Molecular Phylogenetics and Evolution 186 (2023) 107838

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Hidden species diversity and mito-nuclear discordance within the Mediterranean cone snail, *Lautoconus ventricosus*

Samuel Abalde [a,b,*], Fabio Crocetta [c], Manuel J. Tenorio [d], Salvatore D'Aniello [e], Giulia Fassio [f], Paula C. Rodríguez-Flores [b,g], Juan E. Uribe [b], Carlos M.L. Afonso [h], Marco Oliverio [f], Rafael Zardoya [b,*]

[a] Department of Zoology, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden
[b] Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain
[c] Department of Integrative Marine Ecology (EMI), Stazione Zoologica Anton Dohrn, Villa Comunale, I-80121 Napoli, Italy
[d] Departamento CMIM y Q. Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cádiz, 11510 Puerto Real, Cádiz, Spain
[e] Department of Biology and Evolution of Marine Organisms (BEOM), Stazione Zoologica Anton Dohrn, Villa Comunale, I-80121 Napoli, Italy
[f] Department of Biology and Biotechnologies "Charles Darwin", Sapienza University of Rome, Zoology–Viale dell'Università 32, 00185 Rome, Italy
[g] Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA
[h] Centre of Marine Sciences (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005 - 139 Faro, Portugal

ARTICLE INFO

Keywords:
Cryptic species
Mito-nuclear discordance
Species delimitation
Phylogenomics
SNPs
Mitochondrial genomes

ABSTRACT

The Mediterranean cone snail, *Lautoconus ventricosus*, is currently considered a single species inhabiting the whole Mediterranean basin and the adjacent Atlantic coasts. Yet, no population genetic study has assessed its taxonomic status. Here, we collected 245 individuals from 75 localities throughout the Mediterranean Sea and used *cox1* barcodes, complete mitochondrial genomes, and genome skims to test whether *L. ventricosus* represents a complex of cryptic species. The maximum likelihood phylogeny based on complete mitochondrial genomes recovered six main clades (hereby named blue, brown, green, orange, red, and violet) with sufficient sequence divergence to be considered putative species. On the other hand, phylogenomic analyses based on 437 nuclear genes only recovered four out of the six clades: blue and orange clades were thoroughly mixed and the brown one was not recovered. This mito-nuclear discordance revealed instances of incomplete lineage sorting and introgression, and may have caused important differences in the dating of main cladogenetic events. Species delimitation tests proposed the existence of at least three species: green, violet, and red + blue + orange (i.e., cyan). Green plus cyan (with sympatric distributions) and violet, had West and East Mediterranean distributions, respectively, mostly separated by the Siculo-Tunisian biogeographical barrier. Morphometric analyses of the shell using species hypotheses as factor and shell length as covariate showed that the discrimination power of the studied parameters was only 70.2%, reinforcing the cryptic nature of the uncovered species, and the importance of integrative taxonomic approaches considering morphology, ecology, biogeography, and mitochondrial and nuclear population genetic variation.

1. Introduction

Cryptic species are complexes of morphologically indistinguishable but genetically divergent species. Because taxonomic work has been traditionally based on morphological characters, it is likely that many cryptic species have been largely overlooked, raising great interest about their discovery (Bickford et al., 2007; Struck et al., 2018). However, discovering cryptic species is not that straightforward, and the current wide application of molecular markers in taxonomic studies has promoted the description of many complexes of cryptic species (Struck et al., 2018), which ended not been such after more careful morphological evaluations (Karanovic et al., 2016). Cryptic species, however, do exist, and have been discovered within all animal groups (Bickford et al., 2007; Pfenninger and Schwenk, 2007). Importantly, the discovery of these previously overlooked species has profound effects in many fields beyond taxonomy, such as in biodiversity estimates and

* Corresponding authors at: Department of Zoology, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden (S. Abalde). Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain (R. Zardoya).
E-mail addresses: saabalde@gmail.com (S. Abalde), rafaz@mncn.csic.es (R. Zardoya).

https://doi.org/10.1016/j.ympev.2023.107838
Received 6 October 2022; Received in revised form 15 May 2023; Accepted 31 May 2023
Available online 5 June 2023

monitoring efforts used in conservation (Bickford et al., 2007; Peters et al., 2013), ecological studies (Fišer et al., 2018; Pante et al., 2015), evolutionary inferences (Struck et al., 2018), and more generally, any comparative biology study.

The phylum Mollusca is the second most diverse phylum on Earth in terms of number of species, and thus, it is of wide interest to both researchers and amateur naturalists, making mollusk taxonomy a focus of intense study since the time of Linnaeus (Ponder et al., 2020). Yet, mollusk taxonomy has been built over the centuries on morphological descriptions of sometimes misleading characters like shell patterns and colorations, which has led to the incorrect definition of species within many lineages. In fact, the application of molecular markers to the study of species diversity within this phylum has uncovered numerous instances of morphological convergence (e.g., Chang and Liew, 2021; Donald and Spencer, 2016), strong intraspecific variation (e.g., Johannesson et al., 2010; Pardos-Blas et al., 2019), and cryptic species (e.g., Aissaoui et al., 2016; Calvo et al., 2009; Lemer et al., 2014; Moles et al., 2021; Yang et al., 2018). Cone snails are not an exception, as the different species had been mostly defined based on shell phenotypes and can be considered a point example of the importance of including molecular markers in taxonomic studies (Abalde et al., 2017a).

The family Conidae J. Fleming, 1822 is a hyperdiverse group of worldwide distributed marine snails, including almost 1,000 species (WoRMS Editorial Board, 2023). Because of their rich species diversity, key ecological role, and sophisticated venom system, cones are an important model system for evolutionary studies in the marine realm. Cone snail phylogenomics, evolution, and venomics have been studied using target-capture data (Phuong et al., 2019; Phuong and Mahardika, 2018), transcriptomes (e.g., Abalde et al., 2020; Dutertre et al., 2014; Menting et al., 2016; Pardos-Blas et al., 2019), and recently, two chromosomal-level genomes (Pardos-Blas et al., 2021; Peng et al., 2021). Yet, cone snail taxonomy still relies almost exclusively on the description of shell characters, which are known to be poor predictors for species delimitation. For instance, a phylogenetic study based on complete mitochondrial genomes of cone snails from West Africa (an important hotspot concentrating 10% of the species diversity of the group) has drawn a concerning scenario. Almost half of the total species of this region were synonymized, whilst new species were found hidden under the convergent morphology of their shells (Tenorio et al., 2020). These results have profound implications for designing conservation strategies in the Cabo Verde archipelago (Peters et al., 2013) and even in more general studies focusing on cone snail evolution (Phuong et al., 2019). The case of the *Virroconus ebraeus / judaeus* species complex is another particularly relevant example of cryptic diversity. These two species live sympatrically and share identical shell morphologies, although they have different radular morphologies and feeding habits (Duda et al., 2009a, 2009b), making them a sound model system to study venom evolution (Pardos-Blas, Tenorio, Galindo, & Zardoya, 2022). Although an improvement from exclusively morphological taxonomy, molecular systematics in cone snails mostly relies on mitochondrial markers, with very few examples corroborating their results with nuclear data (Duda et al., 2008; Puillandre et al., 2014a).

Animal mitochondrial and nuclear genomes follow two distinct modes of inheritance, which results in different evolutionary dynamics and, in some cases, notorious mito-nuclear discordance in the genetic structure of populations (Toews and Brelsford, 2012). This discrepancy is particularly troublesome in the case of cryptic species. Because they are morphologically indistinguishable, genetic structure is often the only way to differentiate species. However, if there is discordance between mitochondrial and nuclear markers in a given group, it will lead to an incorrect species delimitation, thus over- or under-estimating its true diversity (Bernardo et al., 2019; Hinojosa et al., 2019; Shults et al., 2022). Thus far, the only evidences of discordance between mitochondrial and nuclear data in cone snails have been reported at the family and genus levels, with markers from the two genomes rendering contrasting tree topologies, but their implications at the populational level and in species delimitation have not been further investigated (Phuong et al., 2019; Wood and Duda, 2021). Here, we explore this putative scenario using as case study another species of cone snail still in need of a careful taxonomic evaluation, *Lautoconus ventricosus* (Gmelin, 1791). The Mediterranean cone snail, *L. ventricosus*, is rapidly becoming a model for the study of cone snail venom evolution, upon the high-quality assembly of its genome, several transcriptomes, and breakthrough discoveries like the presence of conotoxin gene expression in tissues outside the venom gland (Pardos-Blas et al., 2021); hence the need to re-assess its taxonomic status based on a robust multilocus approach. Its restricted distribution allows for an extensive geographical sampling, making it a sound model to test the population genetic structure using both mitochondrial and nuclear markers in a wide area.

The Mediterranean cone snail was one of the earliest described species of cones, and for many years, either under the name of *Conus ventricosus* or as (the now synonymized) *Conus mediterraneus* Hwass in Bruguière 1792, this species was thought to be distributed from the Mediterranean Sea to the neighboring Atlantic regions, including the Canary Islands and the African coast down to Senegal (for more details see Bandel and Wils, 1977; Pin and Leung Tack, 1995). However, a recent molecular study of the cone snails endemic to Senegal showed that early records of *C. mediterraneus* in this country actually corresponded to a different species, restricting *L. ventricosus* distribution to the Mediterranean Sea up to the south coast of Portugal (Abalde et al., 2017b). Likewise, despite the many attempts to describe different cone species inhabiting the Mediterranean Sea, the general consensus in the last decades has been to consider them all as junior synonyms of *L. ventricosus* (WoRMS Editorial Board, 2023).

However, the only molecular phylogeny published to date including more than one specimen of *L. ventricosus* has challenged this view (Abalde et al., 2017a). In fact, the molecular divergence in mitochondrial DNA estimated for two specimens was higher than in any other pair of sister species included in the phylogeny, suggesting that "*L. ventricosus*" could actually represent a species complex. The present study builds on that hypothesis and proposes, after an exhaustive sampling and a careful analysis of mitochondrial and genome-wide data, the existence of at least three species inhabiting the Mediterranean Sea. Moreover, this study shows evidence of mito-nuclear discordance in cone snails and discusses its implications in species delimitation, prompting for a careful evaluation of the family systematics based on the comparison of both mitochondrial and nuclear markers.

2. Materials and methods

2.1. DNA extraction, PCR amplification and sequencing

A total of 245 individuals of the Mediterranean cone snail *L. ventricosus* were collected from 75 localities (Fig. 1; Supp. Mat. Table 1). Unfortunately, no specimen from the easternmost coast of the Mediterranean Sea and the northern coast of Africa could be collected. However, individuals from Djerba and Bizerte, in Tunisia, were accessible through the malacological collection of the Natural History Museum of Paris (MNHN), and one partial mitochondrial *cox1* gene sequence from a cone snail collected in Israel is publicly available on GenBank. DNA was extracted using the QIAGEN DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany). The barcoding region of the mitochondrial *cox1* gene was successfully amplified using universal primers (Folmer et al., 1994). The PCR reaction contained: 2.5 μL of 10 × buffer, 0.5 μL of dNTPs (2.5 mM each), 0.5 μL of each primer (10 mM), 1 ng of template DNA, 0.2 μL of TaKaRa Taq™ (R001A) DNA polymerase, and sterilized distilled water up to 25 μL. The PCR cycle was: a denaturation step at 94 °C for 60 s; 45 cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 30 s, and extension at 72 °C for 90 s; followed by a final extension step at 72 °C for 5 min. These short fragments were Sanger sequenced at Secugen (Madrid, Spain).

Based on a reconstructed phylogenetic tree (Supp. Mat. Fig. 1), a

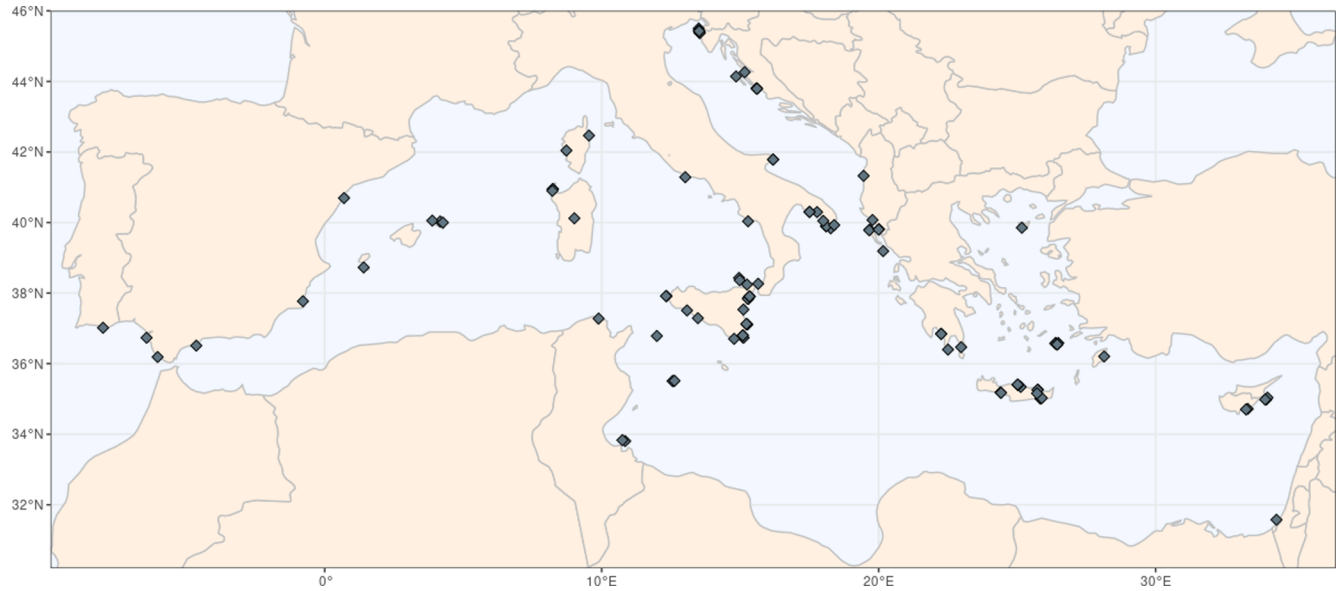

Fig. 1. Map of the sampling locations where individuals of L. ventricosus included in the *cox1* tree were collected. The map was drawn on R using the ggplot2, sf, rnaturalearth, rgeos, and spData packages (Bivand et al., 2022; Bivand and Rundel, 2021; Pebesma, 2018; South, 2022; Wickham, 2016).

subset of 22 specimens representing six main clades (see results) was selected. Their nearly complete mitochondrial genomes (only lacking the control region and the neighboring regions) were amplified using two alternative approaches. For those having DNA of good quality, an approach based on long PCRs was used, following the methods described in Uribe et al. (2017). For those cases in which DNA was too fragmented, standard PCRs were performed using the primers and methods described in Tenorio et al. (2020). Short fragments were directly Sanger sequenced whereas long PCRs were mixed in equimolar concentrations to construct NEXTERA XT DNA libraries (Illumina, San Diego, CA, USA) that were sequenced in an Illumina HiSeq2000 platform at Sistemas Genómicos (Valencia, Spain).

Finally, a subset of 15 specimens from the six clades recovered in the mitogenome-based phylogenetic tree (see results) was selected for genome skimming i.e., low coverage (1–3×) genome sequencing, along with two outgroups from the sister genus *Africonus* Petuch, 1975 endemic to Cabo Verde islands (Abalde et al., 2017a). DNA extracts for genome skimming were used to construct indexed libraries with the TruSeq Nano DNA library prep kit (Illumina, San Diego, CA, USA). The libraries were sequenced in a NovaSeq 6000 platform at Macrogen (Seoul, South Korea). For each individual, roughly 10 Gb of 150 base pairs (bp) PE reads were produced, corresponding to an approximate 3× coverage based on the reported genome size of this species (Pardos-Blas et al., 2021). Note that, because of an error during the sequencing, we only received the forward raw reads of the sample SZN064. In practice, the only consequence of this error was the loss of half of the information for this specimen, but despite that, all the results from this sample seem to be correct.

2.2. Assembly and annotation

The Sanger sequences corresponding to the forward and reverse primers were assembled into individual fragments using Geneious Prime 2019.0.3. The Illumina reads generated from the long PCRs were assembled in Geneious Prime 2019.0.3 as follows: first, clean reads were mapped to the corresponding *cox1* sequence with a minimum identity of 99% and a minimum overlapping fragment of 75 bp; then, this fragment was extended through successive mapping iterations of reads with a minimum identity of 100% and a minimum overlap of 100 bp until reaching the end of the amplified fragment; finally, all mitogenomes were aligned to locate any potential assembly error, and later confirmed and corrected by looking at the mapping stats. The Illumina reads corresponding to mitochondrial DNA obtained from low coverage genomes were assembled using NOVOPlasty v.4.2 (Dierckxsens et al., 2017). All mitogenomes, regardless of their source, were annotated with the option "Annotate from Database" in Geneious Prime 2019.0.3 and using the mitochondrial genomes from closely related species as reference. Protein-coding genes were manually inspected to ensure the presence of start and stop codons, tRNA genes were annotated using the MITOS2 webserver (Donath et al., 2019), and rRNA genes were assumed to extend to the boundaries of adjacent genes as in other cone snail mitogenomes previously described (e.g., Abalde et al., 2017a).

Previous to the assembly of the nuclear genomes, mitochondrial reads were discarded by mapping them to the assembled mitochondrial genomes using Bowtie2 v.2.2.6 (Langmead and Salzberg, 2012). Nuclear gene sequences were cleaned using Prinseq v.0.20.3 (Schmieder and Edwards, 2011) with the following parameters: reads with a quality mean below 25, with N's in more than 25% of the sequence, or sequences of low complexity (minimum entropy of 50) were filtered out; in addition, sequences were trimmed in the 3′ and 5′-end until the PHRED quality score was above 30. Clean reads were assembled in SPAdes v.3.14.1 (Bankevich et al., 2012) using five Kmer lengths (21, 33, 55, 75, and 121). The resulting genomes were scaffolded with RagTag v.2.0.1 (Alonge et al., 2021) using the high-quality genome of *L. ventricosus* (Pardos-Blas et al., 2021) as reference. Genome completeness was assessed with BUSCO v.3.0.2 (Seppey et al., 2019) and the metazoa_odb9 database. Finally, the ORFs of protein-coding genes were annotated and extracted using Scipio v.1.4.1 (Keller et al., 2008) and the annotated proteins of the *L. ventricosus* genome as reference.

2.3. Phylogenomic inference

Phylogenetic analyses were performed over three datasets. First, newly produced as well as GenBank-downloaded *cox1* nucleotide sequences of *L. ventricosus* (249 in total, Supp. Mat. Table 1) were aligned with 16 outgroups using the MAFFT algorithm and cleaned with Gblocks (allowing gap positions within the final blocks, but not many contiguous non-conserved positions) in TranslatorX (Abascal et al., 2010). A maximum likelihood (ML) phylogenetic tree was inferred using IQ-TREE v.1.6.12 (Nguyen et al., 2015), which detects the best-fit substitution model using ModelFinder (Kalyaanamoorthy et al., 2017) and the Bayesian Information Criterion (BIC) (Schwarz, 1978). Nodal support

was calculated using 1,000 ultrafast bootstrap replicates and the SH-like approximate likelihood ratio test with 1,000 replicates (Guindon et al., 2010; Hoang et al., 2018).

Second, 13 protein-coding and two rRNA genes were used for phylogenetic inference based on complete mitochondrial genomes. The nucleotide sequences of the protein-coding genes were codon-aligned using the MAFFT algorithm and cleaned using Gblocks in TranslatorX. The nucleotide sequences of the rRNA genes were aligned using MAFFT v.7 (Katoh and Standley, 2013) and cleaned with Gblocks 0.91b (Castresana, 2000). The 15 gene alignments were converted into PHYLIP format in the ALTER webserver (Glez-Pena et al., 2010) and concatenated in Geneious Prime 2019.0.3. The best model-fit partition scheme was selected using the IQ-TREE built-in algorithm ModelFinder (Kalyaanamoorthy et al., 2017) with the BIC and separating the protein-coding genes per codon position, but allowing the algorithm to merge the partitions with the same model. Finally, the ML phylogenetic tree was inferred using IQ-TREE v.1.6.12 (Nguyen et al., 2015) with 1,000 ultrafast bootstrap replicates (Hoang et al., 2018). Additionally, a phylogenetic tree was reconstructed under Bayesian Inference (BI) using PhyloBayes v.1.5 (Lartillot et al., 2013). Specifically, we ran three independent chains, under the CAT-GTR model with four gamma distributions and not removing constant sites. The chains ran until they reached convergence around 14,000 generations, which was measured with the built-in commands *bpcomp* and *tracecomp*, removing the first 1000 generations as burn-in and sampling every ten generations. Following the recommendations from the manual, we aimed at a 0.1 maximum difference and 100 minimum effective size.

Third, the protein-coding genes extracted from the nuclear assembly were analyzed at the nucleotide level (thus avoiding potential errors in the identification of exon–intron boundaries that would break the reading frame of the protein). Gene redundancy was removed by merging sequences similar >99% with CD-HIT v.4.6.4 (Fu et al., 2012). The filtered genes were classified into orthogroups using OrthoFinder v.2.4.1 (Emms and Kelly, 2019). Among the 23,193 orthogroups identified, only those that were single-copy and were present in all individuals were considered. Each orthogroup was aligned with MAFFT v.7 and cleaned using BMGE v.1.12 (Criscuolo and Gribaldo, 2010), removing all positions with more than 70% of missing data as well as specimens lacking at least half of the sequence length, but allowing fast-evolving sites (-h 1 -w 1). Given the topological mismatch between the mitochondrial and nuclear phylogenies (see results), the nuclear genes used for phylogenetic inference were tested for several known systematic errors following the recommendations by Mongiardino Koch (2021), whose genesortR script aims to identify those genes that maximize phylogenetic information, while minimizing the influence of systematic errors. The best 500 genes were selected, and symmetry tests were performed over these genes with IQ-TREE v.2.1.3 (Minh et al., 2020), to finally keep 437 genes that respect the assumptions of stationarity and homogeneity. Phylogenetic inference under ML was performed using IQ-TREE v.1.6.12, allowing the program to infer the best substitution model for each partition using BIC, and ASTRAL v.5.5.7 (Zhang et al., 2017), after inferring a gene tree for each alignment using IQ-TREE v.1.6.12. An additional phylogenetic tree under BI was carried out with PhyloBayes v.1.5 (Lartillot et al., 2013). Due to difficulties to reach convergence over the 437 genes dataset, we ran this analysis over Datasets 1 and 2 (n = 50 genes) prepared for analyses of species delimitation (see below). The same parameters as in the mitochondrial dataset were applied, but the chain ran for 17,000 and 35,000 generations, respectively.

In addition, pairwise genome-wide distances were inferred using Skmer (Sarmashghi et al., 2017), which first generates the kmer spectra (k = 21) for each genome, and then computes the distances using the Jaccard Index. Uncorrected distances were then converted into JC69 and used to infer a neighbor-joining tree with FastME (Lefort et al., 2015). Skmer is designed to calculate genomic distances based on raw read data. However, a careful inspection of the kmer spectra generated showed that in some individuals the kmer diversity was overestimated, probably due to poor DNA quality. Hence, assembled genomes were used instead of raw reads as input for the analysis.

2.4. Single nucleotide polymorphism (SNP) analysis

The clean reads generated for the nuclear genome assembly were mapped to the reference genome of *L. ventricosus* using the BWA-MEM algorithm from BWA v.0.7.16 (Li and Durbin, 2009), sam files were converted to bam and sorted using samtools v.1.5 (Danecek et al., 2021), and duplicate reads were marked using Picard v.2.25.5 (Institute, 2019). Variant calling was performed with freebayes v9.9.2–27-g5d5b8ac including all samples (Garrison and Marth, 2012). Only variants that passed the following filters were considered: minimum coverage four, exclude unobserved genotypes, minimum mapping quality 30, and minimum base quality 20. Called variants were quality filtered using vcftools v.0.1.17 (Danecek et al., 2011) and vcffilter v.1.0.0 (Garrison et al., 2021). Only variants present in 14 out of 17 samples were considered (–max-missing 0.8) and then filtered out if they did not fulfill any of the following criteria: minimum quality score of 30 (–minQ 30); minimum minor allele count, minimum depth and Minor Allele frequency of three (–mac 3; –minDP 3 –maf 3); average number of mapped reads of at least 15 (–min-meanDP 15); Allele Balance, the relative support of each allele, between 0.2 and 0.8, but keeping homozygotic alleles (AB > 0.2 & AB < 0.8 | AB < 0.1); the ratio between the reference and the variant allele qualities must be around one (MQM/MQMR > 0.9 & MQM/MQMR < 1.05); reads pairs must be concordantly mapped. Finally, following the recommendations by Li and Wren (2014) to avoid overestimation of mapping quality due to high mapping depths, read depth of a given allele must be lower than the average depth plus the square root of the average depth times three, and its mapping quality lower than two times its read depth. Only SNPs were considered in subsequent analyses.

Plink v.1.07 (Purcell et al., 2007) was used to detect and discard all SNPs under linkage disequilibrium, using a sliding window of 100,000 bp, moving 10 bp in each step, and pruning any variant with a r2 greater than 0.1. Then, a Principal Component Analysis (PCA) was performed over this dataset in Plink and visualized in R using the package ggplot2 v.3.3.5 (Wickham, 2016). Due to the differences observed between the two outgroups and the samples SZN047 and SZN064 (see results), a new PCA was performed excluding these four samples from the analysis. SNP data were also used for phylogenetic analysis. Briefly, those SNPs present in all samples were converted into a PHYLIP matrix using vcf2phylip v.2.7 (Ortiz, 2019) and analyzed under ML in IQ-TREE v.1.6.12 allowing the program to infer the best model-fit using BIC but applying the ascertainment bias correction (Lewis, 2001).

2.5. Species delimitation

A species delimitation analysis was performed using BPP v.4.4.0 (Yang & Rannala, 2010), which considers that exchange of migrants ceases as soon as species separate and tests accordingly the Bayesian posterior probabilities of all potential species delimitation hypotheses. The initial dataset consisted of all the nuclear genes used for phylogenomic analyses after applying all filters. Instead of analyzing the 437 genes altogether, five smaller datasets containing each 50 genes randomly sampled were created and analyzed independently to allow the detection of alternative species delimitation hypotheses from genes with different evolutionary histories (Supp. Mat. Table 2). BPP requires each individual to be assigned to a species. However, according to the null hypothesis, all of them belong to the same species, *L. ventricosus*. Instead, assuming an alternative hypothesis, individuals were assigned to the clades recovered in the mitogenome phylogenetic tree (see results), as these represent the potential maximum number of cryptic species within *L. ventricosus*. The only exception was the brown clade, which was not recovered in the nuclear tree (see results), suggesting that

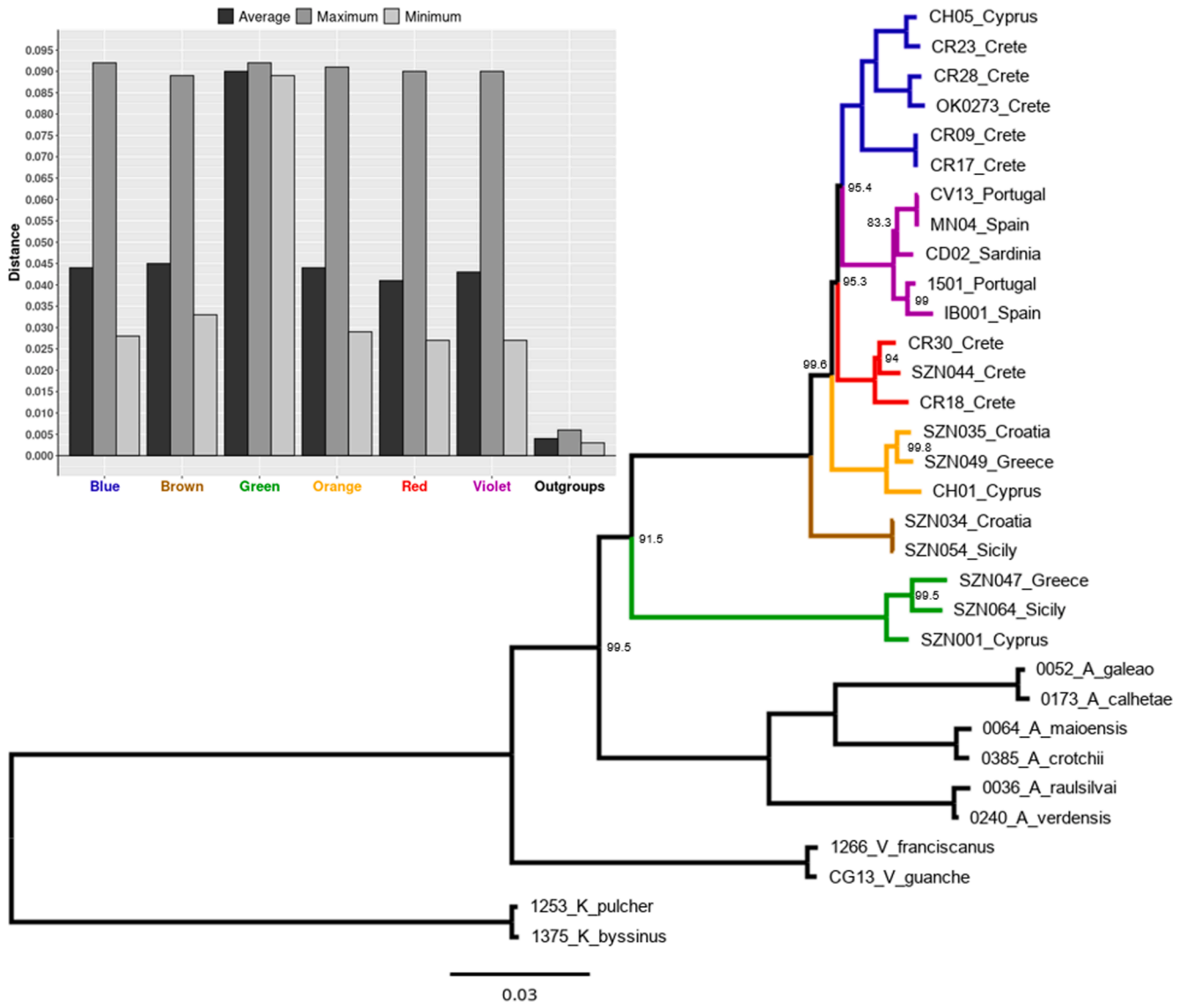

Fig. 2. Maximum likelihood phylogenetic tree based on complete mitochondrial genomes and inferred with IQ-TREE. All nodes have maximum statistical support unless otherwise specified. The scale bar represents substitutions per site. The inset shows the average, maximum, and minimum pairwise distance observed between the six main clades. Outgroups were not compared to the ingroup clades but distances were measured between outgroup sister pairs.

it actually represents an artifact due to introgression and/or Long Branch Attraction (LBA). The two specimens of the brown clade were assigned to the violet and orange clades, respectively, following the nuclear tree topology (see results).

Species delimitation hypotheses were tested on each of the five datasets using either the tree generated during phylogenomic analyses as guide tree (model A10) or letting the program to infer its own guide tree (model A11). In both cases, algorithm 0 (with $\varepsilon = 2$) and algorithm 1 ($\alpha = 2$, and $m = 0.5$) were employed. Two independent runs per analysis were performed to assess convergence; however, for dataset 2 using algorithm 1, another two additional runs were necessary to achieve chain convergence. Following the BPP tutorial (Flouri et al., 2020), the *thetha* and *tau* priors were set as $\alpha = 3$ and $\beta = 0.004$, respectively. The option to fine tune all priors was active in all analyses, which ran for a total of 10 million generations, sampling every 10 generations, resulting in a MCMC chain of 1 million samples. The first 10,000 samples were discarded as burn-in.

2.6. Estimation of divergence times

Bayesian molecular dating was performed over the nuclear and mitochondrial datasets on MCMCtree (Rannala and Yang, 2007) within PAML v.4.10 (Yang, 2007). The nuclear data consisted of three sets of 50 genes (datasets 1, 2, and 4 in BPP), in each case considering each gene an independent partition. The mitochondrial matrix was divided into the four partitions identified by PartitionFinder during phylogenetic analyses: all protein-coding genes together but separating the three codon positions, and the two rRNA genes together. The same three calibration points were defined for both analyses: first, the origin of the oldest island in the Cabo Verde archipelago, Sal, 28 million years ago (Mya) (Holm et al., 2008), was used to calibrate the node separating the genera *Africonus* and *Lautoconus* Monterosato, 1923. A skew-normal distribution ranging between 20 and 28 and with mode in 26 Mya was defined, with scale 1.85 and shape = −10, and allowing a 1% probability for this node to be younger than the predefined range. Second, the node including the last common ancestor of the *L. ventricosus* species complex was calibrated with a uniform distribution between 10 (the oldest fossil of the species is dated in the middle-upper Miocene (Tucker and Tenorio,

2009) and 20 Mya (the lower end of the preceding node), with a probability of 10% of being older than that. Finally, the separation between the two outgroup species, *Africonus crotchii* and *Africonus maioensis,* was calibrated based on previous estimates (Abalde et al., 2017a). The prior distribution of this node was defined with a uniform distribution ranging from 0.5 to 2 Mya, a 0.08 offset, and a 5% probability of the node age falling outside this threshold, either younger or older. Additionally, because calibrating the root is a requirement of the program, we calibrated the divergence time between *Kalloconus* da Motta, 1991 and all other cones in the mitochondrial analysis with a uniform distribution between 30 and 40 Mya and a 10% probability of falling outside this range based on previous estimates (Abalde et al., 2017a). All the calibration parameters were defined with the MCMCtreeR package (Puttick, 2019). We assumed an uncorrelated lognormal clock and a uniform distribution of the birth and death rates ($\lambda = \mu = 1$, and $\rho = 0$). The most complex substitution model integrated in MCMCtree, HKY with five gamma categories and a 0.5 alpha, was applied to the analysis. Finally, the alpha prior of the mean substitution rate was set on 2, while the beta prior was estimated from the data (nuclear: rgene_gamma = 2 4200 1; mitochondrial: rgene_gamma = 2 560 1). Briefly, a phylogenetic tree from the two matrices was inferred using the same partitions and model parameters applied in MCMCtree, and the branch lengths between *Africonus* and *Lautoconus* were used to calculate beta (beta = alpha * divergence time/branch length). MCMCtree ran for 11 million generations, sampling every 100 generations, and discarding the first million as burn-in. Tracer (Rambaut and Drummond, 2007) was used to ensure chain convergence and to check that the ESS of all parameters was above 200.

2.7. Morphometric analyses

For comparisons of shell morphometry, we performed analyses of the covariance (ANCOVA) for different shell parameters, namely: maximum diameter (MD), height of the maximum diameter (HMD), and spire height (SH), using species hypotheses as factor and shell length (SL) as covariate. Additionally, the mean values of SL using t- and U-tests were also compared statistically. The corresponding morphometric parameters of 100 shells selected among voucher specimens from the *cox1* phylogeny were measured on their digital photos using the software GIMP 2.10.20 (The GIMP Development Team, 2019). All measurements can be found in Supp. Mat. Table 3. Statistical tests were carried out among mitochondrial (but brown) and nuclear clades with STATGRAPHICS CENTURION 18 software (Statgraphics Technologies, USA) after ensuring all measures passed the normality tests.

3. Results

3.1. Mitochondrial phylogenetics

The final *cox1* matrix was formed by 265 sequences (including 249 individuals of the *L. ventricosus* complex and 16 outgroups from closely related cone snail genera) and consisted of 543 characters. The BIC best-fit model was TN + F + G4. The inferred ML phylogenetic tree revealed the presence of six main clades that could potentially represent different species (Supp. Mat. Fig. 1). However, weak nodal support prompted the generation of more data to confirm this result (see below). These clades are hereafter named according to the color code used in all figures: blue, brown, green, orange, red, and violet, respectively. More than half of the specimens analyzed (144 out of 249, 57.83%) belonged to the violet clade, whereas the blue one included almost one fifth of the specimens (46 individuals, 18.47%). The other four clades contained <10% of the samples each (orange: 23 individuals, red: 14, green: 12, and brown: 10). It was noteworthy that, although tentatively considered independent because of its long branch, the brown clade was deeply nested within the violet clade. The phylogenetic analyses recovered the green clade as sister to another clade including all other lineages and separated by a relatively long branch (indicating important sequence divergence). Within the second clade, inter-clade relationships were unclear due to weak nodal support (Supp. Mat. Fig. 1).

In order to further confirm these clades, the complete (15) and almost complete (five) mitochondrial genomes from representatives of each clade were sequenced. In total, 20 new plus two already available mitochondrial genomes were analyzed: six from the blue clade, five from the violet, three from each the green, orange and red, and two from the brown. The final matrix, concatenating 13 protein-coding and two ribosomal genes, was 13,490 positions long. According to BIC, the best-fit partition scheme included six partitions: all first codon positions, but those of *cob* and *cox* genes, plus the ribosomal genes (HKY + F + R2); all second codon positions but those of *cox1* and *cox2* (TVM + F + R2); all third codon positions but those of *cob*, *nad4*, and *nad5* (TPM3 + F + G4); and the remaining first (TN + F + R2), second (HKY + F), and third (GTR + F + G4) codon positions.

The tree topology based on mitochondrial genomes was partly discordant with that based only on the *cox1* matrix, but had higher statistical support for the six main clades and all but five nodes received maximum statistical support (Fig. 2, Supp. Mat. Fig. 2). As in the *cox1* phylogenetic tree, the green clade was sister to all other lineages and clades violet and blue were sister to each other. The red clade was recovered as sister to the violet + blue clades, with the orange clade sister to the three (Fig. 2). The main difference with respect to the *cox1* phylogenetic tree involved the brown clade, which was recovered outside the violet clade, as one of the first offshoots of the ingroup, branching right after the green clade (Fig. 2). The two specimens forming this clade had exactly the same mitogenome sequence (the divergent BAU_1151_2_PUG from this clade could neither be PCR amplified nor Illumina sequenced due to low DNA quality). Regarding branch lengths, in all cases branches leading to each of the main clades were longer than those of the species pairs included in the outgroup, suggesting that *L. ventricosus* could actually represent a complex of six species. The geographical distribution in the Mediterranean Sea of these six clades is shown in Supp. Mat. Fig. 3.

3.2. Genome-wide analyses

The low coverage genomes of 15 individuals from the six clades plus two outgroups from the genus *Africonus* were sequenced and assembled. Roughly 10 Gb were sequenced per specimen, generating between 61 and 76 million reads (Table 1). After removing all mitochondrial reads and filtering by quality, above 80% of the reads were maintained in all individuals but five, leaving between 43 and 65 million reads available for assembly. The specimen with the lowest read quality was SZN035, keeping only 68.5% of all reads. This metric did not have an effect in the final coverage of the assemblies (between 6 × and 7 × in all individuals), but in genome sizes. Although all assembled genomes were similar in size (between 1.1 and 1.4 Gbp), those returned from samples with more reads available for assembly were longer. Therefore, specimen SZN035 presented the shortest genome. These and other genomic statistics are shown in Table 1.

The number of proteins annotated in the genomes varied between 14,505 and 21,600, later reduced to 13,576 and 19,854 after removing duplicates. Among the 23,193 orthogroups identified by OrthoFinder, only 2,867 were single-copy and retrieved in all specimens, of which 2,447 were retained after all cleaning steps. The best 500 genes selected by the genesortR script (Mongiardino Koch, 2021), which is meant to minimize the presence of systematic errors in the data, were selected for further processing (Supp. Mat. Fig. 4 shows gene properties as measured by genesortR for all 2,447 genes sorted from best to worst, left to right). Another 63 genes were further removed for violating the assumptions of

stationarity and homogeneity. The final matrix contained 18 taxa (17 newly sequenced plus the reference genome of *L. ventricosus*) and 437 genes, with a total length of 524,651 nucleotides. Concatenation- (IQ-TREE, Fig. 3) and coalescence-based (ASTRAL, Supp. Mat. Fig. 5) phylogenetic analyses were performed. The two recovered tree topologies were congruent, with only slight differences in the relationships within the violet clade and in branch lengths. Nodal support was generally lower in the coalescence tree, and topology differences were associated with nodes with low statistical support. PhyloBayes recovered a polytomy in the ingroup that precluded topological comparisons (Supp. Mat. Fig. 2). In contrast to the mitochondrial trees, the nuclear topologies only recovered four out of the six clades (Fig. 3 and Supp. Mat. Fig. 5). The mitochondrial blue and orange clades were recovered as a single one, with specimens thoroughly mixed within the clade. Moreover, the brown clade was not recovered in the nuclear phylogenetic trees (Fig. 3 and Supp. Mat. Fig. 5). The two specimens were placed distantly in the tree (SZN054 as sister to the violet clade, and SZN034 nested deep within the blue + orange clade). Regarding inter-clade relationships, the only concordance with the mitochondrial tree was the position of the green clade as sister to the others. In the nuclear trees, the next offshoot was the violet clade, followed by the red one, sister to the mixed blue plus orange clades (Fig. 3 and Supp. Mat. Fig. 5).

A total of 115,797,225 variants were annotated in the samples. After applying several quality filters, the variants were reduced to 169,625 (0.15%), most of which were SNPs (156,774; 92.42%). The final dataset was composed of 11,256 SNPs, as the remaining SNPs were identified as evolving under linkage disequilibrium and discarded. A PCA analysis identified three major clusters on these data: the two outgroups, the two individuals of the green clade, and all other clades together, although showing some clustering (Fig. 4A). A second PCA analysis, this time based only on the 13 samples of the latter cluster, also showed some spatial separation among clades (Fig. 4A). The red and violet clades were recovered relatively far from the others, whereas the differences between orange and blue clades were smaller. Yet, the samples of these two clades were not as mixed as it could be expected based on phylogenomic analyses. The brown clade was not recovered, and its two samples clustered with the violet and blue + orange clades, respectively, in line with phylogenomic results.

A phylogenetic tree was inferred from the SNP data based on an alignment of 1,945 sites with no missing data. This tree was generally well resolved at the tips, but with internal nodes connecting main clades having low support (Supp. Mat. Fig. 6). It presented many similarities with the one based on the phylogenomic data set (Fig. 3), including the long branch leading to the green clade, as first offshoot of the ingroup, and the recovery of the violet (including the specimen SZN054 from the brown clade, in this case nested deep within the clade and not as sister group) and red clades. In contrast, the major differences observed were related to the blue and orange taxa. While most of them grouped within a single clade, two of the specimens were scattered along the tree: the orange CH01 was the sister group to all other taxa but the green clade, and the blue CR17 was recovered as sister to the red clade (Supp. Mat. Fig. 6).

Genome-wide distances among individuals were measured using a kmer approach to compare the individuals and the Jaccard index to calculate pairwise distances corrected by the JC69 model. Pairwise distances were relatively low in all cases, including the ingroup-outgroup comparisons (Fig. 4B). According to the matrix, the only noticeable distances were those including either the two specimens of the green clade (0.0297 and 0.018 in average for SZN064 and SZN047, respectively) or the outgroup taxa with respect to all the other samples. In fact, the largest distance observed was that between the specimen SZN064 and the two outgroups (0.035). Although most distances were rather similar, the phylogenetic tree based on this distance matrix (Fig. 4C) recovered exactly the same topology of the phylogenomic tree reconstructed based on the concatenated nuclear dataset (Fig. 3).

Table 1
Summary of the statistics of the assembled low coverage genomes, including number of reads, contiguity, completeness and number of annotated proteins.

	CD02	CH01	CR17	CR18	CR29	CR30	CV0064	CV0385	CV1501	MN04	OK0273	SZN034	SZN035	SZN047	SZN049	SZN054	SZN064
# reads	71,920,328	72,997,240	71,263,624	74,810,651	69,010,439	61,530,298	72,776,725	73,194,618	62,923,546	61,121,958	62,283,748	66,140,646	62,972,421	76,317,726	73,547,579	67,997,492	67,070,129
# nuclear reads	60,966,552	60,345,249	58,967,922	65,456,991	59,974,512	52,372,937	64,930,856	65,127,572	54,151,808	53,911,590	53,230,999	50,416,104	43,156,891	60,084,361	58,212,288	52,519,953	61,805,537
Genome length (Mbp)	1,309.44	1,282.91	1,262.07	1,335.64	1,302.26	1,242.97	1,366.13	1,418.97	1,271.13	1,227.47	1,201.14	1,195.75	1,079.39	1,348.21	1,304.58	1,218.68	1,366.49
# contigs	379,342	437,915	432,776	423,305	414,295	419,563	474,102	505,641	349,490	365,049	367,412	460,775	393,491	429,093	482,577	412,398	1016,251
Largest contig (Mbp)	74.85	72.60	71.95	76.04	74.92	69.25	77.17	79.94	73.20	70.21	68.63	67.54	61.10	76.80	73.82	68.96	72.53
Average contig length (bp)	3,451.88	2,929.58	2,916.23	3,155.27	3,143.31	2,962.54	2,881.51	2,806.28	3,637.09	3,362.49	3,269.19	2,595.07	2,743.10	3,141.99	2,703.37	2,955.10	1,344.63
N50	31,507,434	31,021,304	30,188,614	32,428,743	31,223,897	28,899,097	32,247,255	33,237,901	31,292,829	29,770,023	28,416,344	27,755,941	24,893,376	31,476,897	30,440,373	28,550,715	29,059,095
Coverage[†]	6.98	7.06	7.01	7.35	6.91	6.32	7.13	6.89	6.39	6.59	6.65	6.32	6.00	6.69	6.69	6.46	6.78
# annotated proteins	21,167	20,555	19,741	21,496	20,948	19,496	19,564	14,505	21,600	21,123	20,077	18,959	17,948	20,659	19,560	19,789	16,031

[†] Calculated as the number of nuclear reads times the read length (150 bp) divided by genome length.

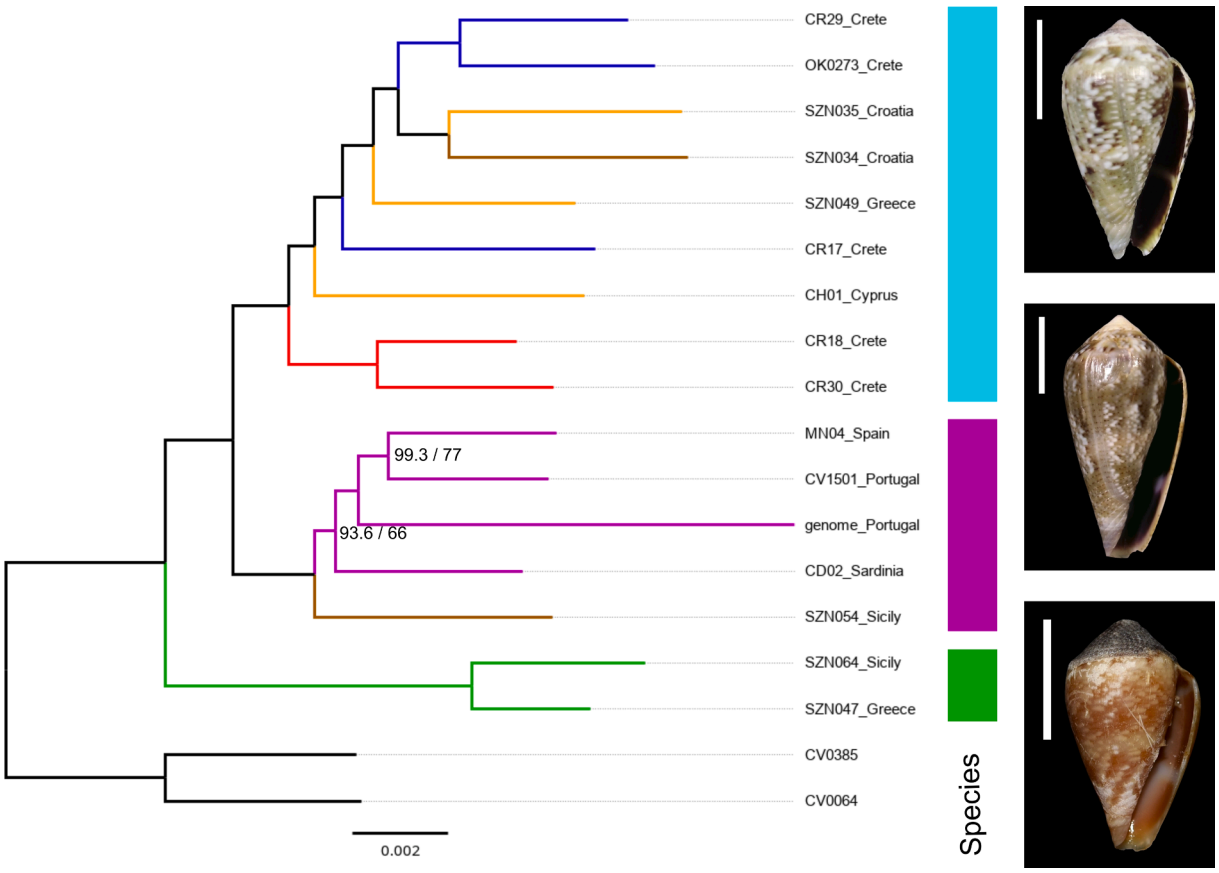

Fig. 3. Phylogenomic tree inferred from 437 nuclear genes filtered for different systematic biases and that passed symmetry tests. Scale bar represents substitutions per site. All nodes received maximum statistical support unless otherwise specified. Branch colors follow the same coding as in the mitochondrial genome tree in Fig. 2. The three color bars represent the species identified by BPP. The three photos to the right correspond to shells of representatives of the three species identified by BPP, from top to bottom: CR18, cyan; SC002, violet; and BAU1550.1, green. In the three photos the scale is 10 mm long. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Species delimitation

Five independent datasets were prepared aiming at exploring how many species would delimit each one of them. Species delimitation analyses returned a mixed signal across datasets, but supported the presence of more than one species under the name *L. ventricosus* (Supp. Mat. Table 4). The two species delimitation hypotheses inferred using BPP propose the existence of three or four species, depending on the dataset, which would coincide with the clades identified in the nuclear-based phylogenomic analyses. Three species were identified by dataset 2, which would correspond to the clades green, violet, and a combination of blue, orange, and red. In contrast, datasets 1, 3, and 4 identified four species, separating the red clade from blue plus orange. No differences were observed between the A10 (species tree provided) and the A11 (species tree inferred from the data) analyses, except for dataset 5. In this case, method A10 identified four species, whereas A11 only three. The posterior probabilities of these species delimitation analyses were generally over 0.80, although between 0.60 and 0.75 with method A11 applied to dataset 3 and method A10 to dataset 5 (Supp. Mat. Table 4). Due to the mixed signal detected in the data, we follow a conservative approach and consider only three species within the *L. ventricosus* species complex. Hereafter, the combination of the specimens from the blue, orange, and red mitochondrial clades will be referred to as the "cyan species", in line with Fig. 3. The other two species would correspond to the green and violet clades in all analyses, and therefore will be referred to as the "green species" and "violet species".

3.4. Divergence time estimations

The three nuclear datasets recovered almost identical divergence times (Supp. Mat. Fig. 7), and those from Dataset 2 were selected for comparison. The mitochondrial and nuclear trees presented essentially different topologies, making main cladogenetic events not readily comparable. Yet, a clear pattern emerged, as the divergence time estimations based on nuclear data were generally older than those inferred with the mitogenomes (Fig. 5, Supp. Mat. Figs. 7 and 8). This was particularly obvious in the estimated age of the node connecting the two outgroup species, *A. maioensis* and *A. crotchii*, which was roughly five times older in the nuclear tree (mean = 7.89 Mya) than in the mitochondrial tree (1.56 Mya). The dates for the separation of the genera *Africonus* and *Lautoconus* were more similar, estimated at 24.52 (nuclear) and 23.52 (mitochondrial) Mya (Fig. 5, Supp. Mat. Fig. 7 and 8). The main date differences between the two chronograms were observed within *L. ventricosus*. Although the two analyses were more or less congruent respect to the origin of the green clade (nuclear = 17.26 Mya; mitochondrial = 20.18 Mya), the estimates of the origins of the other main clades were notably different depending on the dataset (Fig. 5, Supp. Mat. Figs. 7 and 8). The divergence of the violet and cyan species was estimated at 14.53 Mya based on the nuclear data, whereas the origin of the main mitochondrial clades was estimated 7.70 Mya. Within the latter, the last common ancestor of each of the six main clades was estimated to have lived between 0.05 (brown) and 3.81 (blue) Mya. In contrast, the last common ancestors of the three main clades identified in the nuclear phylogeny were estimated to have lived between 6.49 (green) and 13.80 (cyan) Mya. Detailed information of the

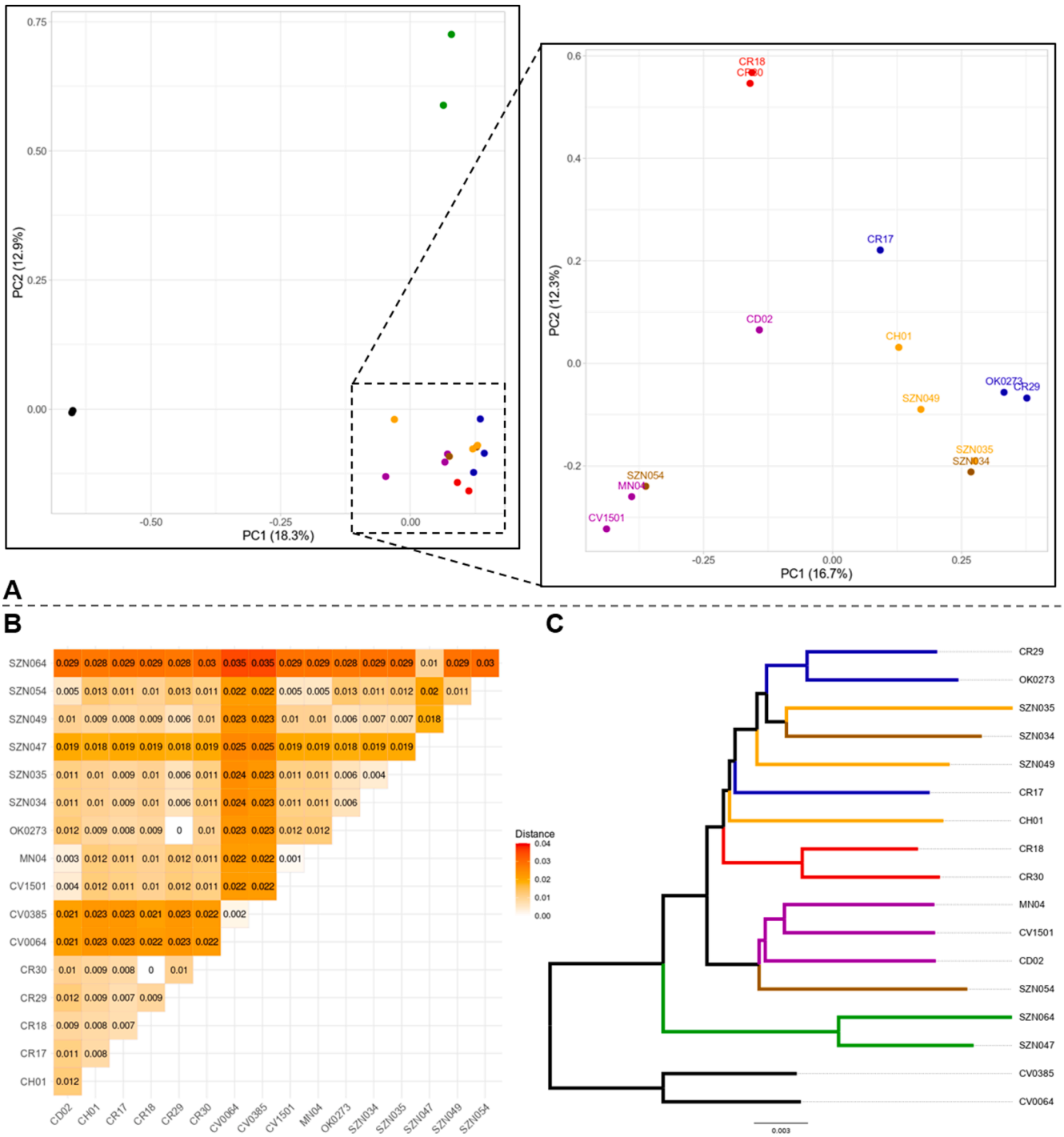

Fig. 4. Summary of the genome-wide analyses. (A) Principal Component Analysis of the SNP data for the full dataset (left) and for all samples but the outgroups and green clade (right). (B) Distance matrix inferred with Skmer. (C) NJ JC69 distance tree inferred from the same data. The scale bar represents the number of differences per site. In all figures, the color coding follows the one shown in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

diversification events of all nodes, including the mean and the 95% lower and upper credible intervals, can be found in Supp. Mat. Fig. 8.

3.5. Morphometric analyses

The shell length (SL) and three morphometric parameters (MD: maximum diameter; HMD: heights of maximum diameter; and SH: spire height) were statistically compared among the three species hypotheses (green, violet, and cyan) identified within the *L. ventricosus* species complex. Pairwise comparisons between the violet and cyan species did not show statistically significant differences in mean SL ($t = -0.107$, $p = 0.915$; $U = 114.0$, $p = 0.791$), as the value was 22 ± 2 mm for both species, but the green species, with a mean SL of 16 ± 2 mm, was clearly smaller and statistically different to both violet ($t = 2.186$, $p = 0.033$; $U = 56$, $p = 0.015$) and cyan ($t = -2.527$, $p = 0.015$; $U = 226$, $p = 0.008$) species. This result, however, must be taken with caution due to the small number of specimens measured for the green species ($n = 6$).

We tested for statistically significant differences in MD, HMD, and SH

by means of analysis of the covariance (ANCOVA) using species hypothesis as factor, the corresponding morphometric parameter as variable, and SL as covariate. We did not find any statistically significant differences in morphometric parameters corrected for differences in SL for any comparison involving the green species (Table 2). In contrast, statistically significant differences were found in MD and SH between the violet and cyan species (Table 2). The shells of the specimens from the cyan species were slightly broader and with a lower spire (least squares mean values: MD = 12.5 ± 0.1 mm; SH = 4.2 ± 0.1 mm) than those of the violet one (least squares mean values: MD = 11.7 ± 0.1 mm; SH = 4.6 ± 0.1 mm). Although subtle, these differences were statistically significant and more apparent in larger individuals, whereas the smaller shells showed a higher degree of overlap. A discriminant function analysis using SL, MD, HMD, and SH as variables only classified correctly 70.2% of the specimens. All the mitochondrial clades but Violet-Green, Orange-Blue, and Blue-Green were statistically different for MD, SH, or both, but never for HMD (Supp. Mat. Table 5).

4. Discussion

The use of partial gene sequences, particularly the *cox1* barcode, is an invaluable resource for identifying species and remains as the main molecular tool in integrative taxonomic studies (Atherton and Jondelius, 2021; Perea et al., 2020; Recuero and Rodríguez-Flores, 2020; Rodríguez-Flores et al., 2021). The steady drop of shotgun sequencing prices, however, is shifting this paradigm towards the use of genome-wide data. In this regard, SNP sequencing has become a common approach to define species limits (Corral-Lou et al., 2022; Gabrielli et al., 2020; Parker et al., 2021; Recuerda et al., 2021), and genome skimming is becoming a popular practice. The advantage of this technique consists in its wide applicability, from the assembly of complete organelle genomes to the annotation of regions of interest such as SNPs, nuclear genes, or ultraconserved regions for its use in phylogenomics (Fernández-Álvarez et al., 2022; Malukiewicz et al., 2021; Ren et al., 2017; Sanchez et al., 2021). In fact, several authors have proposed the use of genome skims in substitution to the traditional *cox1* barcodes (Coissac et al., 2016; Straub et al., 2012). Molecular cone snail systematics has been thus far mostly limited to the use of few PCR-amplified gene fragments (Duda et al., 2008; Duda & Kohn, 2005; Puillandre et al., 2014a), and more recently, complete mitochondrial genomes (Abalde et al., 2017a, 2017b; Tenorio et al., 2020). Here, we have used barcodes, complete mitochondrial genomes, and genome skims to study the *L. ventricosus* species complex, and compared the results obtained from these complementary approaches.

The *cox1* tree inferred from almost 250 barcodes of samples spanning most of the extension of the Mediterranean Sea recovered six main clades, but branches leading to some of them were extremely short and only a handful of nodes were statistically supported. Hence, representatives from each of these clades were selected for complete mitogenome sequencing aiming to generate a more informative matrix. The mitogenome tree recovered the same six clades but sister relationships between clades were different from those observed in the *cox1* tree. All nodes received strong support. Barcodes are an easy way of generating molecular information from hundreds of specimens, but because of their small size the phylogenetic signal may be very limited. In contrast, mitochondrial genomes have proven to be a more cost-effective way for phylogenetic inference than single barcode data (Zaharias et al., 2020), rendering well-resolved trees up to the family level, as it is known that phylogenetic inference increases with the number of genes analyzed (Rokas and Carroll, 2005). Therefore, the tree inferred from complete mitogenomes was considered as our best working hypothesis and up to six clades potentially representing different species were identified.

In order to perform sequence divergence comparisons among main clades within *L. ventricosus*, outgroups were carefully selected to

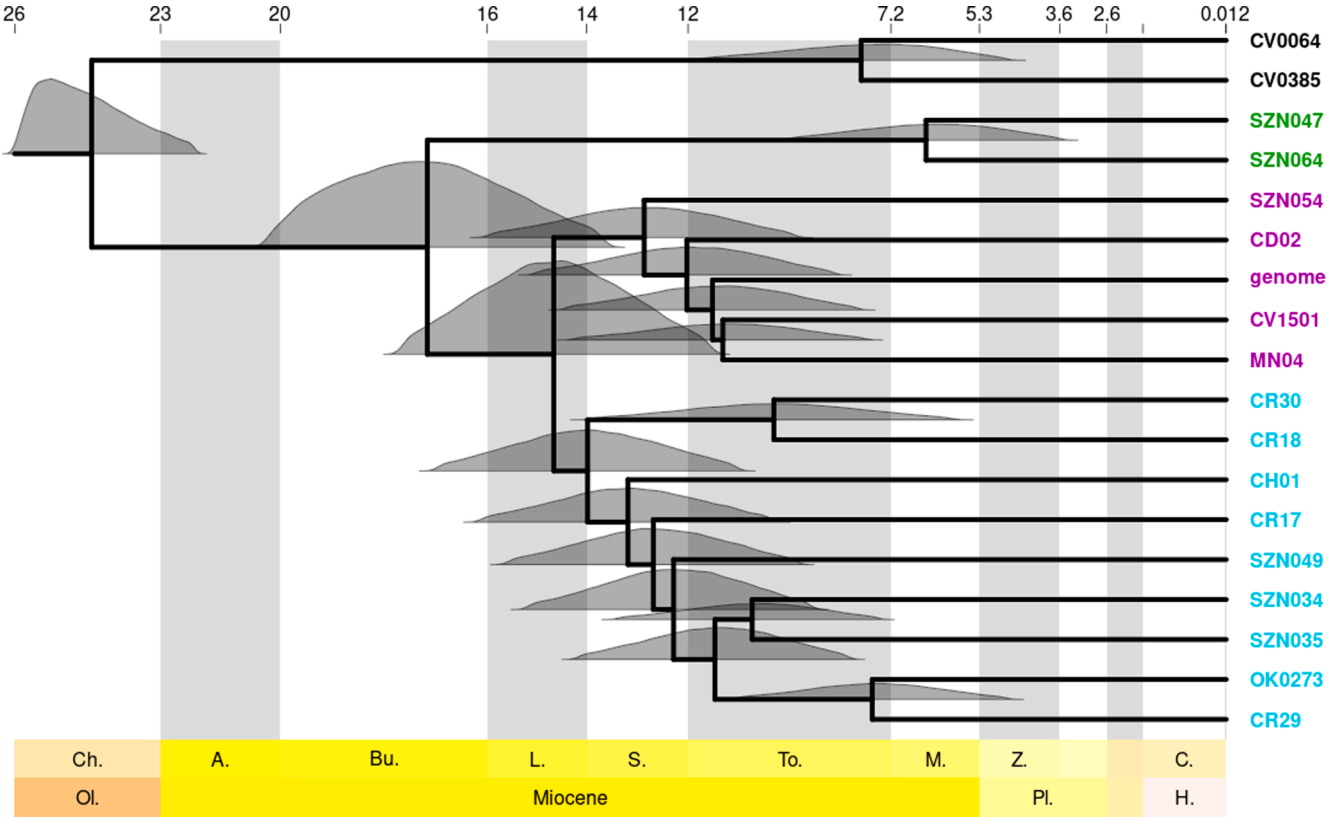

Fig. 5. Chronogram inferred from 50 nuclear genes with MCMCtree (Dataset 2). Each node is located at the mean age inferred for that particular node, with the full distribution of the posterior probabilities within the 95% credible interval attached to each node. Tip colors correspond to each of the three species identified within the *L. ventricosus* species complex. The scale at the top is in million years ago, and the bar at the bottom represent geological epochs and ages.

Table 2
Results of the ANCOVA analyses for the pairwise comparisons of the morphometric parameters of the shells from the three species. Species was used as factor, shell length (SL) as covariate, and all other measurements as variables.

Comparison	Parameter †	F-score	p-value
Violet–Green	MD	0.26	0.61
	HMD	0.45	0.51
	SH	0.71	0.40
Green–Cyan	MD	0.29	0.59
	HMD	0.2	0.66
	SH	2.35	0.13
Violet–Cyan	MD	22.59	0.00
	HMD	3.78	0.05
	SH	6.86	0.01

† MD: maximum diameter; HMD: height of maximum diameter; SH: spire height.

represent pairs of accepted cone snail species. Uncorrected pairwise distances were calculated over the mitogenome matrix with Mega 11.0.8 (Tamura et al., 2021). In all cases, the average, maximum and minimum inter-clade distances within the ingroup were more than 10 times higher than the distance observed between the species pairs from the outgroup taxa (Fig. 2; Supp. Mat. Table 6). With distances between 8.9% and 9.2%, the green clade stands out among all others. The next highest distance, 3.6%, was observed between blue and brown clades, mainly as a consequence of the long branch of the brown clade (Supp. Mat. Table 6). In cone snails, minimum distances of 0.2% (Tenorio et al., 2020) to 1% (Abalde et al., 2017a) have been used to separate species, whereas in other groups, distances between 0.5 and 2% would justify a careful taxonomic evaluation of the specimens involved (Roux et al., 2016), and of 3% would generally be considered enough to validate species as different (Hartop et al., 2022). Here, all six main clades presented a minimum distance above 2.7%, which would be sufficient to justify their consideration as independent species under the above-mentioned criteria.

Besides mitochondrial data, genome skims generated from 15 specimens plus two outgroups were analyzed in different ways. First, 437 nuclear genes were used to infer phylogenomic trees. Second, the high-quality genome of *L. ventricosus* from Olhão, Portugal (violet clade), was used to generate a dataset with thousands of SNPs. Finally, genomic distances among individuals were measured using a kmer approach based on methods employed for organism identification using genome skims. All these approaches rendered generally congruent results, identifying only four out of the six mitochondrial clades, as the blue and orange were thoroughly mixed and the two individuals from the brown clade were recovered in distant positions within the nuclear trees. The mixture of blue and orange individuals within the same clade may indicate incomplete lineage sorting phenomena. The brown clade is the smallest one in the *cox1* tree and the distribution of its individuals is restricted to a narrow region in the central part of the Mediterranean Sea, along the Adriatic Sea and down to the Eastern coast of Sicily (Supp. Mat. Fig. 3). Altogether, the conflicting mitochondrial and nuclear tree topologies and the narrow distribution suggest that the brown clade actually may correspond to individuals from different clades showing effects of an introgression event, and that their position in the mitochondrial tree may be the result of a LBA artifact. The only instance in the nuclear data disagreeing with this general pattern is the PhyloBayes analysis, which failed to resolve the deepest nodes of the tree. The CAT-GTR model is a sophisticated site-specific substitution model, able to perform well even in the presence of LBA artifacts. This inability to reach a strongly supported topology might be the result of insufficient (due to the small size of the dataset) or conflicting signal in the data, which is often neglected by traditional partition analyses. A comprehensive study based on simulation data has shown that, despite the general robustness of PhyloBayes, small datasets are challenging for CAT-GTR, especially in the absence of intra-gene heterogeneity, favoring partition analyses over this algorithm (Whelan and Halanych, 2017). The PhyloBayes tree does not return a contrasting topology, but just an unresolved one, and the existence of four well-supported clades in the IQ-TREE topology is supported by independent analyses based on SNP and Kmer data. Thus, this disagreement between PhyloBayes and the other analyses seems to be an artifact due to the small dataset used to run this analysis.

The definition of three to four species instead of six is congruent with the species delimitation analysis using BPP, arguably the most accurate delimitation method based on the Multispecies Coalescent Model (Jackson et al., 2017; Rabiee and Mirarab, 2020). Although three out of the five analyzed subsets identified four species (green, violet, red, and blue + orange clades), the signal was unclear in the other two, which suggested gene flow between the red and the blue + orange clades, and that these three clades could be considered as a single species (i.e., the "cyan species"). In order to obtain a more detailed perspective of the haplotypes shared by the different clades, we tried to use the Admixture software. However, the limited size of the dataset and the skewed representation of each clade precluded this approach. Admixture is known to be unable to distinguish populations when the number of samples is poor, and even more importantly when the different populations are asymmetrically represented (Lawson et al., 2018), as would be the case here. Considering all results, we tentatively propose the definition of three species of cone snails inhabiting the Mediterranean Sea: green, violet, and cyan (the latter a combination of the red, orange, and blue clades in the mitochondrial trees), but note that a fourth one, represented by the red clade, should be subject to further consideration upon the generation of more data. Importantly, it is worth considering that analyses of species delimitation do not test for reproductive isolation but find clusters of more closely related populations. Hence, it would be important to confirm these results in future studies by testing whether these clades meet the biological concept of species.

The application of high-throughput sequencing techniques to the study of species diversity has allowed an unprecedented understanding of the speciation continuum, providing fine-detailed information about the structure of populations within species and leading to the discovery of many species previously overlooked (Irisarri et al., 2021; Janzen et al., 2017; Kornilios et al., 2020; Wagner et al., 2013). However, this level of resolution often blurs the limits of what should be considered within- and between-species structure. In fact, simulation studies have shown that delimitation methods based on the coalescent model frequently lead to the over-splitting of species in presence of strong geographic structure of genetically isolated populations (Chambers and Hillis, 2020; Sukumaran and Knowles, 2017). Moreover, these methods assume that the sampling effort has been sufficient to accurately represent the genetic variation among populations. When this assumption is not met, for instance when intermediate sites connecting different populations are missing, these models tend to define relatively isolated populations as separate species (Mason et al., 2020). In this study, we made a big effort to sample most of the extension of the Mediterranean Sea, and thus the obtained molecular evidence presented here seems to be robust enough to justify the definition of at least three species within the *L. ventricosus* species complex.

Although enough sequence divergence in both mitochondrial and nuclear markers is often sufficient to justify alone the definition of a new species, the common practice in an integrative taxonomy approach is to further support such molecular evidence with other independent sources of information, such as morphology, ecology, or biogeography (Hartop et al., 2022; Janzen et al., 2017; Parker et al., 2021). In an attempt to identify other characters that could be used to differentiate clades as species, we ran a morphometric analysis to compare shell morphology among the main clades recovered in phylogenomic analyses. The contrasting results between the cluster, without any *a priori* assumption of the number of species, and ANCOVA analyses, designed to test the differences between our predefined species, supported the putative existence of a complex of cryptic species. Without previous information, all

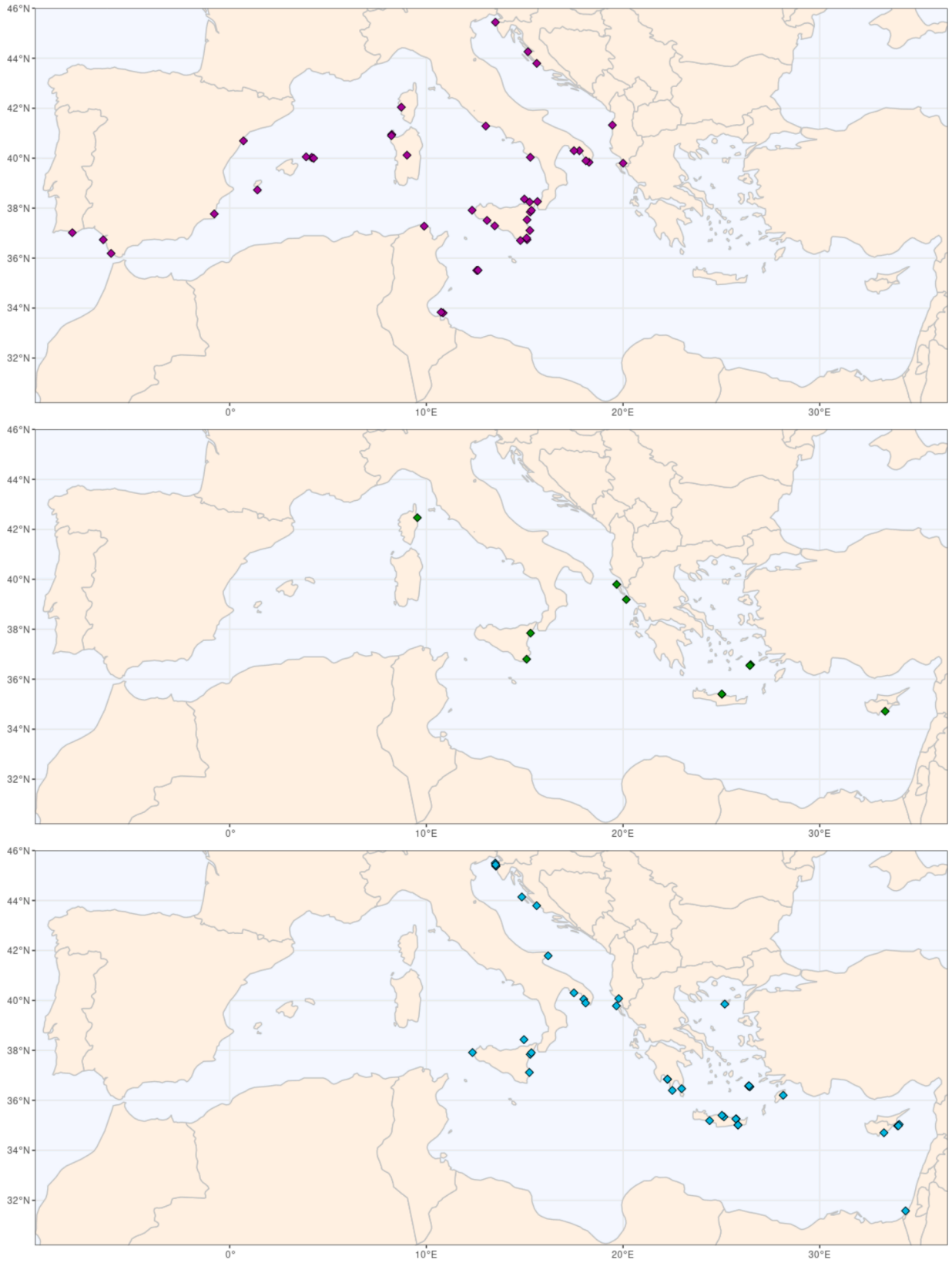

Fig. 6. Distribution of the three species proposed in this study. The maps only include the location of those individuals whose cox1 gene has been sequenced. The color coding follows that of Fig. 3. The maps were drawn on R using the ggplot2, sf, rnaturalearth, rgeos, and spData packages (Bivand et al., 2022; Bivand and Rundel, 2021; Pebesma, 2018; South, 2022; Wickham, 2016).

analyzed individuals were too similar to tell apart (data not shown), but statistically significant differences in the morphometric traits manifested when the molecular clades were considered. Importantly, although all the specimens of the green species were visibly (and statistically significant) smaller than those of the other two species, the limited sampling size advises to take this result with caution, and to test it in the future with an extended sampling. Finally, the discriminant function analysis confirmed the difficulty of telling the species of this complex apart using morphological data alone, and thus their cryptic nature. Despite the statistically significant differences observed between the violet and cyan species, the discrimination power of these parameters was only 70.2%, meaning that in order to confidently discriminate species, morphometric differences must be compared between specimens of the two species and/or used in combination with other traits, such as biogeography (see below). The differences are slightly more obvious when the mitochondrial clades are considered, as only three of the comparisons do not return statistically significant difference for any shell metric. Interestingly, significant differences are observed between red and both orange and blue, but not between the latter two, supporting the existence of a fourth species (red clade) as suggested by BPP. Nevertheless, these comparisons must be taken cautiously given the reduced number of samples per clade, and this alternative species delimitation hypothesis must be considered upon collecting more specimens for these clades.

Given the morphological uniformity of the shells of these cone snails, we also explored the use of ecological data for species identification within *L. ventricosus*. Previous studies have shown that ecological and physiological traits can also be used as diagnostic characters when morphological differences are subtle or even absent (Damm et al., 2010; Derycke et al., 2016). Specifically in cone snails, differences in venom composition where used, in combination with morphological and molecular data, to describe a new species (Puillandre et al., 2014b), and more recently the venoms of two complexes of cryptic species were shown to be markedly different (Himaya et al., 2022; Pardos-Blas et al., 2022). Likewise, another study exploring venom variation in a phylogenetic gradient has shown different degrees of variation in venom composition between intra- and inter-specific comparisons (Abalde et al., 2020). Therefore, we annotated all possible conotoxins in the assembled genomes using Scipio (Keller et al., 2008) and Spaln (Iwata and Gotoh, 2012), two programs designed to annotate genes in a genome using proteins from a closely related species as reference. However, either because conotoxins have fast evolutionary rates or because these genomes were highly fragmented, all proteins annotated were incomplete or contained different types of sequencing errors, and no identical conotoxins were shared between individuals (data not shown). Despite this unsuccessful attempt, we still believe that this approach is worth exploring in future studies. Another way to test this hypothesis would be sequencing several venom gland transcriptomes from the three newly identified species and run comparative analyses as in previous studies (e.g., Himaya et al., 2022; Pardos-Blas et al., 2022). Alternatively, target-capture sequencing has been shown to be a reliable tool for the annotation of conotoxins (Phuong and Mahardika, 2018), and it would also be a cost-effective approach to generate enough molecular information to test the validity of the clades/species here proposed.

The existence of three cryptic species within *L. ventricosus* could be further supported by their geographical distribution (Fig. 6). The violet and cyan species have almost complementary distributions, occupying the Western and Eastern regions of the Mediterranean Sea, respectively. Many other marine invertebrate and vertebrate species and the seagrass *Posidonia oceanica* share an East-West distribution within the Mediterranean Sea due to the presence of a well-known biogeographical barrier at the Siculo-Tunisian Strait that prevents gene flow since the last Pleistocene ice age (Arnaud-Haond et al., 2007; Borrero-Pérez et al., 2011; Moussa et al., 2022; Quéré et al., 2012). The green species is mostly sympatric with the cyan one. The three species present overlapping distributions around the Italian peninsula, and are found in the same localities from the East coast of the Adriatic Sea to the South coast of Italy and around Sicily. Further, the green species was found in the East coast of Corsica, also inhabited by the violet one. The green species shows the highest molecular divergence from the others in the complex, three times higher than that between violet and cyan in the mitogenome and two to three times higher in the nuclear genome. Hence, even if living in the same localities there seems to be a reproductive barrier between them, which should be tested in empirical studies. The case of the violet and cyan species could be subject to a more careful debate, because of the lower divergence, sister relationship in the trees, and contact zone, but this overlap covers less than one third of the total distribution of these species, they present a mitochondrial divergence of 3%, and morphometric differences have been found between them, supporting the hypothesis of them representing two different species.

The nuclear and mitochondrial chronograms inferred markedly different dates, with the nuclear tree displaying much older nodes than the mitochondrial one, which contradicts previous findings. Mitogenomes are known to evolve under higher substitution rates than nuclear genomes, normally resulting in older diversification estimates (Fulton and Strobeck, 2010; Zheng et al., 2011). In a comparison of the time estimates inferred by the two genomes, Zheng et al. (2011) proposed that the models would not capture correctly the substitution dynamics of mitochondrial genomes, leading to this difference. In our case, because of the size of the dataset, we originally analyzed all genes as a single unit, whereas mitochondrial genes were analyzed as independent partitions. Thus, it may be plausible that the models fit better the mitochondrial than the nuclear data, which could have caused the observed discrepancies. Furthermore, if the intron–exon boundaries were not well defined during the annotation step, genomic regions with strikingly different evolutionary dynamics would be present in the dataset, worsening the model misfit, and could also explain in part an overestimation of the divergence times. However, further testing with data subsets where genes were analyzed as independent partitions and a careful inspection of the alignments discarded this hypothesis. Because the same calibrations were used in the two analyses, and mitochondrial estimates are generally congruent with those inferred by BEAST in previous studies (Abalde et al., 2017a), we disregarded the use of MCMCtree and the selection of the calibration points as a cause of conflict in the analysis. Therefore, we could not find any methodological reason that would justify the observed differences between the nuclear and mitochondrial data, and the reason might simply be a discrepancy between the two genomes due to incomplete lineage sorting phenomena.

Despite these differences, some conclusions can be drawn from the two chronograms. From an evolutionary perspective, four main processes have been proposed to explain the existence of cryptic species: recent diversification, parallelism, convergence, and stasis (Struck et al., 2018), of which only two, recent diversification and stasis, can apply to *L. ventricosus*. The first refers to species that have diverged recently, without enough time to accumulate evident morphological differences, whereas the second refers to sister species that retain a similar morphology over extended periods of time. The other two, parallelism and convergence, usually describe complexes of non-sister species. The diversification event leading to the green species was estimated between 18.36 and 19.26 Mya, and to the other two species roughly 16.36 Mya (nuclear data) or between 5.10 and 6.59 Mya (mitochondrial data). Cone snails are known for their extraordinary divergence in shell coloration and patterns even in shorter periods, including many instances of high intra-specific variability and phenotypic plasticity (Abalde et al., 2017a; Abalde et al., 2017b; Pardos-Blas et al., 2019), suggesting that morphological stasis might be the evolutionary mechanisms behind the morphological homogeneity within the *L. ventricosus* species complex. Taken together, the results here presented indicate contrasting morphological evolutionary patterns within the family, and call for a careful re-evaluation of the shell as main morphological

character used in cone snail taxonomy.

5. Conclusions

The application of molecular techniques to taxonomic studies of cone snails has allowed the discovery of incorrectly delimited species boundaries, highlighting the importance of using such tools to define species and weight the reliability of the morphological characters thus far used during species delimitation (Abalde et al., 2017a; Duda et al., 2009b). Despite recent studies on cone snails ecology and evolution using a genomic perspective (Pardos-Blas et al., 2021; Phuong et al., 2019; Wang et al., 2017), and although Weese and Duda (2015) proposed a set of SNP data specifically selected for cone snails, molecular systematics of the group has been mostly limited thus far to the use of mitochondrial markers. Here, we have studied the *L. ventricosus* species complex using *cox1* barcodes, complete mitochondrial genomes, and genome-wide data from genome skims, observing discordance between mitochondrial and nuclear markers. Because of the different evolutionary rates and inheritance patterns driving the evolution of the two genomes, this pattern is not uncommon and may reflect gene flow, admixture, introgression, and incomplete lineage sorting processes (Bernardo et al., 2019; Firneno et al., 2020; Kimball et al., 2021; Toews and Brelsford, 2012). In *L. ventricosus*, mitochondrial data was paramount for the discovery of a potential species complex, later confirmed with nuclear phylogenomics, SNP analyses, and inferences of genomic distances, but these recognized a lower number of species. This study supports the use of traditional barcodes as first approach for identifying species in cone snails, but it calls for a careful evaluation of additional sources of sequence data in order to make fine-detail inferences of species boundaries, particularly when results involve important updates of cone snail systematics with conservation implications. This integrative approach should be extended to other studies of cone snail systematics, with special emphasis on species-rich clades (e.g., *Africonus*) and cases of cryptic species, whose classification is so far solely based on mitochondrial data.

CRediT authorship contribution statement

Samuel Abalde: Methodology, Formal analysis, Investigation, Writing – original draft, Data curation, Visualization. **Fabio Crocetta:** Methodology, Investigation, Writing – review & editing, Data curation, Funding acquisition. **Manuel J. Tenorio:** Methodology, Formal analysis, Investigation, Resources, Writing – review & editing, Data curation. **Salvatore D'Aniello:** Methodology, Investigation, Writing – review & editing, Data curation. **Giulia Fassio:** Investigation, Writing – review & editing. **Paula C. Rodríguez-Flores:** Investigation, Writing – review & editing. **Juan E. Uribe:** Investigation, Writing – review & editing. **Carlos M.L. Afonso:** Investigation, Writing – review & editing. **Marco Oliverio:** Conceptualization, Investigation, Resources, Writing – review & editing, Data curation. **Rafael Zardoya:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing, Data curation, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data important to understand the manuscript are attached as supplementary material. Raw data have been uploaded to GenBank. All the code has been shared through GitHub (release upon publication). All molecular data are available at GenBank under accession numbers: cox1 (ON951339 - ON951583), mitochondrial genomes (ON968966-ON968984, ON975007-ON975008), and low coverage genomes raw data (BioProject: PRJNA856832; accession numbers: SRR20206919 - SRR20206935). Additionally, detailed instructions (including the necessary files) to replicate the analyses and R scripts used to create the figures were uploaded to GitHub: https://github.com/saabalde/2023_Lventricosus_Species_Complex.

Acknowledgements

We are grateful to Ángel Luque, José Templado, Anabel Perdices, Annie Machordorm, David Buckley, Violeta López, and Iván Acevedo for sharing with us cone snails captured during their own sampling campaigns. Likewise, we thank María Ángeles Carpintero-Carcedo from Spain and Giuseppe Colamonaco (†), Emilio Riginella, Francesca Strano, Valentina Bigini, and Valeria Russini from Italy for kindly providing specimens and sequences necessary to complete this study. Samuel Abalde thanks María Recuerda and Silvia Perea for their advice during SNP analyses, Yi Yang and José Ramón Pardos for their help in the molecular lab, and Sandra Álvarez-Carretero for her insightful advice to run and interpret the MCMCtree analyses. We thank the staff at MNHN for providing access to samples and unpublished sequences of Tunisia specimens collected during the MNHN workshop "GABES 2012-2013", and Philippe Maestrati, Chrifa Aissaoui, Emmanuel Vassard, Gianni Spada, and Jean-Pierre Miquel, for their help in sampling. 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 all analyses of NGS data and phylogenetic inference. This work was mainly supported by the Spanish Ministry of Science, Industry, and Innovation (CGL2016-75255-C2-1-P (AEI/FEDER, UE) and PID2019-103947GB-C22/AEI/10.13039/501100011033 to RZ; BES-2014-069575 and EEBB-I-16-11775 to SA), the Comunidad de Madrid (2019-T2/AMB-13166 to JEU), and partially complemented by the European Union (TD1209 ALIEN Challenge to Rhodes and Crete COST Actions to FC).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107838.

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 Evol. Biol. 17, 231. https://doi.org/10.1186/s12862-017-1069-x.

Abalde, S., Tenorio, M.J., Afonso, C.M.L., Zardoya, R., 2017b. Mitogenomic phylogeny of cone snails endemic to Senegal. Mol. Phylogenet. Evol. 112, 79–87. https://doi.org/10.1016/j.ympev.2017.04.020.

Abalde, S., Tenorio, M.J., Afonso, C.M.L., Zardoya, R., 2020. Comparative transcriptomics of the venoms of continental and insular radiations of West African cones. Proc. R. Soc. B Biol. Sci. 287, 20200794. https://doi.org/10.1098/rspb.2020.0794.

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

Aissaoui, C., Puillandre, N., Bouchet, P., Fassio, G., Modica, M.V., Oliverio, M., 2016. Cryptic diversity in Mediterranean gastropods of the genus aplus (Neogastropoda: Buccinidae). Sci. Mar. 80, 521–533. https://doi.org/10.3989/scimar.04422.12A.

Alonge, M., Lebeigle, L., Kirsche, M., Aganezov, S., Wang, X., Lippman, Z.B., Schatz, M. C., Soyk, S., 2021. Automated assembly scaffolding elevates a new tomato system for high-throughput genome editing. bioRxiv.

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van De Vliet, M.S., Alberto, F., Procaccini, G., Duarte, C.M., Serrão, E.A., 2007. Vicariance patterns in the Mediterranean Sea: east-west cleavage and low dispersal in the endemic seagrass Posidonia oceanica. J. Biogeogr. 34, 963–976. https://doi.org/10.1111/j.1365-2699.2006.01671.x.

Atherton, S., Jondelius, U., 2021. Phylogenetic assessment and systematic revision of the acoel family Isodiametridae. Zool. J. Linn. Soc. 1–25.

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

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V.,

Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.
Bernardo, P.H., Sánchez-Ramírez, S., Sánchez-Pacheco, S.J., Álvarez-Castañeda, S.T., Aguilera-Miller, E.F., Mendez-de la Cruz, F.R., Murphy, R.W., 2019. Extreme mito-nuclear discordance in a peninsular lizard: the role of drift, selection, and climate. Heredity (Edinb). 123, 359–370. https://doi.org/10.1038/s41437-019-0204-4.
Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22, 148–155. https://doi.org/10.1016/j.tree.2006.11.004.
Bivand, R., Rundel, C., 2021. rgeos: Interface to Geometry Engine - Open Source ('GEOS').
Bivand, R., Nowosad, J., Lovelace, R., 2022. spData: Datasets for Spatial Analysis.
Borrero-Pérez, G.H., González-Wangüemert, M., Marcos, C., Pérez-Ruzafa, A., 2011. Phylogeography of the Atlanto-Mediterranean sea cucumber Holothuria (Holothuria) mammata: the combined effects of historical processes and current oceanographical pattern. Mol. Ecol. 20, 1964–1975. https://doi.org/10.1111/j.1365-294X.2011.05068.x.
Calvo, M., Templado, J., Oliverio, M., MacHordom, A., 2009. Hidden Mediterranean biodiversity: Molecular evidence for a cryptic species complex within the reef building vermetid gastropod Dendropoma petraeum (Mollusca: Caenogastropoda). Biol. J. Linn. Soc. 96, 898–912. https://doi.org/10.1111/j.1095-8312.2008.01167.x.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.
Chambers, E.A., Hillis, D.M., 2020. The multispecies coalescent over-splits species in the case of geographically widespread taxa. Syst. Biol. 69, 184–193. https://doi.org/10.1093/sysbio/syz042.
Chang, Z.Y., Liew, T.S., 2021. A molecular phylogeny of Geotrochus and Trochomorpha species (Gastropoda: Trochomorphidae) in Sabah, Malaysia reveals convergent evolution of shell morphology driven by environmental influences. PeerJ 9, 1–23. https://doi.org/10.7717/peerj.10526.
Coissac, E., Hollingsworth, P.M., Lavergne, S., Taberlet, P., 2016. From barcodes to genomes: extending the concept of DNA barcoding. Mol. Ecol. 25, 1423–1428. https://doi.org/10.1111/mec.13549.
Corral-Lou, A., Perea, S., Perdices, A., Doadrio, I., 2022. Quaternary geomorphological and climatic changes associated with the diversification of Iberian freshwater fishes: the case of the genus Cobitis (Cypriniformes, Cobitidae). Ecol. Evol. 12, 1–22. https://doi.org/10.1002/ece3.8635.
Criscuolo, A., Gribaldo, S., 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210.
Damm, S., Schierwater, B., Hadrys, H., 2010. An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. Mol. Ecol. 19, 3881–3893. https://doi.org/10.1111/j.1365-294X.2010.04720.x.
Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10, 1–4. https://doi.org/10.1093/gigascience/giab008.
Derycke, S., De Meester, N., Rigaux, A., Creer, S., Bik, H., Thomas, W.K., Moens, T., 2016. Coexisting cryptic species of the Litoditis marina complex (Nematoda) show differential resource use and have distinct microbiomes with high intraspecific variability. Mol. Ecol. 25, 2093–2110. https://doi.org/10.1111/mec.13597.
Dierckxsens, N., Mardulyn, P., Smits, G., 2017. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucl. Acids Res. 45 https://doi.org/10.1093/nar/gkw955.
Donald, K.M., Spencer, H.G., 2016. Phylogeographic patterns in New Zealand and temperate Australian cantharidines (Mollusca: Gastropoda: Trochidae: Cantharidinae): trans-Tasman divergences are ancient. Mol. Phylogenet. Evol. 100, 333–344. https://doi.org/10.1016/j.ympev.2016.04.029.
Donath, A., Jühling, F., Al-Arab, M., Bernhart, S.H., Reinhardt, F., Stadler, P.F., Middendorf, M., Bernt, M., 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucl. Acids Res. 47, 10543–10552. https://doi.org/10.1093/nar/gkz833.
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. Mol. Phylogen. Evol. 49, 867–876. https://doi.org/10.1016/j.ympev.2008.08.009.
Duda, T.F., 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., Kohn, A.J., 2005. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus Conus. Mol. Phylogen. Evol. 34, 257–272. https://doi.org/10.1016/j.ympev.2004.09.012.
Duda, T.F., Kohn, A.J., Matheny, A.M., 2009b. Cryptic species differentiated in Conus ebraeus, a widespread tropical marine gastropod. Biol. Bull. 217, 292–305, 217/3/292 [pii].
Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nat. Commun. 5, 3521.
Emms, D.M., Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20, 238.
Fernández-Álvarez, F.Á., Taite, M., Vecchione, M., Villanueva, R., Allcock, A.L., 2022. A phylogenomic look into the systematics of oceanic squids (order Oegopsida). Zool. J. Linn. Soc. 194, 1212–1235. https://doi.org/10.1093/zoolinnean/zlab069.
Firneno, T.J., O'Neill, J.R., Portik, D.M., Emery, A.H., Townsend, J.H., Fujita, M.K., 2020. Finding complexity in complexes: assessing the causes of mitonuclear discordance in a problematic species complex of Mesoamerican toads. Mol. Ecol. 29, 3543–3559. https://doi.org/10.1111/mec.15496.
Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486.
Flouri, T., Rannala, B., Yang, Z., 2020. A tutorial on the use of BPP for species tree estimation and species delimitation. In: Scornavacca, C., Delsuc, F., Galtier, N. (Eds.), Phylogenetics in the Genomic Era. No Commercial Publisher, pp. 5.6:1–5.6:16.
Folmer, O., Black, M., Hoeh, W., 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.
Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565.
Fulton, T.L., Strobeck, C., 2010. Multiple fossil calibrations, nuclear loci and mitochondrial genomes provide new insight into biogeography and divergence timing for true seals (Phocidae, Pinnipedia). J. Biogeogr. 37, 814–829. https://doi.org/10.1111/j.1365-2699.2010.02271.x.
Gabrielli, M., Nabholz, B., Leroy, T., Milá, B., Thébaud, C., 2020. Within-island diversification in a passerine bird. Proc. R. Soc. B Biol. Sci. 287 https://doi.org/10.1098/rspb.2019.2999.
Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. arXiv 1–9.
Garrison, E., Kronenberg, Z.N., Dawson, E.T., Pedersen, B.S., Prins, P., 2021. Vcflib and tools for processing the VCF variant call format. bioRxiv 2021.05.21.445151.
Glez-Pena, D., Gomez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., Posada, D., 2010. ALTER: program-oriented conversion of DNA and protein alignments. Nucl. Acids Res. 38, W14–W18. https://doi.org/10.1093/nar/gkq321.
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. Syst. Biol. 59, 307–321.
Hartop, E., Srivathsan, A., Ronquist, F., Meier, R., 2022. Large-scale Integrative Taxonomy (LIT): resolving the data conundrum for dark taxa. Syst. Biol. syac033.
Himaya, S.W.A., Arkhipov, A., Yum, W.Y., Lewis, R.J., 2022. Comparative venomics of C. flavidus and C. frigidus and closely related vermivorous cone snails. Mar. Drugs 20, 209.
Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., Pitteloud, C., Dincă, V., Alvarez, N., Vila, R., 2019. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris. Mol. Ecol. 28, 3857–3868. https://doi.org/10.1111/mec.15153.
Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.
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. J. Geophys. Res. 113 https://doi.org/10.1029/2007jb005339.
Institute, B., 2019. Picard Toolkit. Broad Institute, Github Repository: https://github.com/broadinstitute/picard.
Irisarri, I., Darienko, T., Pröschold, T., Fürst-Jansen, J.M.R., Jamy, M., De Vries, J., 2021. Unexpected cryptic species among streptophyte algae most distant to land plants. Proc. R. Soc. B Biol. Sci. 288 https://doi.org/10.1098/rspb.2021.2168.
Iwata, H., Gotoh, O., 2012. Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. Nucl. Acids Res. 40, 1–9. https://doi.org/10.1093/nar/gks708.
Jackson, N.D., Carstens, B.C., Morales, A.E., O'Meara, B.C., 2017. Species delimitation with gene flow. Syst. Biol. 66, 799–812. https://doi.org/10.1093/sysbio/syw117.
Janzen, D.H., Burns, J.M., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert, P.D.N., Grishin, N.V., 2017. Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology. Proc. Natl. Acad. Sci. 114, 8313–8318. https://doi.org/10.1073/pnas.1621504114.
Johannesson, K., Panova, M., Kemppainen, P., André, C., Rolan-Alvarez, E., Butlin, R.K., 2010. Repeated evolution of reproductive isolation in a marine snail: unveiling mechanisms of speciation. Philos. Trans. R. Soc. B Biol. Sci. 365, 1735–1747. https://doi.org/10.1098/rstb.2009.0256.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.
Karanovic, T., Djurakic, M., Eberhard, S.M., 2016. Cryptic species or inadequate taxonomy? Implementation of 2D geometric morphometrics based on integumental organs as landmarks for delimitation and description of copepod taxa. Syst. Biol. 65, 304–327. https://doi.org/10.1093/sysbio/syv088.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Keller, O., Odronitz, F., Stanke, M., Kollmar, M., Waack, S., 2008. Scipio: Using protein sequences to determine the precise exon/intron structures of genes and their orthologs in closely related species. BMC Bioinformatics 9, 1–12. https://doi.org/10.1186/1471-2105-9-278.

Kimball, R.T., Guido, M., Hosner, P.A., Braun, E.L., 2021. When good mitochondria go bad: cyto-nuclear discordance in landfowl (Aves: Galliformes). Gene 801, 145841. https://doi.org/10.1016/j.gene.2021.145841.
Kornilios, P., Thanou, E., Lymberakis, P., Ilgaz, Ç., Kumlutaş, Y., Leaché, A., 2020. A phylogenomic resolution for the taxonomy of Aegean green lizards. Zool. Scr. 49, 14–27. https://doi.org/10.1111/zsc.12385.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.
Lartillot, N., Rodrigue, N., Stubbs, D., Richer, J., 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst. Biol. 62, 611–615. https://doi.org/10.5061/dryad.c459h).
Lawson, D.J., van Dorp, L., Falush, D., 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nat. Commun. 9, 1–11. https://doi.org/10.1038/s41467-018-05257-7.
Lefort, V., Desper, R., Gascuel, O., 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. Mol. Biol. Evol. 32, 2798–2800. https://doi.org/10.1093/molbev/msv150.
Lemer, S., Buge, B., Bemis, A., Giribet, G., 2014. First molecular phylogeny of the circumtropical bivalve family Pinnidae (Mollusca, Bivalvia): Evidence for high levels of cryptic species diversity. Mol. Phylogenet. Evol. 75, 11–23. https://doi.org/10.1016/j.ympev.2014.02.008.
Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925. https://doi.org/10.1080/106351501753462876.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li, H., Wren, J., 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 30, 2843–2851. https://doi.org/10.1093/bioinformatics/btu356.
Malukiewicz, J., Cartwright, R.A., Dergam, J.A., Igayara, C.S., Nicola, P.A., Pereira, L.M. C., Ruiz-Miranda, C.R., Stone, A.C., Silva, D.L., Silva, F. de F.R. da, Varsani, A., Walter, L., Wilson, M.A., Zinner, D., Roos, C., 2021. Genomic skimming and nanopore sequencing uncover cryptic hybridization in one of world's most threatened primates. Sci. Rep. 11, 1–10. https://doi.org/10.1038/s41598-021-96404-6.
Mason, N.A., Fletcher, N.K., Gill, B.A., Funk, W.C., Zamudio, K.R., 2020. Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. Syst. Biodivers. 18, 269–280. https://doi.org/10.1080/14772000.2020.1730475.
Menting, J.G., Gajewiak, J., MacRaild, C.A., Chou, D.H., Disotuar, M.M., Smith, N.A., Miller, C., Erchegyi, J., Rivier, J.E., Olivera, B.M., Forbes, B.E., Smith, B.J., Norton, R.S., Safavi-Hemami, H., Lawrence, M.C., 2016. A minimized human insulin-receptor-binding motif revealed in a Conus geographus venom insulin. Nat. Struct. Mol. Biol. 23, 916–920. https://doi.org/10.1038/nsmb.3292.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., Teeling, E., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
Moles, J., Derkarabetian, S., Schiaparelli, S., Schrödl, M., Troncoso, J.S., Wilson, N.G., Giribet, G., 2021. An approach using ddRADseq and machine learning for understanding speciation in Antarctic Antarctophilinidae gastropods. Sci. Rep. 11, 1–14. https://doi.org/10.1038/s41598-021-87244-5.
Mongiardino Koch, N., 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. Mol. Biol. Evol. 38, 4025–4038. https://doi.org/10.1093/molbev/msab151.
Moussa, M., Choulak, S., Rhouma-Chatti, S., Chatti, N., Said, K., 2022. First insight of genetic diversity, phylogeographic relationships, and population structure of marine sponge Chondrosia reniformis from the eastern and western Mediterranean coasts of Tunisia. Ecol. Evol. 12, 1–13. https://doi.org/10.1002/ece3.8494.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.
Ortiz, E.M., 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. 10.5281/zenodo.2540861.
Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., Chenuil, A., Destombe, C., Forcioli, D., Valero, M., Viard, F., Samadi, S., 2015. Species are hypotheses: avoid connectivity assessments based on pillars of sand. Mol. Ecol. 24, 525–544. https://doi.org/10.1111/mec.13048.
Pardos-Blas, J.R., Irisarri, I., Abalde, S., Tenorio, M.J., Zardoya, R., 2019. Conotoxin diversity in the venom gland transcriptome of the magician's cone. Pionocon. magus. Mar. Drugs 17. https://doi.org/10.3390/md17100553.
Pardos-Blas, J.R., Irisarri, I., Abalde, S., Afonso, C.M.L., Tenorio, M.J., Zardoya, R., 2021. The genome of the venomous snail Lautoconus ventricosus sheds light on the origin of conotoxin diversity. Gigascience 10, 1–15. https://doi.org/10.1093/gigascience/giab037.
Pardos-Blas, J.R., Tenorio, M.J., Galindo, J.C.G., Zardoya, R., 2022. Comparative venomics of the cryptic cone snail species virroconus ebraeus and virroconus judaeus. Mar. Drugs 20. https://doi.org/10.3390/md20020149.
Parker, E., Dornburg, A., Struthers, C.D., Jones, C.D., Near, T.J., 2021. Phylogenomic species delimitation dramatically reduces species diversity in an antarctic adaptive radiation. Syst. Biol. 1–20. https://doi.org/10.1093/sysbio/syab057.
Pebesma, E., 2018. Simple features for R: standardized support for spatial vector data. R. J. 10, 439–446. https://doi.org/10.32614/RJ-2018-009.
Peng, C., Huang, Y., Bian, C., Li, J., Liu, J., Zhang, K., You, X., Lin, Z., He, Y., Chen, J., Lv, Y., Ruan, Z., Zhang, X., Yi, Y., Li, Y., Lin, X., Gu, R., Xu, J., Yang, J., Fan, C., Yao, G., Chen, J.S., Jiang, H., Gao, B., Shi, Q., 2021. The first Conus genome assembly reveals a primary genetic central dogma of conopeptides in C. betulinus. Cell Discov. 7 https://doi.org/10.1038/s41421-021-00244-7.
Perea, S., Sousa-Santos, C., Robalo, J., Doadrio, I., 2020. Multilocus phylogeny and systematics of Iberian endemic Squalius (Actinopterygii, Leuciscidae). Zool. Scr. 49, 440–457. https://doi.org/10.1111/zsc.12420.
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. https://doi.org/10.1371/journal.pone.0083353.
Pfenninger, M., Schwenk, K., 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol. Biol. 7, 1–6. https://doi.org/10.1186/1471-2148-7-121.
Phuong, M.A., Alfaro, M.E., Mahardika, G.N., Marwoto, R.M., Prabowo, R.E., von Rintelen, T., Vogt, P.W.H., Hendricks, J.R., Puillandre, N., 2019. Lack of signal for the impact of conotoxin gene diversity on speciation rates in cone snails. Syst. Biol. 68, 781–796. https://doi.org/10.1093/sysbio/syz016.
Phuong, M.A., Mahardika, G.N., 2018. Targeted sequencing of venom genes from cone snail genomes improves understanding of conotoxin molecular evolution. Mol. Biol Evol.
Pin, M., Leung Tack, K.D., 1995. The Cones of Senegal. Conchiglia.
Ponder, W.F., Lindberg, D.R., Ponder, J.M., 2020. Biology and Evolution of the Mollusca, vol. I. Taylor and Francis Group.
Puillandre, N., Bouchet, P., Duda, T.F., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014a. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Mol. Phylogenet. Evol. 78, 290–303. https://doi.org/10.1016/j.ympev.2014.05.023.
Puillandre, N., Stöcklin, R., Favreau, P., Bianchi, E., Perret, F., Rivasseau, A., Limpalaër, L., Monnier, E., Bouchet, P., 2014b. When everything converges: Integrative taxonomy with shell, DNA and venomic data reveals Conus conco, a new species of cone snails (Gastropoda: Conoidea). Mol. Phylogenet. Evol. 80, 186–192. https://doi.org/10.1016/j.ympev.2014.06.024.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795.
Puttick, M.N., 2019. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. Bioinformatics 35, 5321–5322. https://doi.org/10.1093/bioinformatics/btz554.
Quéré, N., Desmarais, E., Tsigenopoulos, C.S., Belkhir, K., Bonhomme, F., Guinand, B., 2012. Gene flow at major transitional areas in sea bass (Dicentrarchus labrax) and the possible emergence of a hybrid swarm. Ecol. Evol. 2, 3061–3078. https://doi.org/10.1002/ece3.406.
Rabiee, M., Mirarab, S., 2020. SODA: Multi-locus species delimitation using quartet frequencies. Bioinformatics 36, 5623–5631. https://doi.org/10.1093/bioinformatics/btaa1010.
Rambaut, A., Drummond, A.J., 2007. Tracer v1. 4: MCMC Trace Analyses Tool.
Rannala, B., Yang, Z., 2007. Inferring speciation tunes under an episodic molecular clock. Syst. Biol. 56, 453–466. https://doi.org/10.1080/10635150701420643.
Recuerda, M., Illera, J.C., Blanco, G., Zardoya, R., Milá, B., 2021. Sequential colonization of oceanic archipelagos led to a species-level radiation in the common chaffinch complex (Aves: Fringilla coelebs). Mol. Phylogenet. Evol. 164 https://doi.org/10.1016/j.ympev.2021.107291.
Recuero, E., Rodríguez-Flores, P.C., 2020. A new Mediterranean species of Dolistenus (Diplopoda, Platydesmida, Andrognathidae), with an updated key for the genus and the first contribution for a barcode database of European Platydesmida. Zootaxa 4718, 123–133. https://doi.org/10.11646/zootaxa.4718.1.10.
Ren, Z., Harris, A.J., Dikow, R.B., Ma, E., Zhong, Y., Wen, J., 2017. Another look at the phylogenetic relationships and intercontinental biogeography of eastern Asian – North American Rhus gall aphids (Hemiptera: Aphididae: Eriosomatinae): evidence from mitogenome sequences via genome skimming. Mol. Phylogen. Evol. 117, 102–110. https://doi.org/10.1016/j.ympev.2017.05.017.
Rodríguez-Flores, P.C., MacPherson, E., MacHordom, A., 2021. Revision of the squat lobsters of the genus Phylladiorhynchus Baba, 1969 (Crustacea, Decapoda, Galatheidae) with the description of 41 new species. Zootaxa. 10.11646/zootaxa.5008.1.1.
Rokas, A., Carroll, S.B., 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22, 1337–1344. https://doi.org/10.1093/molbev/msi121.
Roux, C., Fraisse, 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.
Sanchez, G., Fernández-Álvarez, F.Á., Taite, M., Sugimoto, C., Jolly, J., Simakov, O., Marlétaz, F., Allcock, L., Rokhsar, D.S., 2021. Phylogenomics illuminates the evolution of bobtail and bottletail squid (order Sepiolida). Commun. Biol. 4, 1–9. https://doi.org/10.1038/s42003-021-02348-y.
Sarmashghi, S., Bohmann, K., Gilbert, T.P., M., Bafna, V., Mirarab, S.,, 2017. Assembly-free and alignment-free sample identification using genome skims. Genome Biol. 20, 1–20. https://doi.org/10.1101/230409.
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. Ann. Stat. 6, 461–464.
Seppey, M., Manni, M., Zdobnov, E.M., 2019. BUSCO: assessing genome assembly and annotation completeness. In: M., K. (Ed.), Gene Prediction. Methods in Molecular Biology. Humana Press, New York.

Shults, P., Hopken, M., Eyer, P.A., Blumenfeld, A., Mateos, M., Cohnstaedt, L.W., Vargo, E.L., 2022. Species delimitation and mitonuclear discordance within a species complex of biting midges. Sci. Rep. 12, 1–13. https://doi.org/10.1038/s41598-022-05856-x.

South, A., 2022. rnaturalearth: World Map Data from Natural Earth.

Straub, S.C.K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., Liston, A., 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am. J. Bot. 99, 349–364. https://doi.org/10.3732/ajb.1100335.

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.H., Liow, L.H., Nowak, M.D., Stedje, B., Bachmann, L., Dimitrov, D., 2018. Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33, 153–163. https://doi.org/10.1016/j.tree.2017.11.007.

Sukumaran, J., Knowles, L.L., 2017. Multispecies coalescent delimits structure, not species. Proc. Natl. Acad. Sci. U.S.A. 114, 1607–1611. https://doi.org/10.1073/pnas.1607921114.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 38, 3022–3027. https://doi.org/10.1093/molbev/msab120.

Tenorio, M.J., Abalde, S., Pardos-Blas, J.R., Zardoya, R., 2020. Taxonomic revision of West African cone snails (Gastropoda: Conidae) based upon mitogenomic studies: implications for conservation. Eur. J. Taxon. 663, 1–89. https://doi.org/10.5852/ejt.2020.663.

The GIMP Development Team, 2019. GIMP.

Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x.

Tucker, J.K., Tenorio, M.J., 2009. Systematic Classification of Recent and Fossil Conoidean Gastropods, with Keys to the Genera of Cone Shells. ConchBooks.

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. https://doi.org/10.1016/j.ympev.2016.10.008.

Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22, 787–798. https://doi.org/10.1111/mec.12023.

Wang, S., Zhang, J., Jiao, W., Li, J., Xun, X., Sun, Y., Guo, X., Huan, P., Dong, B., Zhang, L., Hu, X., Sun, X., Wang, J., Zhao, C., Wang, Y., Wang, D., Huang, X., Wang, R., Lv, J., Li, Y., Zhang, Z., Liu, B., Lu, W., Hui, Y., Liang, J., Zhou, Z., Hou, R., Li, X., Liu, Y., Li, H., Ning, X., Lin, Y., Zhao, L., Xing, Q., Dou, J., Li, Y., Mao, J., Guo, H., Dou, H., Li, T., Mu, C., Jiang, W., Fu, Q., Fu, X., Miao, Y., Liu, J., Yu, Q., Li, R., Liao, H., Li, X., Kong, Y., Jiang, Z., Chourrout, D., Li, R., Bao, Z., 2017. Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat. Ecol. Evol. 1, 120. https://doi.org/10.1038/s41559-017-0120.

Weese, D.A., Duda, T.F., 2015. Transcriptomic resources for three populations of Conus miliaris (Mollusca: Conidae) from Easter Island, American Samoa and Guam. Mol. Ecol. Resour. 15, 228–229.

Whelan, N.V., Halanych, K.M., 2017. Who let the CAT out of the bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. Syst. Biol. 66, 232–255. https://doi.org/10.1093/sysbio/syw084.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis.

Wood, A.W., Duda, T.F., 2021. Reticulate evolution in Conidae: evidence of nuclear and mitochondrial introgression. Mol. Phylogenet. Evol. 161, 107182 https://doi.org/10.1016/j.ympev.2021.107182.

WoRMS Editorial Board, 2023. World Register of Marine Species [WWW Document]. Accessed 4.11.22. https://www.marinespecies.org.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591. https://doi.org/10.1093/molbev/msm088.

Yang, Y., Li, Q., Kong, L., Yu, H., 2018. Comparative mitogenomic analysis reveals cryptic species in Reticunassa festiva (Neogastropoda: Nassariidae). Gene 662, 88–96. https://doi.org/10.1016/j.gene.2018.04.001.

Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. U.S.A. 107, 9264–9269. https://doi.org/10.1073/pnas.0913022107.

Zaharias, P., Pante, E., Gey, D., Fedosov, A.E., Puillandre, N., 2020. Data, time and money: evaluating the best compromise for inferring molecular phylogenies of non-model animal taxa. Mol. Phylogenet. Evol. 142, 106660 https://doi.org/10.1016/j.ympev.2019.106660.

Zhang, C., Sayyari, E., Mirarab, S., 2017. ASTRAL-III: Increased Scalability and Impacts of Contracting Low Support Branches, in: J., M. (Ed.), Comparative Genomics. RECOMB-CG 2017. Lecture Notes in Computer Science. Springer, Cham.

Zheng, Y., Peng, R., Kuro-O, M., Zeng, X., 2011. Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of salamanders (Order Caudata). Mol. Biol. Evol. 28, 2521–2535. https://doi.org/10.1093/molbev/msr072.