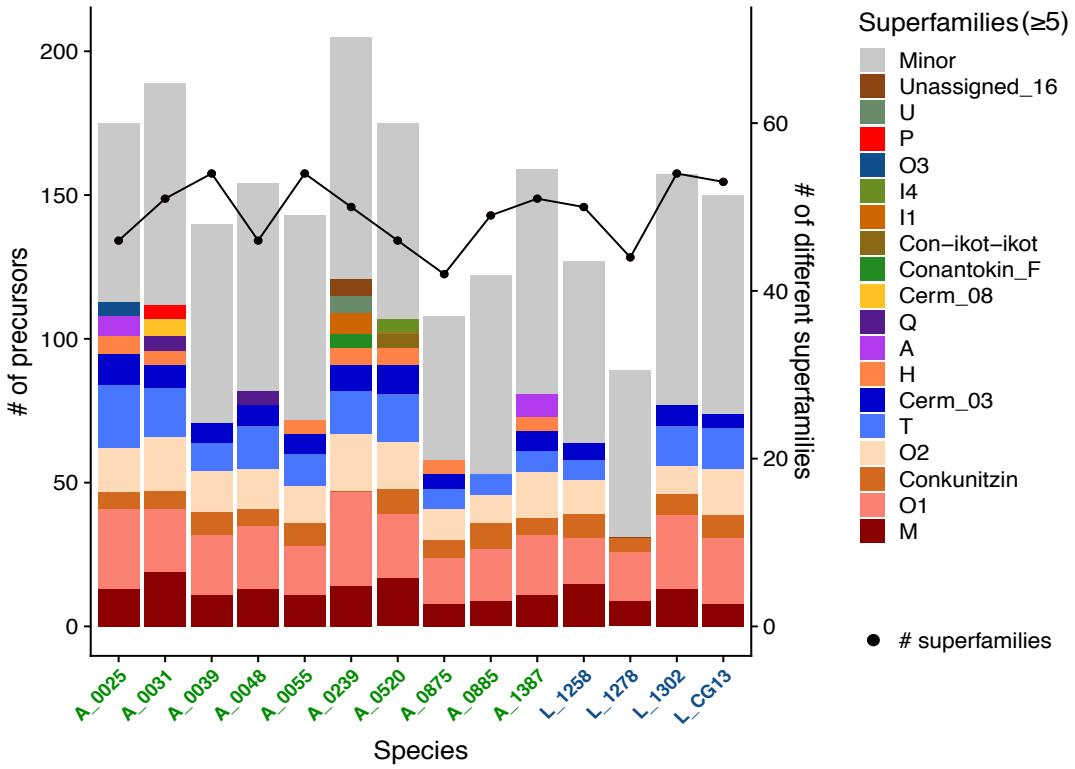


Figure 2. Venom composition of each of the 14 specimens studied. The bars represent the total number of conotoxin precursors and hormones. The proportion of those superfamilies with more than five members is shown in colours. The bar line represents the number of different superfamilies identified in the venom. The species codes in green belong to the genus *Africonus*, and those in blue to *Lautoconus*.

The RNAs of 14 samples corresponding to 12 species of the genera *Lautoconus* and *Africonus* were sequenced and their transcriptomes assembled. The main statistics associated to the sequencing and assembly procedures are summarized in Table 1. The number of raw reads sequenced per sample varied between 22.7 and 56.7 millions with a mean of 33.5 millions, and most reads were kept after cleaning (Table 1). The number of contigs generated after the assembly varied between 39,935 and 102,227 with a mean of 72,139. The BLASTX searches retrieved between 479 and 1269 putative conotoxin precursors, hormones and other venom proteins per species, and after all curation steps, we kept between 109 and 239 of them with a mean of 175,5 (Table 1). The number of conotoxin precursors varied between 89 and 205 with a mean of 149.4.

3.2. Venom cataloguing of West African cones

A total of 2,575 transcripts were identified in the venom gland transcriptomes of the 14 cone specimens: 2,070 were conotoxin precursors, 71 were hormones, and 434 were designed as other venom proteins. The species that presented the highest diversity of contoxin precursors were *A. verdensis* (205) and *A, raulsilvai* (189) whereas the least diversity was found in *L. reticulatus* (89; Table 1 and Fig. 2). All conotoxin precursors were classified into 60 distinct superfamilies and 142 groups of paralogy (hereafter "families") taking into consideration sequence divergences in the signal rand pro-peptide regions and the clades recovered in the reconstructed guide tree. Most of the conotoxin precursor superfamilies were recovered as monophyletic in this guide tree, with the exception of the M superfamily (two lineages), Cerm_03 (three), Q (three), and N (two; Supp. Mat. Fig. 1). The superfamilies that presented more diversity of members

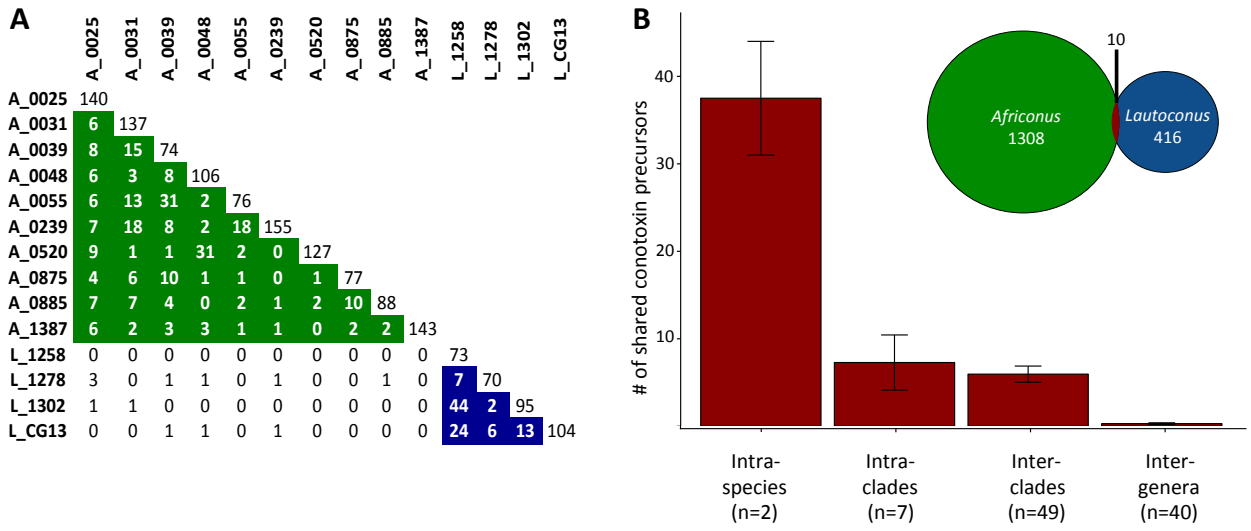


Figure 3. Differences in venom composition along the phylogeny. A) pairwise comparisons of the number of shared identical conotoxin precursors. The number of conotoxins exclusive for an individual is represented in the diagonal. B) Average number of identical conotoxins shared by individuals, intra- and interclade species, and genera. The number of comparisons is shown below each bar. A Venn diagram representing the number of unique and common conotoxin precursors for both genera is shown as inset.

were conkunitzin (13), O2 (11), and O1 (9). A total of 70 peptides could not be assigned to any known superfamily, and were grouped into seven new unassigned superfamilies (their signal sequences, cysteine frameworks and other features are reported in Table 2). Several precursors previously reported as valid conotoxin precursor superfamilies such as R, W, Z (Lavergne et al., 2013) and Cerm_17 (Abalde et al., 2018), among others, were found to be fragments of other proteins once the right ORFs were identified using TBLASTX.

The species that presented the highest diversity of conotoxin precursor superfamilies were *A. maioensis* (A_0055, 54; A_0039, 54) and *L. mercator* (54) whereas the least complex venom was found in *A. miruchae* (42). The superfamilies with highest number of members were O1, O2, T, M, Conkunitzin, and Cerm_03 (Fig. 2). They were found in all the species with the following exceptions: *A. verdensis* lacked the Conkunitzin; *A. antoniomonteroi* had no Cerm_03; and *L. reticulatus* lacked the O2, T and Cerm_03 (Fig. 2). All other conotoxin precursor superfamilies in the species of the genus *Lautoconus* had less than five members. However, the species in the genus *Africonus* (except *A. antoniomonteroi* and the specimen A_0039 of *A. maioensis*) showed another 12 superfamilies with five or more distinct conotoxin precursors. These expanded superfamilies were differently distributed: four were exclusive to *A. verdensis* (Conantokin F, I1, U, and Unassigned_16); two to *A. raulsilvai* (Cerm_08 and P); two to *A. boavistiensis* (Con-ikot-ikot and I4); and one to *A. infinitus* (O3); the remaining (H, A, and Q) were found in more than one species (Fig. 2). The 14 species presented up to 71 hormone sequences that were classified into six superfamilies (none of them present five or more members): Conopressin, Conorfamide, Insulins 1-5, Prohormone-4, Thyrostimulin hormone alpha, and Thyrostimulin hormone beta 5 (Suppl. Mat. File 1 and Suppl. Mat. Table 1).

Finally, we identified up to 434 transcripts that were assigned to 28 protein families of various function. Among them, the Protein Disulfide Isomerase (81 sequences), Ferritin (50) and Conodipine (41) were the most diverse. Interestingly, we found in the venom, several members of a conserved superfamily of cysteine-rich secretory, antigen 5, and pathogenesis-related 1 proteins (CAP) that are often secreted and have a

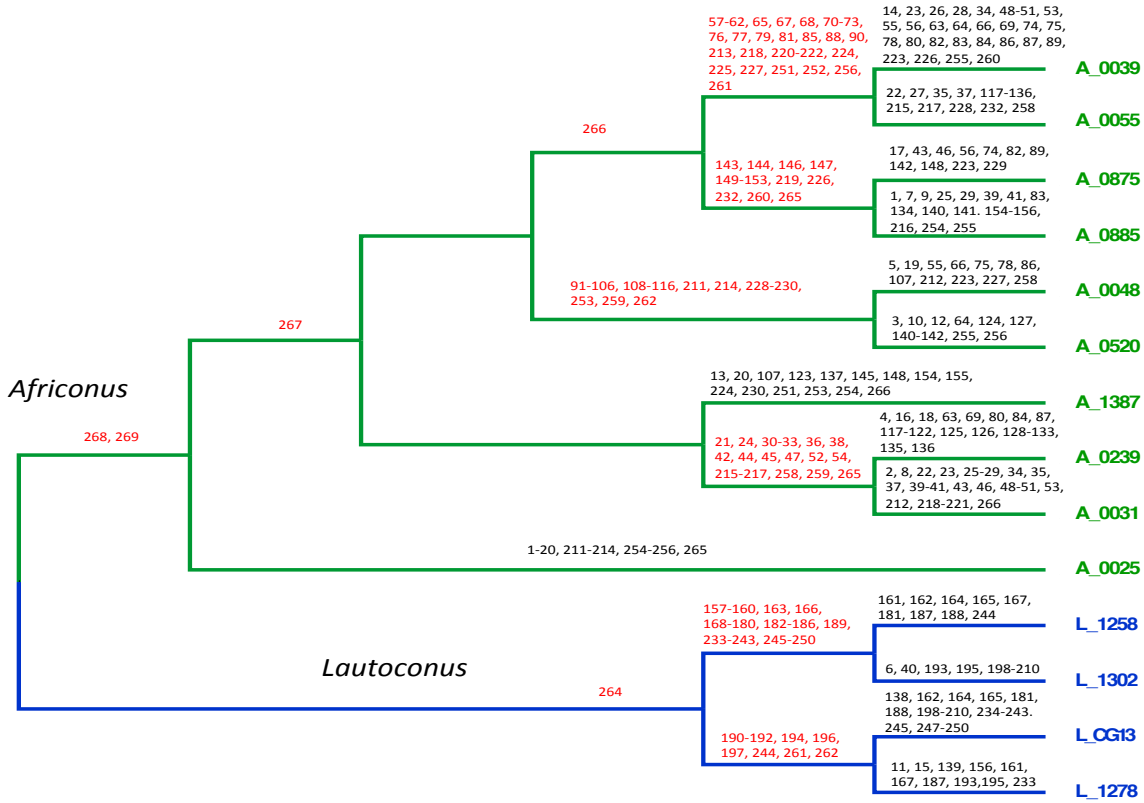


Figure 4. Reconstruction under parsimony of the conotoxin precursors present in the different common ancestors in the phylogeny of cones from West Africa. The numbers in the nodes correspond to conotoxin precursors shared by at least two taxa and shown in Suppl. Mat. Table 2.

protease activity with an extracellular endocrine or paracrine function in a wide range of organisms including venomous ones such as ants, wasps and snakes (Bateman et al., 2004). These proteins were previously reported in the molluscivorous *Cylinder textile* and the vermivorous *Conus marmoreus* and may be important for venom function ((Milne et al., 2003; Hansson et al., 2006); Suppl. Mat. Fig. 2).

3.3. Variations in venom composition according to phylogenetic divergence

The venom composition of the 14 specimens was pairwise compared at different hierarchical levels (Fig. 3). For two species (*A. maioensis* and *L. mercator*), we could compare venom composition between individuals. The two specimens of *A. maioensis* shared 31 conotoxin precursors, which represent 37-39% of the total sequences. The two specimens of *L. mercator* had 44 conotoxin precursors in common (46-60%; Fig. 3). The number of shared sequences between pairs of species from the same clade within *Africonus* varied between 1 to 31, with a mean of 8.7 (8% of the mean total sequences). Four pairwise comparisons at this level rendered common sequences more than average: *A. boavistensis* versus *A. galeao* (clade III; 31 shared sequences), *A. verdensis* versus *A. raulsilvai* (Clade II; 18), *A. maioensis* A_0039 versus *A. miruchae* (Clade IV; 10), and *A. antoniomonteroi* versus *A. miruchae* (Clade IV; 10). The number of shared sequences between pairs of species from different clades within the same genus varied between 0 and 18 with a mean of 4.61 (4.1% of the mean total sequences). Four pairwise comparisons at this level rendered common sequences more than average: *A. verdensis* versus *A. maioensis* A_0055 (18 shared sequences), the two specimens of *A. maioensis* versus *A. raulsilvai* (13 and 15), and the two specimens of *L. mercator* versus *L. guanche* (13-24; Fig. 3). The species *A. grahami* was the one sharing fewer sequences with other species (1-6; Fig. 3). The number

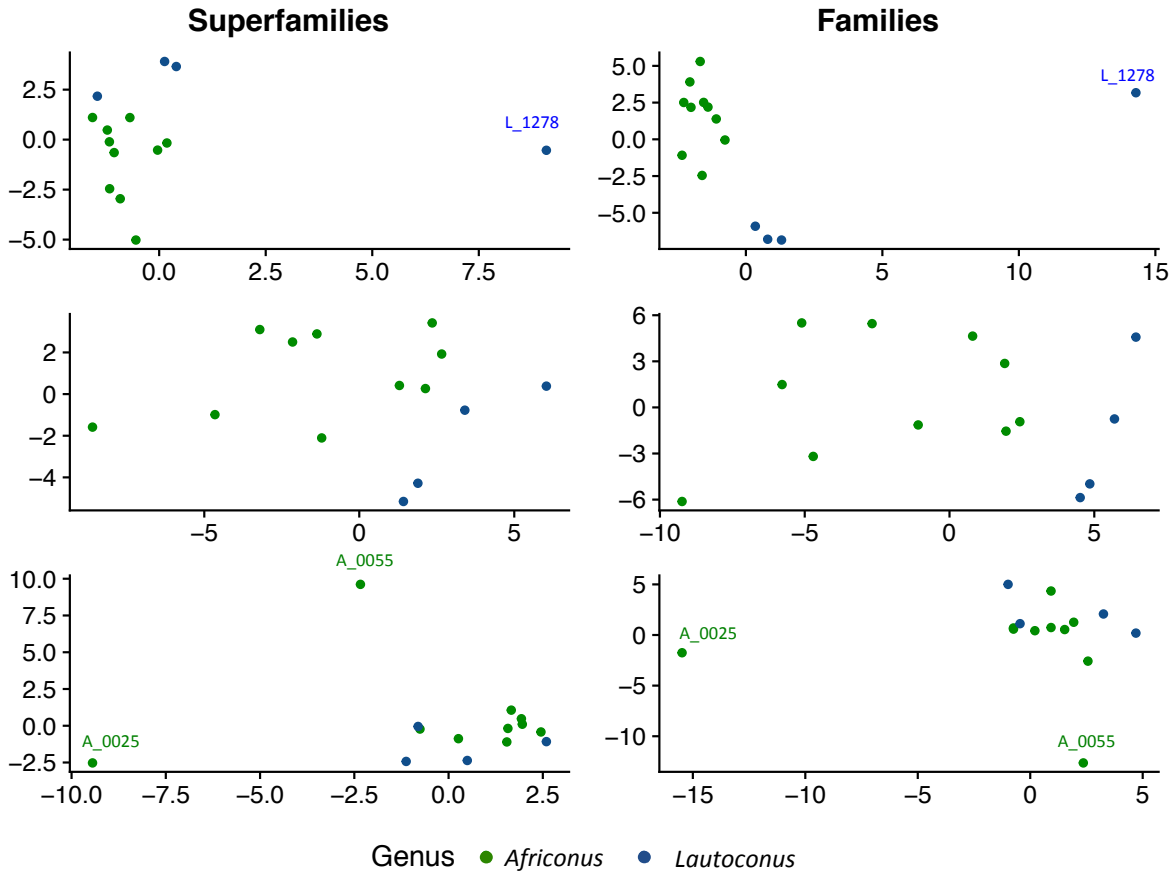


Figure 5. Principal Component Analysis comparing the venom composition among species and genera. The venom composition was defined as presence/ absence of conotoxin precursors (first row), number of members (second row), and relative expression levels (third row). The three comparisons were made at the superfamily and family (groups of paralogy) levels.

of shared sequences between genera was nine (0.7-2% of the mean total sequences). The total diversity of conotoxin precursors in *Africonus* triplicated that of *Lautoconus* (1367 versus 426; Fig. 3).

An ancestral reconstruction analysis under parsimony was performed to determine contoxin precursors in most recent common ancestors in the phylogeny and detect instances of convergence (Fig. 4). According to the reconstruction, the ancestor of the species belonging to clade IV of *Africonus* had the Cerm_10 superfamily. The ancestor of Clade III had members of B2, C, I1, J, M, O1, O2, P, T, U, V, Y, Conkunitzin, Con-ikot-ikot, Rimp_01, Rimp_03, Rmil_01, Tpra_06, Pamg_02, Cerm_02, Cerm_03 and Cerm_011 superfamilies (Fig. 4). The ancestor of *A. raulsilvai* and *A. verdensis* in Clade II had members of Conantokin F, B2, F, H, M, O1, O2, O3, T, Conkunitzin, Cerm_03, Cerm_10, and Unassigned_07 superfamilies (Fig. 4). The ancestor of clades II-IV had a member of the O2 superfamily. The ancestor of all *Africonus* species had a member of O2 and Pmag_02 superfamilies. The ancestor of *L. guanche* and *L. reticulatus* had members of A2, M, O1, O3, Conkunitzin, Pmag_02, Rimp_01, Rimp_04, and Rmil_02 superfamilies (Fig. 4). The ancestor of all *Lautoconus* species had a member of the O2 superfamily. Finally, *A. raulsilvai* (clade II; Maio) and *A. maioensis* A_0039 (clade IV; Maio) shared aeveral conotoxin precursors (members of B1, I3, M, O1, Con-ikot-ikot, Crerm_08, Cerm_011, and Unaasigned_04 superfamilies) despite distantly related in the phylogeny.

Pairwise comparisons were also performed taking superfamily or family classification into consideration and subjected to PCAs. Three different metrics were

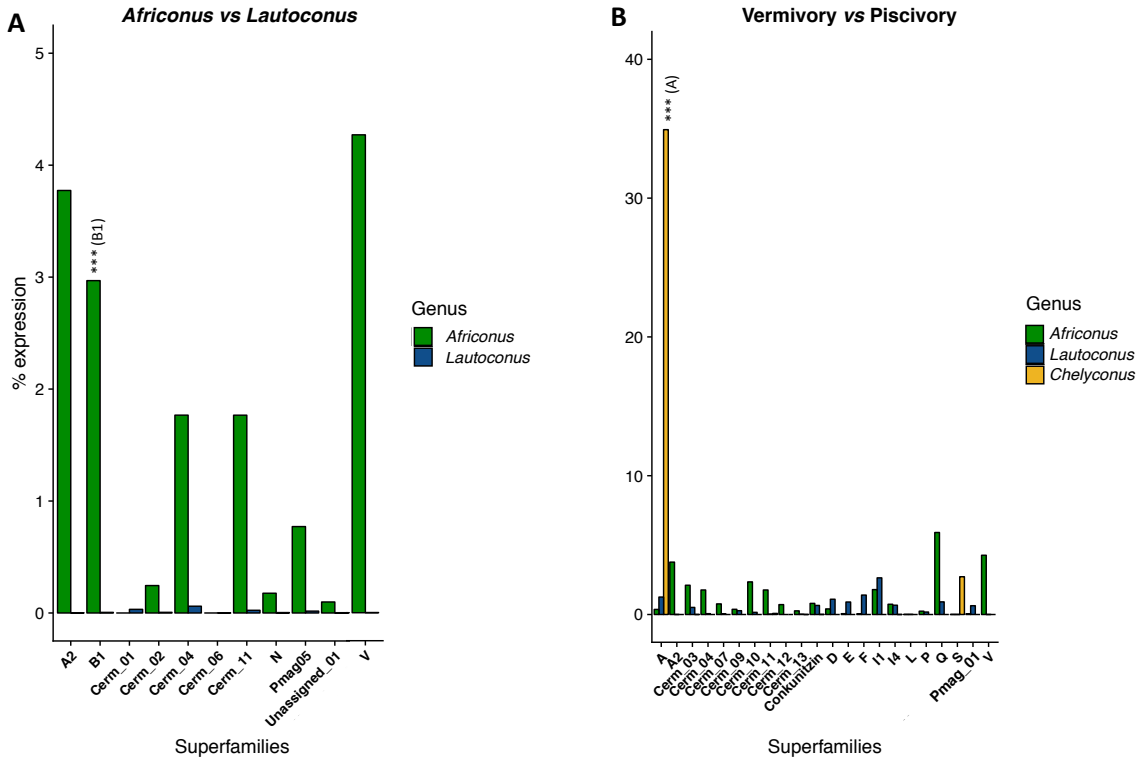


Figure 6. A) Differential Expression (measured in TPMs) of superfamilies between A) *Africonus* and *Lautoconus* and B) the 14 specimens here sequenced as biological replicates for vermivory and the three individuals of *Chelyconus ermineus* from (Abalde et al., 2018) representing piscivory, Bar plot depicting the superfamilies differentially expressed are shown. The genus *Africonus*, *Lautoconus*, and *Chelyconus* are depicted in green, blue, and yellow, respectively. The three asteriks represent a significant *p*-value = 0 following the ANOVA analyses.

analyzed at both levels: presence/absence, number of members, and the relative expression levels (calculated as TPMs) (Fig. 5). According to the PCAs, species from each genera cluster together when the presence/ absence of conotoxin superfamilies and families is tested, although *L. reticulatus* is clearly an outlier; Fig. 5). The PCA of the relative abundance (number of members) of each superfamily/ family revealed no overlapping between genera (Fig. 5). Finally, relative expression levels are similar in both genera, although *A. infinitus* and *A. maioensis* A_0055 are outliers (Fig. 5).

3.4. Differential expression of conotoxins and hormones in the two genera and in vermivores versus piscivores

We estimated whether any superfamily was differentially expressed between the two genera using the species of each genus as biological replicates and the sums of the expression of each superfamily as variables. We detected 11 superfamilies differentially expressed between genera (Fig. 6): A2, B1, Cerm_01, Cerm_02, Cerm_04, Cerm_06, Cerm_11, N, Pmag_05, Unassigned_01, and V. All these superfamilies were overexpressed in *Africonus*, except Cerm_01 and Cerm_06, which were in *Lautoconus*. After an ANOVA test, only the overexpression of B1 superfamily in *Africonu*s remained significant (p-value = 0).

The same tests were performed to detect differential expression between the 14 individuals of vermivorous species here studied and three individuals of the Atlantic piscivorous species *C. ermineus* (Fig. 6). Using the same threshold as above, we found 22 superfamilies differentially expressed. Among them, only the A and S superfamilies were overexpressed in *Chelyconus ermineus*. The ANOVA test only identified the expression levels of the A superfamily as significantly different between diets (*p*-value = 0).

4 | DISCUSSION

At present, high-throughput sequencing techniques are producing massive amounts of raw read sequence data, which are assembled and automatically annotated using a variety of bioinformatic pipelines. However, correct annotation of genes relies entirely on the quality of the reference database (Salzberg, 2019). In particular, in the case of conotoxins, and given their intrinsic variability, several authors have warned about the need of inspecting carefully automated annotation results, since assembly artifacts could over-estimate final conotoxin diversity (Phuong et al., 2016; Abalde et al., 2018) and, even more dangerous, they could propagate once incorporated into updated reference databases (Li et al., 2017; Abalde et al., 2018; Salzberg, 2019). Here, we carefully inspected all the ORFs initially rendered by the BLASTX searches through manual comparison to previously published sequences obtained from cone venoms or from venom gland transcriptomes. At the assembly level, after mapping the reads back to the transcripts, we found in many instances, regions of the assembled transcript (particularly at the tails) mapped only by very few reads (sometimes even only one), which could represent errors of the sequencing process and led to frame shifts that generated spurious variability (Phuong et al., 2016). At the annotation level, we implemented a TBLASTX step that found instances of wrong ORF assignation. For example, the ORF of the R superfamily (Lavergne et al., 2013) when translated in a different frame corresponds to the proteasome subunit alpha, and the conotoxin precursor "Bt-23" from (Peng et al., 2016) corresponds to a 60S ribosomal protein (other examples are shown in Suppl. Mat. File 1)..

Remarkably, we found in different species several sequences that were identical, except for the presence of one or several copies in tandem of a duplicated fragment within the pro-peptide region. It is the case, for example, of the B1 superfamily (Conantokin) sequences A_0239_306, A_0039_179 and A_0055_201 (these and other examples are found in Suppl. Mat. file 1). The mapping of the reads over the corresponding regions of the transcript confirmed high coverage and, in some cases, the reads covered the entire duplicated region, thus discarding an assembly artefact (not shown). The signal, mature and post regions of the precursors are encoded each by a single exon in the genome, whereas the pro-peptide region could be encoded by up to six different exons (Phuong and Mahardika, 2018). Hence, one potential explanation for the observed duplications could be the inclusion more than once of the same exon during the formation of the mRNA, although confirming this hypothesis would require experimental validation.

The number of raw reads sequenced per sample varied between 23-57 millions. Generated sequence data fit well within the range recommended by previous studies testing the optimal sequencing depth for *de novo* assembly, which concluded that 20-30 million reads would report most of the expressed genes in a given tissue while minimizing the number of artifacts (Francis et al., 2013). The number of raw reads generated per species did not correlate neither with the number of assembled contigs nor with the number of conotoxin precursors, hormone, and other venom proteins. For example, the assembly of the *L. mercator* (L_1302) and *A. maioensis* (A_0039) transcriptomes rendered similar number of contigs (77-78,000) but started from 28 and 52 million raw reads, respectively. Similarly, *A. miruchae* and *L. reticulatus* started from a similar sequencing depth (24 million reads) but presented a 17.5% difference in the number of conotoxin precursors (108 and 89, respectively).

The number of conotoxin precursors present on the venom varied from 89 in *L. reticulatus* to 205 in *A. verdensis*. These numbers are in good agreement with those reported for other species of cones (Hu et al., 2011; Barghi et al., 2015; Peng et al., 2016; Li et al., 2017; Abalde et al., 2018; Jin et al., 2019). It has ben proposed that larger sets of conotoxins are associated to broader diets (Elliger et al., 2011; Phuong et al., 2016;

Phuong and Mahardika, 2018). Hence, the richer diversity of conotoxins of *A. verdensis* could reflect a broader diet, which hitherto is known to be based on worms but largely unstudied. In this regard, ecological studies on *Miliariconus miliaris* showed that the individuals of this species inhabiting the remote Eastern Island presented a considerably broader diet of worms, which could have evolved through ecological release in the absence of congeners (Duda and Lee, 2009). This could also be the case for *A. verdensis*, which is the only species inhabiting the island of Santiago in Cabo Verde. However, this plausible explanation does not seem to fully apply to *A. raulsilvai* and *A. infinitus*, with 189 and 175 different conotoxin precursors, respectively. In these cases, while the observed conotoxin precursor diversity could still reflect a broader diet, it could not have evolved by ecological release, as both species are endemic to Maio, one of the islands of Cabo Verde where more cone species cohabit (Abalde et al., 2017a).

Conotoxins are well known for their accelerated rates of evolution, which in turn generate high sequence divergences even between individuals of the same species (Peng et al., 2016; Abalde et al., 2018; Jin et al., 2019), and are the basis of the reported general lack of common peptides between cone species, and the extended notion that virtually each species has produces a unique venom cocktail (e.g., (Gao et al., 2017). However, thus far, these conclusions were based mostly on comparisons between distantly related species (for an exception see (Li et al., 2017), whereas the present study brings the opportunity of comparing two clades (genera) of up to 12 closely related species sharing relatively recent common ancestors. Individuals of the same species (although representing distinct shell phenotypes previously described as species) showed only half (*L. mercator*) to one-third (*A. maioensis*) common conotoxin precursor sequences. These proportions are similar to those reported for intraspecific comparisons in *D. betulinus* (Peng et al., 2016), *Rhombiconus imperialis* (Jin et al., 2019) or *C. ermineus* (Abalde et al., 2018). The proportion of shared sequences decreased substantially for the pairwise comparisons between species (but within the range of 2-9% reported for sister species of the genus *Turriconus*; (Li et al., 2017), although there were still enough common sequences to infer which conotoxin precursors were likely present in most recent common ancestors along the phylogeny. In particular, O2 superfamily was found to be the only one already present in the ancestors of *Africonus* and *Lautoconus*. Therefore, our results support that a phylogenetic signal exists in conotoxins above the species level, but it is quickly eroded as lineages diverge and virtually almost no sequence is shared between closely-related genera comparisons (Phuong et al., 2016). On the other hand, there are some instances in which identical conotoxin precursor sequences are found even in species from distantly related genera (see Suppl. Mat. File 1), indicating that those sequences either are subjected to strong balancing selection or reflect cases of convergent evolution. The rather erratic distribution of these sequences in the phylogeny of cones favors the latter hypothesis. Moreover, the presence of common conotoxin precursors in *A. raulsilvai* and *A. maioensis* A_0039, which are distantly related in the phylogeny of cones of Cabo Verde (clades II and IV within *Africonus*; (Abalde et al., 2017a) but inhabit two close bays (Soca and Santana, respectively; Table 1) in the northwest of Maio Island, suggests functional convergence to similar diets (Remigio and Duda, 2008).

The analyses of the composition of the venoms of West African cones in terms of number and type of superfamilies showed that almost all species had a core set of six superfamilies, which are characterized by having five or more members: M, O1, O2, T, Conkunitzin, and Cerm_03. The wider presence of the former four superfamilies in any cone and always showing similar levels in diversity of members (e.g. (Terrat et al., 2012; Lavergne et al., 2013; Peng et al., 2016; Phuong et al., 2016; Li et al., 2017; Robinson

et al., 2017; Abalde et al., 2018) may suggest that the ancestor of living cones already had this core set and that having members of these superfamilies is essential for triggering the minimum physiological responses leading to the capture of a prey, regardless of whether it is a worm, a snail or a fish.

In addition, in West African cones, we found up to 13 different groups of paralogy or families of conkunitzins present in most cases both in *Africonus* and *Lautoconus*. The conkunitzins block voltage-gated potassium channels and were first described in *Pionoconus striatus* (Bayrhuber et al., 2005). They have been also identified in other piscivorous cones such as *C. ermineus* (Abalde et al., 2018) and *G. geographus* (Dutertre et al., 2014) as well as in the vermivore *Dendroconus betulinus* (Peng et al., 2016). Conkunitzins belong to a larger superfamily of kunitz-type fold peptides, which are ubiquitous serine protease inhibitors found in different animals (Ranasinghe and McManus, 2013). We also found a great diversity of paralog groups belonging to Cerm_03 superfamily. This superfamily was recently described in *C. ermineus* (Abalde et al., 2018), and has a typical precursor structure with the three domains and a mature peptide with a cysteine framework type XIV. It has been also found in several vermivore species including *D. betulinus*, *Turriconus praecellens*, and *Elisaconus litteratus* but without the variety of paralogs identified here and in *C. ermineus*. The function of the mature peptide remains to be determined.

While *Lautoconus* species showed an expanded number of members only for the above-mentioned six superfamilies, most *Africonus* species presented in addition the expansion of H superfamily. Not much is known about the function of the conotoxins (with a cysteine frame work VI/VII) of this superfamily, which was first identified in the vermivorous *C. marmoreus* (Dutertre et al., 2013) and constitutes a large proportion of conotoxin mRNA transcripts in the venom gland of the molluscivorous *Cylinder victoriae* (Robinson et al., 2014). Interestingly, *A. verdensis*, the only cone from Santiago Island, is the species with more number of conotoxin precursor superfamilies showing expansion of their members, and these expanded superfamilies do not coincide with those of *A. raulsilvai* from Maio Island, which is its sister species (together with *Africonus gonsaloi*), and the second species with highest diversity of expanded families. The larger set of expanded superfamilies in *Africonus* versus *Lautoconus* could be explained by the species radiation in the archipelago and the potential opportunities for local diet specializations, thus increasing the variability in the compositions of the venoms of these insular species (Abalde et al., 2017a). This would be in agreement with the lack of diversification of *L. guanche* in Canary Islands, which is reflected in a lack of diversity of expanded superfamilies.

PCA has been used in several other animal groups to summarize the information related to venom composition (Gibbs et al., 2013; Lomonte et al., 2014) but not in cones to the best of our knowledge. All PCAs recovered generally non-overlapping patterns for *Africonus* and *Lautoconus*, indicating that species more closely related tend to have the same conotoxin precursor superfamilies, in similar proportions and expressed at similar levels (contradicting the results of (Duda and Remigio, 2008), who reported that differences in expression of conotoxins of closely related Indo-Pacific vermivorous species could not be explained by phylogeny but by functional convergence). The species *L. reticulatus* was as an outlier in the PCA of presence/absence of superfamilies, which could be because it was the one with the least number of conotoxin precursors. Two outliers were found in the PCA of relative expression levels. In the case of *A. infinitus*, the differences in expression could be partly related to the fact that this species from Maio Island was the only representative of its lineage. In the case of the individual A_0055 of *A. maioensis* the interpretation is not that straightforward since the other individual of this species clusters with most species, and they come from neighboring populations.

Although the exact worm species eaten by the different species of *Lautoconus* and *Africonus* are unknown, at least the three clades described within genus *Lautoconus* correlate with different morphologies of the radular teeth suggesting subtle diet specializations (Abalde et al., 2017b). Moreover, the two genera have a dissimilar geographic distribution, with *Africonus* confined to the archipelago of Cabo Verde, and *Lautoconus* present in both the continental coast and the Canary Islands, opening the possibility of different processes of local adaptation. We tested whether the two genera showed differential expression of their venom components, which could be correlated with diet adaptations. Up to 11 superfamilies were differentially expressed, most of them overexpressed in *Africonus*. Most of these superfamilies were only recently described through comparative transcriptomics and only three have been known for longer time. The B1 superfamily (Conantokin) was originally described in the piscivorous *G. geographus* and reported to provoke a "sleeping" phenotype in vertebrates (Olivera et al., 1985). Although this superfamily has been identified in other vermivorous species, its function has not been characterized (Lu et al., 2014; Robinson and Norton, 2014). The N superfamily has been described in the molluscivorous *C. marmoreus*, although its function is unknown (Dutertre et al., 2013; Robinson and Norton, 2014). Finally, the V superfamily was first identified in the venom of the vermivorous *Virgiconus virgo* (Peng et al., 2008) and later in other species. To date, there is no information regarding the structure or function of this superfamily (Robinson and Norton, 2014). Nonetheless, results on differential expression should be interpreted with caution as in some instances the captured signal could rely on the high expression level of just one biological replicate, and may not be considering the variance among replicates. To correct such potential bias, we run an ANOVA test and found that only the B1 superfamily presented significantly different expression between genera.

Similarly, we tested the presence of differential expression between piscivorous (three individuals of *C. ermineus*; (Abalde et al., 2018) and vermivorous cones in West Africa. We found two superfamilies differentially overexpressed in *C. ermineus* (A and S) and 20 in the vermivorous cones. The importance of having different members of the A superfamily for hunting fish has been highlighted previously for several cone species (Lopez-Vera et al., 2007; Olivera et al., 2015; Hoggard et al., 2017; Abalde et al., 2018). The S superfamily was first identified in *G. geographus* and found to inhibit neurotransmitter receptors (England et al., 1998). Later, it was reported as minor component of different cone species, not all necessarily hunting on fish. After the ANOVA test, only the expression of the A superfamily was identified as significantly different between diets.

Acknowledgements

We are grateful to Cabo Verde biology students Paulo Vasconcelos and Stiven Delgado for their valuable help during sampling in Santa Luzia (Cabo Verde). We thank Dr Rui Freitas from the Universidade de Cabo Verde (UniCV) for his continuous support of our research in Cabo Verde. We are indebted to Dr Iderlindo Silva dos Santos and Dra. Sonia Monteiro de Pina Araujo from the Direcçao Nacional do Medio Ambiente of the Ministério do Ambiente, Habitação e Ordenamento do Território (MAHOT) of the Republic of Cabo Verde for their help with collecting permits (Autorizações 07/2013, 26/2013, 01/2104, 04/2015, and 03/2016). We thank Jesús Marco and Aida Palacio, who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing assembling and phylogenetic analyses.

Funding

This work was funded by the Spanish Ministry of Economy, Industry and Competitiveness (CGL2013-45211-C2-2-P

and CGL2016-75255-C2-1-P [AEI/FEDER, UE]) to R.Z.; BES-2014-069575 to S.A.

Authors' contributions

R.Z. conceived the study; M.J.T., C.M.L.A., and R.Z. obtained the samples; S.A. generated, assembled and annotated the trancriptome sequences; S.A. and R.Z. analyzed the sequence data. All authors participated in the writing of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Data availability

Sequence data are available at NCBI (SRA and GenBank). Supplementary Material available at:

https://docs.wixstatic.com/ugd/42cff7_b64fdd27335a488a9a97ce9040f7b6a9.pdf
https://docs.wixstatic.com/ugd/42cff7_b790267e56194c11b3136a9dfdcd94e1.xlsx?dn=Suppl.%20Mat.%20Table%201.xlsx
https://docs.wixstatic.com/ugd/42cff7_3c420c46f2c2489fb540ff115fbc1d9e.xlsx?dn=Suppl.%20Mat.%20Table%202.xlsx
https://docs.wixstatic.com/ugd/42cff7_f2729f1eba304ab8b5023ed5c27d56fc.pdf

5 | REFERENCES

Abalde, S., Tenorio, M.J., Afonso, C.M.L., Uribe, J.E., Echeverry, A.M., Zardoya, R., 2017a. Phylogenetic relationships of cone snails endemic to Cabo Verde based on mitochondrial genomes. BMC Evolutionary Biology 17, 231.

Abalde, S., Tenorio, M.J., Afonso, C.M.L., Zardoya, R., 2017b. Mitogenomic phylogeny of cone snails endemic to Senegal. Molecular Phylogenetics and Evolution 112, 79-87.

Abalde, S., Tenorio, M.J., Afonso, C.M.L., Zardoya, R., 2018. Conotoxin Diversity in *Chelyconus ermineus* (Born, 1778) and the Convergent Origin of Piscivory in the Atlantic and Indo-Pacific Cones. Genome Biology and Evolution 10, 2643-2662.

Ahorukomeye, P., Disotuar, M.M., Gajewiak, J., Karanth, S., Watkins, M., Robinson, S.D., Florez Salcedo, P., Smith, N.A., Smith, B.J., Schlegel, A., Forbes, B.E., Olivera, B., Hung-Chieh Chou, D., Safavi-Hemami, H., 2019. Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. Elife 8, e41574.

Aman, J.W., Imperial, J.S., Ueberheide, B., Zhang, M.-M., Aguilar, M., Taylor, D., Watkins, M., Yoshikami, D., Showers-Corneli, P., Safavi-Hemami, H., Biggs, J., Teichert, R.W., Olivera, B.M., 2014. Insights into the origins of fish hunting in venomous cone snails from studies of Conus tessulatus. Proceedings of the National Academy of Sciences USA 112, 5087-5092.

Andrews, S., 2010. FastQC: a quality control for high throughput sequence data. available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/; last accessed September 05, 2018.

Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2015. High conopeptide diversity in *Conus tribblei* revealed through analysis of venom duct transcriptome using two High-Throughput sequencing platforms. Marine Biotechnology 17, 81-98.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., Studholme, D.J., Yeats, C., Eddy, S.R., 2004. The Pfam protein families database. Nucleic Acids Research 32, D138-141.

Bayrhuber, M., Vijayan, V., Ferber, M., Graf, R., Korukottu, J., Imperial, J., Garrett, J.E., Olivera, B.M., Terlau, H., Zweckstetter, M., Becker, S., 2005. Conkunitzin-S1 is the first member of a new Kunitz-type neurotoxin family. Structural and functional characterization. Journal of Biological Chemistry 280, 23766-23770.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013. GenBank. Nucleic Acids Research 41, D36-42.

Bergeron, Z.L., Chun, J.B., Baker, M.R., Sandall, D.W., Peigneur, S., Yu, P.Y., Thapa, P., Milisen, J.W., Tytgat, J., Livett, B.G., Bingham, J.P., 2013. A 'conovenomic' analysis of the milked venom from the mollusk-hunting cone snail *Conus textile*--the pharmacological importance of post-translational modifications. Peptides 49, 145-158.

Binford, G., 2001. Differences in venom composition between orb-weaving and wandering Hawaiian *Tetragnatha* (Araneae). Biological Journal of the Linnean Society 74, 581-595.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.

Chang, D., Duda, T.F., Jr., 2012. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Molecular Biology and Evolution29, 2019-2029.

Chang, D., Duda, T.F., Jr., 2016. Age-related association of venom gene expression and diet of predatory gastropods. BMC Evolutionary Biology 16, 27.

Conticello, S.G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z., Fainzilber, M., 2001. Mechanisms for evolving hypervariability: the case of conopeptides. Molecular Biology and Evolution 18, 120- 131.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde islands. Systematic Biology 54, 634-650.

Cunha, R.L., Lima, F.P., Tenorio, M.J., Ramos, A.A., Castilho, R., Williams, S.T., 2014. Evolution at a different pace: distinctive phylogenetic patterns of cone snails from two ancient oceanic archipelagos. Systematic Biology 63, 971-987.

Duda, T.F., Jr., 2008. Differentiation of venoms of predatory marine gastropods: divergence of orthologous toxin genes of closely related *Conus* species with different dietary specializations. Journal of Molecular Evolution 67, 315-321.

Duda, T.F., Jr., Chang, D., Lewis, B.D., Lee, T., 2009. Geographic variation in venom allelic composition and diets of the widespread predatory marine gastropod *Conus ebraeus*. PLoS One 4, e6245.

Duda, T.F., Jr., Palumbi, S.R., 2004. Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*. Proceedings of the Royal Society series B 271, 1165-1174.

Duda, T.F., Jr., Remigio, E.A., 2008. Variation and evolution of toxin gene expression patterns of six closely related venomous marine snails. Molecular Ecology 17, 3018-3032.

Duda, T.F., Lee, T., 2009. Isolation and population divergence of a widespread Indo-West Pacific marine gastropod at Easter Island. Marine Biology 156, 1193-1202.

Duda, T.F., Palumbi, S.R., 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 USA 96, 6820-6823.

Duda, T.F., Rolán, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Molecular Ecology 14, 267-272.

Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nature Communications 5, 3521.

Dutertre, S., Jin, A.H., Kaas, Q., Jones, A., Alewood, P.F., Lewis, R.J., 2013. Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. Molecular Cell Proteomics 12, 312-329.

Elliger, C.A., Richmond, T.A., Lebaric, Z.N., Pierce, N.T., Sweedler, J.V., Gilly, W.F., 2011. Diversity of conotoxin types from *Conus californicus* reflects a diversity of prey types and a novel evolutionary history. Toxicon 57, 311-322.

England, L.J., Imperial, J.S., Jacobsen, R.B., Craig, A.G., Gulyas, J., Akhtar, M., Rivier, J., Julius, J., Olivera, B.M., 1998. Inactivation of a serotonin-gated ion channel by a polypeptide toxin from marine snails. Science 281, 575 - 578.

Espiritu, D.J.D., Watkins, M., Dia-Monje, V., Cartier, G.E., Cruz, L.J., Olivera, B.M., 2001. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon 39, 1899-1916.

Francis, W.R., Christianson, L., M., Kiko, R., Powers, M.L., Shaner, N.C., Haddock, S.H.D., 2013. A comparison across non-model animals suggests an optimal sequencing depth for *de novo* transcriptome assembly. BMC Genomics 14, 167.

Gao, B., Peng, C., Yang, J., Yi, Y., Zhang, J., Shi, Q., 2017. Cone snails: a big store of conotoxins for novel drug discovery. Toxins 9, E397.

Gibbs, H.L., Sanz, L., Sovic, M.G., Calvete, J.J., 2013. Phylogeny-based comparative analysis of venom proteome variation in a clade of rattlesnakes (Sistrurus sp.). PLoS One 8, e67220.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29, 644-652.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology 59, 307 - 321.

Hansson, K., Thämlitz, A.M., Furie, B., Furie, B.C., 2006. A single γ-carboxyglutamic acid residue in a novel cysteine-rich secretory protein without propeptide. Biochemistry 45, 12828 - 12839.

Hoggard, M.F., Rodriguez, A.M., Cano, H., Clark, E., Tae, H.S., Adams, D.J., Godenschwege, T.A., Mari, F., 2017. *In vivo* and *in vitro* testing of native alpha-conotoxins from the injected venom of *Conus purpurascens*. Neuropharmacology 127, 253-259.

Hu, H., Bandyopadhyay, P.K., Olivera, B.M., Yandell, M., 2011. Characterization of the *Conus bullatus* genome and its venom-duct transcriptome. BMC Genomics 12, 60.

Jin, A.H., Dutertre, S., Dutt, M., Lavergne, V., Jones, A., Lewis, R.J., Alewood, P.F., 2019. Transcriptomic-proteomic correlation in the predation-evoked venom of the cone snail, *Conus imperialis*. Marine Drugs 17, 177.

Kaas, Q., Westermann, J.C., Craik, D.J., 2010. Conopeptide characterization and classifications: an analysis using ConoServer. Toxicon 55, 1491-1509.

Kaas, Q., Yu, R., Jin, A.H., Dutertre, S., Craik, D.J., 2012. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. Nucleic Acids Research 40, D325-330.

Katoh, K., Rozewicki, J., Yamada, K.D., 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics bbx108.

Kohn, A.J., 1959. The ecology of *Conus* in Hawaii. Ecological Monographs 29, 47-90.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9, 357-359.

Lavergne, V., Dutertre, S., Jin, A.-H., Lewis, R.J., Taft, R.J., Alewood, P.F., 2013. Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. BMC Genomics 14, 708.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., Kendziorski, C., 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics 29, 1035-1043.

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

Li, Q., Barghi, N., Lu, A., Fedosov, A., E., Bandyopadhyay, P.K., Lluisma, A., O., Concepcion, G.P., Yandell, M., Olivera, B.M., Safavi-Hemami, H., 2017. Divergence of the venom exogene repertoire in two sister species of *Turriconus.* Genome Biology and Evolution 9, 2211-2225.

Lomonte, B., Tsai, W.C., Urena-Diaz, J.M., Sanz, L., Mora-Obando, D., Sanchez, E.E., Fry, B.G., Gutierrez, J.M., Gibbs, H.L., Sovic, M.G., Calvete, J.J., 2014. Venomics of New World pit vipers: genus-wide comparisons of venom proteomes across *Agkistrodon*. Journal of Proteomics 96, 103-116.

Lopez-Vera, E., Jacobsen, R.B., Ellison, M., Olivera, B.M., Teichert, R.W., 2007. A novel alpha conotoxin (alpha-PIB) isolated from *C. purpurascens* is selective for skeletal muscle nicotinic acetylcholine receptors. Toxicon 49, 1193-1199.

Lu, A., Yang, L., Xu, S., Wang, C., 2014. Various conotoxin diversifications revealed by a venomic study of *Conus flavidus*. Molecular Cell Proteomics 13, 105-118.

Maddison, W.P., Maddison, D.R., 2018. Mesquite: a modular system for evolutionary analysis. Version 3.51.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10-12.

Milne, T.J., Abbenante, G., Tyndall, J.D., Halliday, J., Lewis, R.J., 2003. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. Journal of Biological Chemistry 278, 31105-31110.

Olivera, B.M., 2006. Conus peptides: biodiversity-based discovery and exogenomics. Journal of Biological Chemistry 281, 31173-31177.

Olivera, B.M., McIntosh, J.M., Clark, C., Middlemas, D., Gray, W.R., Cruz, L.J., 1985. A sleep-inducing peptide from *Conus geographus* venom. Toxicon 23, 277 - 282.

Olivera, B.M., Rivier, J., Clark, C., Ramilo, C.A., Corpuz, G.P., Abogadie, F.C., Mena, E.E., Woodward, S.R., Hillyard, D.R., Cruz, L.J., 1990. Diversity of *Conus* neuropeptides. Science 249, 257 - 263.

Olivera, B.M., Seger, J., Horvath, M.P., Fedosov, A.E., 2015. Prey-capture strategies of fish-hunting cone snails: behavior, neurobiology and evolution. Brain Behavior and Evolution 86, 58-74.

Pahari, S., Bickford, D., Fry, B.G., Kini, R.M., 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 Evolutionary Biology 7, 175.

Pekar, S., Bocanek, O., Michalek, O., Petrakova, L., Haddad, C.R., Sedo, O., Zdrahal, Z., 2018. Venom gland size and venom complexity-essential trophic adaptations of venomous predators: A case study using spiders. Molecular Ecology 27, 4257-4269.

Peng, C., Liu, L., Shao, X., Chi, C., Wang, C., 2008. Identification of a novel class of conotoxins defined as V-conotoxins with a unique cysteine pattern and signal peptide sequence. Peptides 29, 985-991.

Peng, C., Yao, G., Gao, B.-M., Fan, C.-X., Bian, C., Wang, J., Cao, y., Wen, B., zhu, Y., Ruan, Z., Zhao, X., You, X., Bai, J., Li, J., Lin, Z., Zou, S., Zhang, X., Qiu, Y., Chen, J., Coon, S.L., Yang, J., Chen, J.-

S., Shi, Q., 2016. High-throughput identification of novel conotoxins from the Chinese tubular cone snail (*Conus betulinus*) by multitranscriptome sequencing. GigaScience 5, 17.
Phuong, M.A., Alfaro, M.E., Mahardika, G.N., Marwoto, R.M., Prabowo, R.E., von Rintelen, T., Wh Vogt, P., Hendricks, J.R., Puillandre, N., 2019. Lack of signal for the impact of conotoxin gene diversity on speciation rates in cone snails. Systematic Biology syz016.
Phuong, M.A., Mahardika, G.N., 2018. Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. Molecular Biology and Evolution 35, 1210-1224.
Phuong, M.A., Mahardika, G.N., Alfaro, M.E., 2016. Dietary breadth is positively correlated with venom complexity in cone snails. BMC Genomics 17, 401.
Pin, M., Leung Tack, K.D., 1995. The Cones of Senegal. Conchiglia.
Prashanth, J.R., Dutertre, S., Jin, A.H., Lavergne, V., Hamilton, B., Cardoso, F.C., Griffin, J., Venter, D.J., Alewood, P.F., Lewis, R.J., 2016. The role of defensive ecological interactions in the evolution of conotoxins. Molecular Ecology 25, 598-615.
Prashanth, J.R., Dutertre, S., Lewis, R.J., 2017. Pharmacology of predatory and defensive venom peptides in cone snails. Molecular Biosystems 13, 2453-2465.
Prator, C.A., Murayama, K.M., Schulz, J., R., 2014. Venom variation during prey capture by the cone snail, *Conus textile*. PLoS One 9, e98991.
Ranasinghe, S., McManus, D.P., 2013. Structure and function of invertebrate Kunitz serine protease inhibitors. Developmental and Comparative Immunology 39, 219-227.
R Core Team, 2013. R: A language and environment for statistical computing. In: computing, R.F.f.S. (Ed.), Vienna, Austria.
Remigio, E.A., Duda, T.F., Jr., 2008. Evolution of ecological specialization and venom of a predatory marine gastropod. Molecular Ecology 17, 1156-1162.
Robinson, S.D., Li, Q., Lu, A., Bandyopadhyay, P.K., Yandell, M., Olivera, B.M., Safavi-Hemami, H., 2017. The venom repertoire of *Conus gloriamaris* (Chemnitz, 1777), the glory of the sea. Marine Drugs 15, 145.
Robinson, S.D., Norton, R.S., 2014. Conotoxin gene superfamilies. Marine Drugs 12, 6058-6101.
Robinson, S.D., Safavi-Hemami, H., McIntosh, L.D., Purcell, A.W., Norton, R., S., Papenfuss, A.P., 2014. Diversity of conotoxin gene superfamilies in the venomous snail, *Conus victoriae*. PLoS One 9, e87648.
Safavi-Hemami, H., Gajewiak, J., Karanth, S., Robinson, S.D., Ueberheide, B., Douglass, A.D., Schlegel, A., Imperial, J.S., Watkins, M., Bandyopadhyay, P.K., Yandell, M., Li, Q., Purcell, A.W., Norton, R.S., Ellgaard, L., Olivera, B.M., 2015. Specialized insulin is used for chemical warfare by fish-hunting cone snails. Proceedings of the National Academy of Sciences USA 112, 1743 - 1748.
Salzberg, S.L., 2019. Next-generation genome annotation: we still struggle to get it right. Genome Biology 20, 92.
Terrat, Y., Biass, D., Dutertre, S., Favreau, P., Remm, M., Stöcklin, R., Piquemal, D., Ducancel, F., 2012. High-resolution picture of a venom gland transcriptome: Case study with the marine snail *Conus consors*. Toxicon 59, 34-46.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673 - 4680.
Tucker, J.K., Tenorio, M.J., 2009. Systematic classification of recent and fossil conoidean gastropods, with keys to the genera of cone shells. ConchBooks.
Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MdM Publishing.
UniProt, C., 2015. UniProt: a hub for protein information. Nucleic Acids Research 43, D204-212.

4. – Discussion

4. – Discussion

Comparative methods applied to the study of the evolutionary history of organisms offer an invaluable opportunity for uncovering the processes that promote diversification and the mechanisms that allow particular ecological adaptations (Dunwell et al., 2017; Pease et al., 2016; Tsai et al., 2013). For example, comparison of the speciation patterns in three independent groups of gastropods showcased the wide importance of the tectonic activity in the Indo-West Pacific in triggering diversification processes, thus shaping nowadays extraordinary diversity in the region (Williams and Duda, 2008). When applied in a phylogenetic context, comparative methods offer a new level of information by discerning convergent events from clade-specific adaptations, among others (Dunn and Munro, 2016). Despite cone snails diversification and diet evolution have been studied for decades, this comparative approach has been only recently applied in detail (Phuong et al., 2019; Puillandre et al., 2014), and a more general perspective of the evolution of the family is largely wanted. Thus far, most of the studies published to date have been focused on those species from the Indo-Pacific (where we can find more than half the species currently described (Tucker and Tenorio, 2013), and very few molecular studies have drawn their attention towards the cones from West Africa (Cunha et al., 2005; Cunha et al., 2014; Cunha et al., 2008; Duda and Rolán, 2005).

West African cone snails constitute a sound model system for studying evolution and applying phylogenetic-based comparative methods. As presented here, most of the genera endemic to these waters form a monophyletic assemblage, with the exception of *Genuanoconus* and *Monteiroconus*. Also, as explained in Chapters II and III, the genus *Africonus* and the species of *Lautoconus* from Senegal and the Canary Islands diversified through radiation events, thus providing the appropriate framework for applying these methods. Finally, the species inhabiting here display different diet specializations, and therefore can be used to understand the mechanisms driving these adaptations. Since all the results have been properly discussed in their respective chapters, this section provides an overarching view, aimed at drawing general conclusions from this work.

4.1. - Biogeographic patterns underlying the diversification processes

In the past, different studies proposed some mechanisms driving the speciation and extinction processes in cone snails, such as the tectonic activity generating new habitats (Williams and Duda, 2008), or diet specialization, which would reduce competition among species (Stanley, 2008). In addition, ecological factors related to habitat preference also were identified, and the mean temperature of the sea surface could be a key factor determining the local diversity of cone snails (Cunha et al., 2014). Thus far, the results presented here suggest that allopatry could be playing an important role in the origin of cone species.

Speciation by allopatry has been long considered to play a minor role in the diversification of species in the sea, since the general lack of geographical barriers, the big size of the marine populations, and the high dispersal capacity of the species would favour the gene flow between populations (Palumbi, 1994). However, the degree of isolation of marine species depends not only on their own dispersal capacities, but also is determined by other factors, such as the marine currents or the suitability of the surrounding habitats, or even on habitat specialization (Williams and Reid, 2004). For example, although the cone species *Miliariconus miliaris* is found throughout the Indo-Pacific, its population inhabiting Easter Island is partially isolated from the others, and displays higher levels of molecular divergence. This population inhabits the most remote island in the world, which limits the capacity for the larvae of other populations of *M. miliaris* to reach it and maintain the gene flow. In fact, genetic studies suggest that if gene flow still persists is mainly due to the Eastern Island larvae contacting the other populations (Duda and Lee, 2009).

Understanding why species diverge, particularly during radiation events, requires deciphering underlying ecological processes but also considering a temporal perspective, as we need to know how the conditions were and changed by the time the cladogenetic events occurred. For allopatric speciation, the degree of isolation of the species varies along time and this can have different causes. For example, the tectonic (e.g. the closure of the Tethys Sea (Duda and Kohn, 2005) or the emergence of the Isthmus of Panama (Frederich et al., 2013; Lessios, 1981; Lessios et al., 1999) and climatic activities can connect and disconnect populations temporarily. In fact, the fluctuations in the sea level caused by climatic events have already been proposed as a major promoter of speciation by allopatry in intertidal species (Davis et al., 2016; Shen

et al., 2011). As suggested by the fossil record, the evolutionary history of cone snails has been influenced by this kind of events, leading to successive speciation and extinction events (Kohn, 1990), and our results suggest that they also could have had a great influence in shaping the current diversity of cone snails in West Africa.

Although the genera endemic to West Africa have different origin times, the main cladogenetic events in all of them happened approximately at the same time. Briefly, the origin of the genus *Africonus* was estimated to be around 23 mya. Within this genus, the divergences of the four main lineages were dated between 9.4 and 6.9 mya, at the same time as the divergence between the *Kalloconus* from the continent and the archipelago occurred, as well as the first cladogenetic event leading to one of the three clades of the *Lautoconus* species from Senegal. The other two clades of this genus, inhabiting Senegal and Canary Islands appeared about 3 mya, the same time as the species of *Kalloconus* in Cabo Verde and the two lineages described within the four main clades of *Africonus*. Finally, most of the diversification events generating extant species happened in the last 2 million years (during the Pleistocene glacial-interglacial cycles). At the level of the family Conidae, many of the genera appeared at the same time as *Africonus*, and the diversification events within each genus in the last 10 million years. Therefore, cladogenesis in the family Conidae could be generally correlated with cooling events (the Oligocene-Miocene transition, (Davis et al., 2016; Miller et al., 2005; Zachos et al., 1997) or periods of climatic instability (the Plio-Pleistocene glacial-interglacial events, (Raymo et al., 1998; Raymo et al., 1989; Woodruff, 2010). Eustatic sea-level changes have important consequences on the intertidal ecosystems, by isolating and re-connecting areas and populations, thus promoting diversification and colonization of new areas.

Another key factor worth considering is that the cone genera endemic to this region present non-planktotrophic larvae, which would facilitate the isolation of the populations. The species with lecitotrophic larvae are confined in a restricted distribution range, which makes them more sensitive to extinction, but also promotes speciation by isolation (Stanley, 2008). By comparing the distribution ranges of the genera of cone snails endemic to West Africa, we find evidence for this statement. The Cabo Verde archipelago is rather isolated and once the ancestor of the species of the genus *Africonus* arrived, movements from one island to another were mostly promoted by sea level drop connecting the islands during glacial periods. Sea level rise during the

interglacial periods promoted isolation and speciation. In fact, the species of the genus *Africonus* are restricted to single islands or even neighbouring bays within an island (Monteiro et al., 2004; Röckel et al., 1980). The ancestor of the species of *Kalloconus* inhabiting Cabo Verde arrived more recently, lost the planktotrophic larvae, and thus far, there has been time for only four species to diversify, being restricted to the islands (Boa Vista, Maio, and Sal) that are closer to the continent (Cunha et al., 2008). The radiation of *Lautoconus* in the continental coasts of Senegal also follows the same pattern triggered by the poor dispersal capacity of the non-planktotrophic larvae. Similarly, *Lautoconus guanche* has colonized all the Canary Islands from the relatively close coast of Africa without having time to diversify as gene flow persists (Cunha et al., 2014). However, the genera *Chelyconus*, *Monteiroconus*, *Genuanoconus,* and *Kalloconus* from the continent have planktotrophic larvae with high dispersal capacities that allow the few species within each of these genera to have wide range distributions including the continent and some of the archipelagos, and in the case of *Chelyconus ermineus*, an amphi-Atlantic distribution.

4.2. - Evaluating the shell morphology as character for species diagnosis in cones

Cone snail species have been traditionally identified and described based on morphological characters, mainly from the shell. However, these characters are prone to homoplasy. Therefore, the traditional morphological approach has lead to the incorrect delimitation of many species, later uncovered when applying molecular methods (Duda et al., 2008; Duda et al., 2009). Despite that West Africa presents one of the most rich species diversity described thus far (Tucker and Tenorio, 2013) and that the pace of species description of this region is far from slowing down (see Fig 1 in Chapter III), very few studies have applied a molecular approach to assess the correct definition of these species (Cunha et al., 2005; Cunha et al., 2014; Duda and Rolán, 2005). Thanks to the sampling and sequencing effort done during this work, we could propose a new species delimitation hypothesis based on molecular data.

Both in Cabo Verde and Senegal we found the same patterns, with species presenting high morphological variability among populations and many instances of convergent phenotypes. Particularly, in Cabo Verde we re-defined up to 15 species whose (now) morphological variants had been described each as a different species. In an extreme case, we found that *Africonus irregularis* represented a cryptic complex,

whose populations were synonymized to *Africonus crotchii* and *Africonus maioensis*. Similarly, taxonomic changes were necessarily implemented in the genus *Kalloconus* (its species in Cabo Verde were formerly assigned to the genus *Trovaoconus*, but their sequence divergences with respect to *Kalloconus* species from the continent did not justify the validity of *Trovaoconus*) and in the endemic species of Senegal (individuals of *Lautoconus mercator* and *Lautoconus cacao* were present in two of the three main lineages of the genus making both species polyphyletic). Our results also suggested that the genus *Lautoconus* is paraphyletic as currently described and that *Lautoconus ventricosus* from the Mediterranean Sea, the only accepted cone species in the Mediterranean Sea might represent a cryptic complex. Both problems should be studied in more detail with an extended taxon sampling. The taxonomy of diverse group is highly dynamic and molecular analyses will prompt not only many synonymizations, but also there is still room for the description of new species (e.g. (Tenorio et al., 2018; Tenorio et al., 2017).

The results here presented showcase how the use of misleading diagnostic characters could both infra- or overestimate the number of species. More importantly, they highlight the need of defining new morphological diagnostic characters (if available) that would make the delimitation between species more clear, and, in any case, include the use of molecular sequence data (e.g., *cox1* barcoding) as an obligate step for species description. Cones are highly appreciated by shell collectors and the temptation to describe species based on bold shell phenotypes is high by amateurs. Therefore, it is crucial that the results obtained through molecular phylogenetics and proper taxonomic analyses have impact on such community.

Uncovering the real diversity of a region, without inflation or underestimation of the number of species, is of paramount importance, since these deviations could have important consequences in downstream analyses such as e.g. the definition of the status of conservation of the species. The IUCN Red List assesses different parameters, including external factors (human pressure) and those related to the distribution and abundance of the species (such as the size of the populations) to evaluate risk of extinction. After evaluating more than 600 cone species worldwide, all the ones that were categorized as Critically Endangered (CR) or Endangered (EN) happen to occur in Cabo Verde and Senegal (Peters et al., 2013). However, our results indicate that several of the cone species in Cabo Verde in fact are populational variations of the same

species, hence with a wider distribution range, higher number of individuals, and likely more optimistic conservation status. In contrast, the species *L. ventricosus* from the Mediterranean Sea is a Least Concern (LC), but this category could change for some populations if the presence of a cryptic complex of species is confirmed.

Taxonomic inaccuracies also have a great impact when studying the species from an evolutionary point of view. For example, upon the analysis of the rates of speciation in the family Conidae, Phuong et al. (2019) found two contrasting results regarding West African cone snails: when including all the species described morphologically the authors detected an acceleration of the speciation rate, whereas this rate shift was not observed when considering the molecular results here presented. Statistical models like BiSSE or HiSSE were used to infer the influence of particular traits in the rates of speciation and extinction (Beaulieu and O'Meara, 2016; Maddison et al., 2007). If our results are correct and represent a realistic delimitation of the species inhabiting West Africa, the rate acceleration found in the branch leading to this clade could be a false positive, thus overestimating the influence of some of the analysed traits in the diversification of cone snails.

4.3. - Suitability of the used molecular markers for the study of the family Conidae

The need of having a robust phylogenetic framework to tackle long-standing evolutionary questions (such as how and when the diet shifts happened, or what are the mechanisms underlying cones diversification, among others) has been one of the main drivers of this study. Cones evolved through radiation events (Berschauer, 2015; Duda and Rolán, 2005; Kohn, 1990; Pin and Leung Tack, 1995) and thus, using only a few partial gene sequences was not an option for dealing with the expected short nodes of the tree. Therefore, we used complete mitochondrial genomes and nuclear data to reconstruct a robust phylogeny for the family and the endemic genera to West Africa. Here, I will evaluate their performance and suitability for future studies in the family Conidae.

Mitochondrial genomes have proven to be useful to resolve the phylogenies of numerous animal groups at different levels (Galaska et al., 2019; Li et al., 2019; Uribe et al., 2019; Uribe et al., 2017b). They provide particularly good resolution up to the family level and with the use of sophisticated evolutionary models they could be useful

even at higher taxonomic levels. In parallel to this work, mitogenomes were applied to the study of cone snails at the family and superfamily levels with satisfactory results (Uribe et al., 2017a; Uribe et al., 2018). In our case, mitogenomes produced rather resolved trees of the family Conidae, which were further improved with the addition of nuclear sequence data. Yet, some nodes remained recalcitrantly unresolved. Several recent studies applying an exon-capture approach and sequencing hundreds of exons in an extended taxon sampling rendered similar levels of unresolved internal nodes (Phuong et al., 2019; Phuong and Mahardika, 2018), which may suggest that the problem is not the phylogenetic signal present in our data but the radiation processes themselves.

In any case, the mitogenomic approach requires extending the taxon sampling. Incorporating new data has proven to be useful for resolving recalcitrant nodes when overriding long branch attraction artefacts, for instance, in the node connecting Patellogastropoda to other gastropod lineages (Uribe et al., 2019). In our study of the family Conidae, we only included one third of the total diversity of genera, and lacked some key taxa such as *Fraterconus distans*, which was recovered as the first divergent lineage of the family in other studies (Phuong and Mahardika, 2018; Puillandre et al., 2014). Future mitogenomic studies should ambition including as many species as in Phuong et al. (2019), who analysed most of the genera and more than 300 species.

4.4. - Comparative transcriptomics applied to the study of the venom of cone snails

As explained in Chapters IV and V, the application of transcriptomics to venomic studies has produced a quantum leap in the characterization of venom composition. Particularly in cone snails, it meant going from describing a few tens to work now with hundreds of different peptides (e.g. (Barghi et al., 2015; Pi et al., 2006a; Pi et al., 2006b).

The critical step in the transcriptomic approach is the correct identification of conotoxins. Because of the big amount of sequencing data, venomics relies on more or less complex pipelines that automate the identification of those peptides that potentially could be part of the venom (Koua et al., 2012; Lavergne et al., 2013; Li et al., 2018). However, sequencing errors and assembly artefacts are not uncommon in NGS datasets, as well as missannotations. These problems cannot be fully detected and discarded

using a completely automated pipeline. Including these errors, artefacts, and missannotations in the final report could bias the interpretation of the data, and hinder the comparison between studies. Moreover, one main concern should be the propagation of these errors when included in public databases and used in reference data sets (Salzberg, 2019). Hence, the need of implementing some manual curation step as suggested here.

In other fields, comparative transcriptomics has helped disentangle the mechanisms underlying ecological adaptations in numerous organisms. By comparing the expression profiles among species or between conditions, we can identify those genes that are participating differentially in response to a particular stimulus (Bernal et al., 2019; Heras and Aguilar, 2019; Rahi et al., 2019; Young et al., 2019). Hitherto, however, venom-related ecological adaptations in cone snails have been studied mainly from a proteomic point of view (e.g. (Dutertre et al., 2014; Prashanth et al., 2016), and transcriptomes have been mostly relegated to descriptive analyses of conotoxin diversity (Hu et al., 2011; Robinson et al., 2017; Yao et al., 2019). While proteomic analyses allow identifying particular conotoxins that are being secreted under a particular stimulus, that correlation is not straightforward when analysing transcriptomic data, not only because of the big diversity of conotoxins expressed, but because these transcripts are not necessarily translated and included in the final secreted venom (Jin et al., 2019).

Despite its yet scarce utilization in cone snails, the comparison of venom expression profiles between species can help us understanding how the prey capture system is evolving. Phuong et al. (2016) compared the venom composition of 16 species of cone snails and inferred the influence of dietary breadth in venom complexity: the wider was the diet the wider was the diversity of precursors in the venom. However, even that transcriptomic data for cone snails is steadily accumulating, one of the main limitations for this comparative approach is that results reported in different species are not fully equivalent. For instance, the progressive addition of sequences into the databases makes reference data sets more thorough and increases the identification of conotoxins in any new transcriptome. Therefore, a direct comparison of the absence or presence of toxin superfamilies between species whose transcriptomes were analysed at different times becomes unfair. Nonetheless, the results reported here demonstrate that most recent venom gland transcriptomes use common methodological

(Illumina NGS) and analytical (Trinity software) approaches, and that comparing transcriptomic data among species can render useful results to better understand diet specialization in cone snails.

The Chapter IV, which focused on the evolution of piscivory in the species *C. ermineus*, shows the potential of the approach. As inferred through phylogenetic analyses, the piscivory of the genus *Chelyconus* is an adaptation independent from that of the Indo-Pacific species. By comparing the transcriptomes of *C. ermineus* with other piscivorous species (Hu et al., 2012; Terrat et al., 2012), we highlighted the importance of the A superfamily for hunting fish, thus confirming previous hypothesis. Interestingly, although there were published proteomes for both piscivorous clades (Barbier et al., 2004; Endean and Izatt, 1965; Hoggard et al., 2017; Olivera et al., 1985; Quinton et al., 2013), it was not until the comparison of the transcriptomic data (and after having a full profile of the venom of these species) that we could observe the convergent evolution shaping the composition of their venoms. Moreover, the relative importance of the S superfamily for piscivory had gone unnoticed in our results until the comparison to the vermivorous species.

The evolutionary framework established in Chapters II and III allowed us to analyse venom evolution within a phylogenetic context of closely related species, by sequencing the transcriptomes of the venom glands of species representing all the lineages of the genera *Africonus* and *Lautoconus*. We detected a clear phylogenetic signal in the data, as those species more closely related shared more common conotoxins (allowing ancestral reconstruction analyses) and had more similar venom composition and expression profiles. In this regard, such phylogenetic signal could also be detected based on the relative expansion of toxin superfamily genes in closely related species using an exon-capture approach at the genomic level (Phuong et al., 2019). Nevertheless, the phylogenetic signal drops significantly above the genus level as few precursors are shared by species from distantly related genera. In any case, venom composition comparisons also shoed instances of functional convergence both between closely (*Africonsu raulsilvai* and *Africonus maiosensis*) and more distantly related species, (Duda and Remigio, 2008). Despite that a recent study (Phuong et al., 2019) found no correlation between venom diversity and speciation rates, we did detect that *Africonus* presented more expanded superfamilies than *Lautoconus*, being the former more species rich than the latter.

In summary, most of the studies in cone snails had been focused thus far on the Indo-Pacific region, neglecting other interesting areas like the Caribbean Sea and West Africa. The results presented here fill in part this gap and provide a robust framework for the study of the evolution of cone snails. Although the rapid radiation events in the group hinder the phylogenetic inferences, the mitochondrial genomes have been proven to perform well for resolving the intra-generic relationships, and even recover most of the nodes of the tree with a good statistical support. Moreover, we studied for the first time the conotoxin repertoire of these species, which can now be included in future studies of the venom of this family. These results highlight the importance of comparative studies on venom composition to understand ecological adaptations, in particular diet specializations.

4.5. – References

Barbier, J., Lamthanh, H., Le Gall, F., Favreau, P., Benoit, E., Chen, H., Gilles, N., Ilan, N., Heinemann, S.H., Gordon, D., Ménez, A., Molgó, J., 2004. A δ-Conotoxin from *Conus ermineus* Venom Inhibits Inactivation in Vertebrate Neuronal Na+Channels but Not in Skeletal and Cardiac Muscles. Journal of Biological Chemistry 279, 4680-4685.

Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2015. High conopeptide diversity in *Conus tribblei* revealed through analysis of venom duct transcriptome using two High-Throughput sequencing platforms. Marine Biotechnology 17, 81-98.

Beaulieu, J.M., O'Meara, B.C., 2016. Detecting Hidden Diversification Shifts in Models of Trait-Dependent Speciation and Extinction. Systematic Biology 65, 583-601.

Bernal, M.A., Dixon, G.B., Matz, M.V., Rocha, L.A., 2019. Comparative transcriptomics of sympatric species of coral reef fishes (genus: *Haemulon*). PeerJ 7, e6541.

Berschauer, D.P., 2015. A comparison of adaptive radiation in Conidae and Conilithidae (Gastropoda: Conoidea) in the Eastern and Western Atlantic, together with an iconography of the conilithid genus Jaspidiconus. The Festivus 47, 99-113.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde islands. Systematic Biology 54, 634-650.

Cunha, R.L., Lima, F.P., Tenorio, M.J., Ramos, A.A., Castilho, R., Williams, S.T., 2014. Evolution at a Different Pace: Distinctive Phylogenetic Patterns of Cone Snails from Two Ancient Oceanic Archipelagos. Systematic Biology 63, 971-987.

Cunha, R.L., Tenorio, M.J., Afonso, C., Castilho, R., Zardoya, R., 2008. Replaying the tape: recurring biogeographical patterns in Cape Verde *Conus* after 12 million years. Molecular Ecology 17, 885-901.

Davis, K.E., Hill, J., Astrop, T.I., Wills, M.A., 2016. Global cooling as a driver of diversification in a major marine clade. Nature Communications 7, 13003.

Duda, T.F., Bolin, M.B., Meyer, C.P., Kohn, A.J., 2008. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. Molecular Phylogenetics and Evolution 49, 867-876.

Duda, T.F., Jr., Remigio, E.A., 2008. Variation and evolution of toxin gene expression patterns of six closely related venomous marine snails. Molecular Ecology 17, 3018-3032.

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., Kohn, A.J., Matheny, A.M., 2009. Cryptic species differentiated in *Conus ebraeus*, a widespread tropical marine gastropod. Biological Bulletin 217, 292-305.

Duda, T.F., Lee, T., 2009. Isolation and population divergence of a widespread Indo-West Pacific marine gastropod at Easter Island. Marine Biology 156, 1193-1202.

Duda, T.F., Rolán, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Molecular Ecology 14, 267-272.

Dunn, C.W., Munro, C., 2016. Comparative genomics and the diversity of life. Zoologica Scripta 45, 5-13.

Dunwell, T.L., Paps, J., Holland, P.W.H., 2017. Novel and divergent genes in the evolution of placental mammals. Proceedings of the Royal Society B: Biological Sciences 284.

Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nature Communiactions 5, 3521.

Endean, R., Izatt, J., 1965. Pharmacological study of the venom of the gastropod *Conus magus*. Toxicon 3, 81-93.

Frederich, B., Sorenson, L., Santini, F., Slater, G.J., Alfaro, M.E., 2013. Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). American Naturalist 181, 94-113.

Galaska, M.P., Li, Y., Kocot, K.M., Mahon, A.R., Halanych, K.M., 2019. Conservation of mitochondrial genome arrangements in brittle stars (Echinodermata, Ophiuroidea). Molecular Phylogenetics and Evolution 130, 115-120.

Heras, J., Aguilar, A., 2019. Comparative Transcriptomics Reveals Patterns of Adaptive Evolution Associated with Depth and Age Within Marine Rockfishes (Sebastes). Journal of Heredity 110, 340-350.

Hoggard, M.F., Rodriguez, A.M., Cano, H., Clark, E., Tae, H.S., Adams, D.J., Godenschwege, T.A., Mari, F., 2017. *In vivo* and *in vitro* testing of native alpha-conotoxins from the injected venom of *Conus purpurascens*. Neuropharmacology 127, 253-259.

Hu, H., Bandyopadhyay, P.K., Olivera, B.M., Yandell, M., 2011. Characterization of the *Conus bullatus* genome and its venom-duct transcriptome. BMC Genomics 12, 60.

Hu, H., Bandyopadhyay, P.K., Olivera, B.M., Yandell, M., 2012. Elucidation of the molecular envenomation strategy of the cone snail *Conus geographus* through transcriptome sequencing of its venom duct. BMC Genomics 13, 284.

IUCN RedList, 2010. The IUCN red list of threatened species. 10.

Jin, A.H., Dutertre, S., Dutt, M., Lavergne, V., Jones, A., Lewis, R.J., Alewood, P.F., 2019. Transcriptomic-Proteomic Correlation in the Predation-Evoked Venom of the Cone Snail, *Conus imperialis*. Marine Drugs 17.

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

Koua, D., Brauer, A., Laht, S., Kaplinski, L., Favreau, P., Remm, M., Lisacek, F., Stocklin, R., 2012. ConoDictor: a tool for prediction of conopeptide superfamilies. Nucleic Acids Research 40, W238-241.

Lavergne, V., Dutertre, S., Jin, A.-H., Lewis, R.J., Taft, R.J., Alewood, P.F., 2013. Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. BMC Genomics 14, 708.

Lessios, H.A., 1981. Divergence in allopatry: Molecular and morphological differenciation between sea urchins separated by the Isthmus of Panama. Evolution 35, 618 - 634.

Lessios, H.A., Kessing, B.D., Robertson, D.R., Paulay, G., 1999. Phylogeography of the pantropical sea urchin *Eucidaris* in relation to land barriers and ocean currents. Evolution 53, 806 - 817.

Li, Q., Watkins, M., Robinson, S.D., Safavi-Hemami, H., Yandell, M., 2018. Discovery of Novel Conotoxin Candidates Using Machine Learning. Toxins (Basel) 10.

Li, Y., Kocot, K.M., Tassia, M.G., Cannon, J.T., Bernt, M., Halanych, K.M., 2019. Mitogenomics Reveals a Novel Genetic Code in Hemichordata. Genome Biology and Evolution 11, 29-40.

Maddison, W.P., Midford, P.E., Otto, S.P., 2007. Estimating a binary character's effect on speciation and extinction. Systematic Biology 56, 701 - 710.

Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E., Sugarman, P.J., Cramer, B.S., Christie-Blick, N., Pekar, S.F., 2005. The Phanerozoic record of Global Sea-Level change. Science 350, 1293-1298.

Monteiro, A., Tenorio, M.J., Poppe, G.T., 2004. The family Conidae, the west African and Mediterranean species of *Conus*. ConchBooks, Hackenheim, Germany.

Olivera, B.M., McIntosh, J.M., Clark, C., Middlemas, D., Gray, W.R., Cruz, L.J., 1985. A sleep-inducing peptide from *Conus geographus* venom. Toxicon 23, 277 - 282.

Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. The Annual Review of Ecology, Evolution, and Systematics 25, 547 - 572.

Pease, J.B., Haak, D.C., Hahn, M.W., Moyle, L.C., 2016. Phylogenomics Reveals Three Sources of Adaptive Variation during a Rapid Radiation. PLoS Biology 14, e1002379.

Peters, H., O'Leary, B.C., Hawkins, J.P., Carpenter, K.E., Roberts, C.M., 2013. *Conus*: first comprehensive conservation red list assessment of a marine gastropod mollusc genus. PLoS One 8, e83353.

Phuong, M.A., Alfaro, M.E., Mahardika, G.N., Marwoto, R.M., Prabowo, R.E., von Rintelen, T., Wh Vogt, P., Hendricks, J.R., Puillandre, N., 2019. Lack of signal for the impact of conotoxin gene diversity on speciation rates in cone snails. Systematic Biology.

Phuong, M.A., Mahardika, G.N., 2018. Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. Molecular Biology and Evolution.

Phuong, M.A., Mahardika, G.N., Alfaro, M.E., 2016. Dietary breadth is positively correlated with venom complexity in cone snails. BMC Genomics 17, 401.

Pi, C., Liu, J., Peng, C., Liu, Y., Jiang, X., Zhao, Y., Tang, S., Wang, L., Dong, M., Chen, S., Xu, A., 2006a. Diversity and evolution of conotoxins based on gene expression profiling of *Conus litteratus*. Genomics 88, 809-819.

Pi, C., Liu, Y., Peng, C., Jiang, X., Liu, J., Xu, B., Yu, X., Yu, Y., Jiang, X., Wang, L., Dong, M., Chen, S., Xu, A.L., 2006b. Analysis of expressed sequence tags from the venom ducts of *Conus striatus*: focusing on the expression profile of conotoxins. Biochimie 88, 131-140.

Pin, M., Leung Tack, K.D., 1995. The Cones of Senegal. Conchiglia.

Prashanth, J.R., Dutertre, S., Jin, A.H., Lavergne, V., Hamilton, B., Cardoso, F.C., Griffin, J., Venter, D.J., Alewood, P.F., Lewis, R.J., 2016. The role of defensive ecological interactions in the evolution of conotoxins. Molecular Ecology 25, 598-615.

Puillandre, N., Bouchet, P., Duda, T.F., Jr., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Molecular Phylogenetics and Evolution 78, 290-303.

Quinton, L., Servent, D., Girard, E., Molgó, J., Le Caer, J.-P., Malosse, C., Haidar, E.A., Lecoq, A., Gilles, N., Chamot-Rooke, J., 2013. Identification and functional characterization of a novel a-conotoxin (EIIA) from *Conus ermineus*. Analytical and Bioanalytical Chemistry 405, 5341-5351.

Rahi, M.L., Mather, P.B., Ezaz, T., Hurwood, D.A., 2019. The Molecular Basis of Freshwater Adaptation in Prawns: Insights from Comparative Transcriptomics of Three Macrobrachium Species. Genome Biology and Evolution 11, 1002-1018.

Raymo, M.E., Ganley, K., Carter, S., Oppo, D.W., McManus, J., 1998. Millenial-scale climate instability during the early Pleistocene epoch. Nature 392, 699.

Raymo, M.E., Ruddiman, W.F., Backman, J., Clement, B.M., Martinson, D.G., 1989. Late pliocene variation in Northern Hemisphere ice sheets and North Atlantic deep water circulation. Paleoceanography 4, 413 - 446.

Robinson, S.D., Li, Q., Lu, A., Bandyopadhyay, P.K., Yandell, M., Olivera, B.M., Safavi-Hemami, H., 2017. The Venom Repertoire of *Conus gloriamaris* (Chemnitz, 1777), the Glory of the Sea. Marine Drugs 15, 145.

Röckel, D., Rolan, E., Monteiro, A., 1980. Cone shells from Cape Verde islands. A difficult puzzle. I. Feito, Vigo, Spain.

Salzberg, S.L., 2019. Next-generation genome annotation: we still struggle to get it right. Genome Biology 20, 92.

Shen, K.-N., Jamandre, B.W., Hsu, C.-C., Tzeng, W.-N., Durand, J.-D., 2011. Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet Mugil cephalus. BMC Evolutionary Biology 11, 83.

Stanley, S.M., 2008. Predation defeats competition on the seafloor. Paleobiology 34, 1-21.

Tenorio, M.J., Abalde, S., Zardoya, R., 2018. Identification of new species of *Kalloconus* and *Africonus* (Gastropoda, Conidae) from the Cabo Verde Islands through mitochondrial genome comparison. The Festivus 50, 73-88.

Tenorio, M.J., Afonso, C.M.L., Rolán, E., Pires, S., Vasconcelos, P., Abalde, S., Zardoya, R., 2017. DNA sequences disclose a nwe species of *Africonus* cone snail from Sao Vicente (Gastropoda: Conidae). Zoologia Caboverdiana 6, 32-41.

Terrat, Y., Biass, D., Dutertre, S., Favreau, P., Remm, M., Stöcklin, R., Piquemal, D., Ducancel, F., 2012. High-resolution picture of a venom gland transcriptome: Case study with the marine snail *Conus consors*. Toxicon 59, 34-46.

Tsai, I.J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K.L., Tracey, A., Bobes, R.J., Fragoso, G., Sciutto, E., Aslett, M., Beasley, H., Bennett, H.M., Cai, X., Camicia, F., Clark, R., Cucher, M., De Silva, N., Day, T.A., Deplazes, P., Estrada, K., Fernandez, C., Holland, P.W.H., Hou, J., Hu, S., Huckvale, T., Hung, S.S., Kamenetzky, L., Keane, J.A., Kiss, F., Koziol, U., Lambert, O., Liu, K., Luo, X., Luo, Y., Macchiaroli, N., Nichol, S., Paps, J.,

Parkinson, J., Pouchkina-Stantcheva, N., Riddiford, N., Rosenzvit, M., Salinas, G., Wasmuth, J.D., Zamanian, M., Zheng, Y., Taenia solium Genome, C., Cai, J., Soberon, X., Olson, P.D., Laclette, J.P., Brehm, K., Berriman, M., 2013. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496, 57-63.

Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MdM Publishing.

Uribe, J.E., Irisarri, I., Templado, J., Zardoya, R., 2019. New patellogastropod mitogenomes help counteracting long-branch attraction in the deep phylogeny of gastropod mollusks. Molecular Phylogenetics and Evolution 133, 12-23.

Uribe, J.E., Puillandre, N., Zardoya, R., 2017a. Beyond *Conus*: Phylogenetic relationships of Conidae based on complete mitochondrial genomes. Molecular Phylogenetics and Evolution 107, 142-151.

Uribe, J.E., Williams, S.T., Templado, J., Abalde, S., Zardoya, R., 2017b. Denser mitogenomic sampling improves resolution of the phylogeny of the superfamily Trochoidea (Gastropoda: Vetigastropoda). Journal of Molluscan Studies 83, 111-118.

Uribe, J.E., Zardoya, R., Puillandre, N., 2018. Phylogenetic relationships of the conoidean snails (Gastropoda: Caenogastropoda) based on mitochondrial genomes. Molecular Phylogenetics and Evolution 127, 898-906.

Williams, S.T., Duda, T.F., Jr., 2008. Did tectonic activity stimulate oligo-miocene speciation in the Indo-West Pacific? Evolution 62, 1618-1634.

Williams, S.T., Reid, D.G., 2004. Speciation and diversity on tropical rocky shored: a global phylogeny of snails of the genus *Echinolittorina*. Evolution 58, 2227 - 2251.

Woodruff, D.S., 2010. Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. Biodiversity and Conservation 19, 919-941.

Yao, G., Peng, C., Zhu, Y., Fan, C., Jiang, H., Chen, J., Cao, Y., Shi, Q., 2019. High-Throughput Identification and Analysis of Novel Conotoxins from Three Vermivorous Cone Snails by Transcriptome Sequencing. Marine Drugs 17.

Young, R.L., Ferkin, M.H., Ockendon-Powell, N.F., Orr, V.N., Phelps, S.M., Pogány, Á., Richards-Zawacki, C.L., Summers, K., Székely, T., Trainor, B.C., Urrutia, A.O., Zachar, G., O'Connell, L.A., Hofman, H.A., 2019. Conserved transcriptomic profiles underpin monogamy across vertebrates. Proceedings of the National Academy of Sciences 116, 1331 - 1336.

Zachos, J.C., Flower, B.P., Paul, H., 1997. Orbitally paced climate oscillations across the Oligocene/Miocene boundary. Nature 388, 567.

5. – Conclusions

5. – Conclusions

From the results presented in this PhD thesis concerning the phylogeny and comparative transcriptomics of West African cones, the following main conclusions can be drawn:

1. The mitochondrial genomes of 114 specimens of cones from the West African region were sequenced using massive sequencing techniques: 76 *Africonus* species, 19 *Lautoconus* species, 16 *Kalloconus* species, *Chelyconus ermineus*, and *Genuanoconus genuanus*. All the mitogenomes presented the consensus gene order reported for the family Conidae.

2. The transcriptomic sequence data from the venom gland of cones is an invaluable source for annotating mitochondrial and nuclear genes for their use in phylogenetic inference. The pace of publication of these transcriptomes increases the species and genera representation of cones in the phylogenetic analyses.

3. Mitogenomes performed well for resolving the phylogenetic relationships between species and for closely related genera. However, their phylogenetic signal failed to resolve some of the deepest nodes of the tree of the family Conidae.

4. The ancestor of the family Conidae was vermivorous. The diet shifts to molluscivory and to prey on amphinomids happened as a single evolutionary event each. However, the piscivory of the Indo-Pacific and Atlantic cones was likely originated in two independent events.

5. The phylogenetic analyses of the cone species endemic to Senegal (plus *L. guanche* from Canary Islands) recovered three main clades, each one presenting a characteristic radular type, which may indicate different diet specializations.

6. The genus *Africonus* is endemic to the Cabo Verde archipelago and presents four main clades, subdivided each into two lineages. The species from Sal, Santiago, and the westernmost islands (Santa Luzia, São Vicente and Santo Antão) form monophyletic assemblages. The closest sister group to *Africonus* is *L. ventricosus*

and not the one formed by continental species from Senegal. The origin of this genus was dated 23 mya.

7. The genus *Trovaoconus* should be synonymized to *Kalloconus,* as the species in Cabo Verde have not accumulated enough sequence divergence compared to the continental ones. The diversification of *Kalloconus* species in Cabo Verde occurred nine mya.

8. The main cladogenetic events in the family Conidae and the West African genera are coincident with periods of climatic instability, leading to eustatic sea level fluctuations, and fostering speciation by allopatry.

9. Instances of incorrect species description based on shell morphology were found in the three genera: *Africonus*, *Kalloconus* and *Lautoconus*. In the case of *Africonus*, the total number of valid species should be reduced to a half.

10. The transcriptomic analyses confirmed the high intra and interspecific variability in venom composition, as well as the venom duct regionalization in the production of conotoxins of the Atlantic pisicivore cone, *Chelyconus ermineus*. The M, O1, O2 and T conotoxin precursor superfamilies were among the most abundant in the venom of this species.

11. The A superfamily presented the highest levels of expression in the venom duct of *C. ermineus*, which confirms the importance of this superfamily for piscivory. The conotoxins in this species that target the muscle nicotinic-acetylcholine receptors of the prey present a α4/4 cysteine spacing pattern (in contrast with the α3/5 of the Indo-Pacific piscivore cone species), suggesting a convergent origin for this diet.

12. The comparative transcriptome analysis of 12 species from two closely related West African genera (*Africonus* and *Lautoconus*) detected that phylogenetic signal in venom composition is retained between closely related species, which tend to

have similar presence/absence, abundance and expression profiles of conotoxin precursor superfamilies. The venom of the *Africonus* species tend to have more expanded conotoxin precursor superfamilies, which could be related to the greater species diversification rates in the archipelago.

13. The differential expression analyses between vermivorous and piscivorous species detected the overexpression of the A and S superfamilies in the venom of *C. ermineus*, confirming the importance of these superfamilies to prey on fish.

5. – Conclusiones

De los resultados presentados de esta tesis doctoral sobre la filogenia y la transcriptómica comparada de los conos de África Occidental, se pueden extraer las siguientes conclusiones principales:

1. Los genomas mitocondriales de 114 especímenes de conos de la región de África Occidental se secuenciaron mediante técnicas de secuenciación masiva: 76 especies de *Africonus*, 19 especies de *Lautoconus*, 16 especies de *Kalloconus*, *Chelyconus ermineus* y *Genuanoconus genuanus*. Todos los mitogenomas presentaron el orden génico de consenso de la familia Conidae.

2. Los datos de secuencia transcriptómica de la glándula venenosa de conos son una fuente valiosa para anotar los genes mitocondriales y nucleares de cara a su uso en la inferencia filogenética. El ritmo de publicación de estos transcriptomas aumenta la representación de especies y géneros de conos en los análisis filogenéticos.

3. Los mitogenomas funcionaron bien en la resolución de las relaciones filogenéticas entre especies y para géneros estrechamente relacionados. Sin embargo, su señal filogenética no logró resolver algunos de los nodos más profundos del árbol de la familia Conidae.

4. El antepasado de la familia Conidae era vermívoro. Los cambios de la dieta a moluscívoro y comer anfinómidos fueron un único evento evolutivo cada uno. En cambio, la piscivoria de los conos del Indo-Pacífico y del Atlántico probablemente se originó en dos eventos independientes.

5. Los análisis filogenéticos de las especies de conos endémicos de Senegal (más *L. guanche* de las Islas Canarias) recuperaron tres clados principales, cada uno con un tipo radular característico, lo cual puede indicar diferentes especializaciones en la dieta.

6. El género *Africonus* es endémico del archipiélago de Cabo Verde y presenta cuatro clados principales, subdivididos cada uno en dos linajes. Las especies de Sal, Santiago y las islas más occidentales (Santa Luzia, São Vicente y Santo Antão) forman grupos monofiléticos. El grupo hermano más cercano a *Africonus* es *L. ventricosus* y no el formado por las especies continentales de Senegal. El origen de este género ocurrió hace 23 millones de años.

7. El género *Trovaoconus* debe ser sinónimizado con *Kalloconus*, ya que las especies en Cabo Verde no han acumulado suficiente divergencia de secuencia en comparación con las continentales. La diversificación de especies de *Kalloconus* en Cabo Verde ocurrió hace nueve millones de años.

8. Los principales eventos cladogenéticos en la familia Conidae y en los géneros de África Occidental coinciden con los períodos de inestabilidad climática, lo que lleva a fluctuaciones eustáticas en el nivel del mar y fomenta la especiación en alopatría.

9. Los tres géneros, *Africonus*, *Kalloconus* y *Lautoconus* presentaron casos de descripciones incorrectas de especies en base a la morfología de la concha . En el caso de *Africonus*, el número total de especies válidas debe reducirse a la mitad.

10. Los análisis transcriptómicos confirmaron la alta variabilidad intra e interespecífica en la composición del veneno, así como la regionalización del conducto del veneno en la producción de conotoxinas del cono pisicívoro del Atlántico, *Chelyconus ermineus*. Las superfamilias de precursores de conotoxinas M, O1, O2 y T estaban entre las más abundantes en el veneno de esta especie.

11. La superfamilia A presentó los niveles más altos de expresión en el conducto del veneno de *C. ermineus*, lo que confirma la importancia de esta superfamilia para la piscivoría. Las conotoxinas en esta especie que interaccionan con los receptores de acetilcolina de tipo nicotínico del músculo de la presa presentan un patrón de

espaciamiento de cisteína α4 / 4 (en contraste con el α3 / 5 de las especies piscívoras de conos del Indo-Pacífico), lo que sugiere un origen convergente para esta dieta.

12. El análisis comparativo del transcriptoma de 12 especies de dos géneros de África Occidental estrechamente relacionados (*Africonus* y *Lautoconus*) detectó que la señal filogenética en la composición del veneno se mantiene entre las especies estrechamente relacionadas, que tienden a tener perfiles similares de presencia / ausencia, abundancia y expresión de superfamilias de precursores de conotoxinas. El veneno de las especies de *Africonus* tiende a tener superfamilias de precursores de conotoxinas más expandidas, lo que podría estar relacionado con las mayores tasas de diversificación de especies en el archipiélago.

13. Los análisis de expresión diferencial entre las especies vermívoras y piscívoras detectaron la sobreexpresión de las superfamilias A y S en el veneno de *C. ermineus*, lo que confirma la importancia de estas superfamilias para capturar a los peces.

6. – Appendix 1

6. – Appendix 1

All the Chapters here presented link to several supplementary material. However, the size of many of these files was too big to be included in this document. All they are accessible online. Therefor, and with the intention of making easier the access to these files, this appendix contains the hyperlinks to all the supplementary material included in this PhD.

Copy and paste these links into your browser.

Chapter I

Supp. Fig. 1:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0001-FigS1.pdf

Supp. Fig. 2:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0002-FigS2.pdf

Supp. Fig 3:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0003-FigS3.pdf

Supp. Table 1:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0004-TableS1.xlsx

Supp. Table 2:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0005-TableS2.xlsx

Supp. Table 3:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0006-TableS3.xlsx

Supp. Table 4:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0007-TableS4.xlsx

Supp. Table 5:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0007-TableS4.xlsx

Supp. Table 6:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fzsc.12329&file=zsc12329-sup-0009-TableS6.xlsx

Chapter II

Supp. Data 1:

https://ars.els-cdn.com/content/image/1-s2.0-S1055790316303700-mmc1.pdf

Supp. Data 2:

https://ars.els-cdn.com/content/image/1-s2.0-S1055790316303700-mmc2.nwk

Chapter III

Suppl. File 1:

https://static-content.springer.com/esm/art%3A10.1186%2Fs12862-017-1069-x/MediaObjects/12862_2017_1069_MOESM1_ESM.zip

Chapter IV

Supp. Data:

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/gbe/10/10/10.1093_gbe_evy150/3/evy150_supp.zip?Expires=1562152254&Signature=KlTwbZ8b0I4koOsknyQM82FzD-6zozvgLpvyBZuXx-NA-4RBNDu9yXACYMMD5EKqe7FsBMx0X8VN9rLwDEKfh9ZnSN5jwbqVtvNeE8Ke05qCdHZ3apAvzBigewen7nffFyljCUlgOHDiVianl2WliYjRlAPwTQ6iLOls07ZROPZjKdIVU~8wByJkBDi45s0a4BTh3Gm7wt0m3aMHKzmuUr5RxU07BVgH1pkskVe0i4~To03XtwHIFpNMxdvmwcVn4oBpnM0fCl1RUD3IQCMbR1QmTx9CAdMfRdvkBPYmOEGppee4hWsrSu5TcwuGe8pH4sE4TGecFKKhXd8iU0Q5nA__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

Chapter V

Supp. File 1:

https://docs.wixstatic.com/ugd/42cff7_b64fdd27335a488a9a97ce9040f7b6a9.pdf

Supp. Table 1:

https://docs.wixstatic.com/ugd/42cff7_b790267e56194c11b3136a9dfdcd94e1.xlsx?dn=Suppl.%20Mat.%20Table%201.xlsx

Supp. Table 2:

https://docs.wixstatic.com/ugd/42cff7_3c420c46f2c2489fb540ff115fbc1d9e.xlsx?dn=Suppl.%20Mat.%20Table%202.xlsx

Supp. Figs.:

https://docs.wixstatic.com/ugd/42cff7_f2729f1eba304ab8b5023ed5c27d56fc.pdf