Molecular Phylogenetics and Evolution 222 (2026) 108630

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Limits of exon capture efficiency for phylogenomics using nudibranch gastropods (Mollusca: Heterobranchia) as a case study

Sofía Paz-Sedano [a,b,*], Ángel Valdés [c], Carla C Stout [c], Kendall Feliciano [c], Sandra Muro [c], Nerida G. Wilson [d,e], Kara K.S. Layton [f], Jessica A. Goodheart [b]

[a] Department of Ecology and Animal Biology, Faculty of Marine Sciences, University of Vigo, Vigo, Galicia, Spain
[b] Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA
[c] California State Polytechnic University, Pomona, CA, USA
[d] School of Biological Sciences, University of Western Australia, Crawley, WA, Australia
[e] Collections and Research, Western Australian Museum, Welshpool, WA, Australia
[f] Department of Biology, University of Toronto Mississauga, Mississauga, ON, Canada

ARTICLE INFO

Keywords:
Next generation sequencing
Exon capture
Phylogenomics
Cladobranchia
Doridina

ABSTRACT

Target capture is an effective tool for phylogenomic reconstruction of evolutionary histories. This sequencing methodology is cost-effective and allows the use of older or poorly preserved specimens. One important consideration while applying target capture is the rate of divergence of the group under study. The success of this approach depends on the distance of the bait set with the target taxon, affecting the sequencing efficiency and the ability to extract the target loci. However, the threshold distance at which the efficiency decreases is still unclear, varying between studies and taxonomic groups. In nudibranch mollusks, an exon capture bait set was designed for the suborder Doridina. However, the usefulness of this bait set in other clades has not been explored. We investigate the efficiency and taxonomic limits of target capture sequencing in evolutionarily distant taxa, testing the capacity of inferring the evolutionary relationships throughout the order Nudibranchia. Seventy-nine species across two diverse suborders and 12 superfamilies were sequenced using an exon capture bait set designed for Doridina. Although species of the suborder Cladobranchia were successfully sequenced, a decrease in the number of captured genes is observed, showing that the capture efficiency significantly decreases after a threshold bait-to-target DNA distance of 18%. Even so, the results obtained from phylogenetic analyses shed light on relationships among nudibranch taxa, addressing questions that have remained controversial due to the lack of support in previous analyses, such as the position of the genus *Doridoxa* within Doridina, supporting Phyllidioidea as monophyletic or the paraphyletic nature of Chromodoridoidea.

1. Introduction

The development of Next Generation Sequencing (NGS) has allowed the sequencing of hundreds or thousands of loci obtained from different sources, such as transcriptomes, genomes, target genes, or methylated sequences (Kulski, 2016; Lozano-Fernandez, 2022). These progressively evolving techniques have proven to be a powerful approach in reconstructing phylogenetic relationships, improving the resolution of phylogeny by increasing the number of informative characters (Cheon et al., 2020). However, not all techniques provide the same information, and the reliability of a phylogenetic tree depends on the identification of quality orthologous genes, as well as the ability to sequence the same loci among different samples and taxa (Delsuc et al., 2005; Bragg et al., 2016).

Target specific subsets, also known as target enrichment, is a cost-effective approach, where genomic regions are selected prior to NGS (Mamanova et al., 2010). It has proven to be a convenient technique when funds for collecting fresh specimens are low or when older museum specimens are primarily used (Mamanova et al., 2010; McCormack et al., 2016; Erickson et al., 2021). The most popular targeted approaches for the study of evolutionary relationships have been ultra-conserved elements (UCEs) and exons, due to their comparability across taxa (Andermann et al., 2020; Cowman et al., 2020). With appropriate reference taxa during the bait set design, both UCEs and

* Corresponding author.
E-mail address: sofia.paz@uvigo.gal (S. Paz-Sedano).

https://doi.org/10.1016/j.ympev.2026.108630
Received 21 October 2025; Received in revised form 15 April 2026; Accepted 18 April 2026
Available online 19 April 2026
1055-7903/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

exons have shown to be successful in resolving deep (order level) and shallow (genus level) relationships in different groups of organisms (Portik et al., 2016; Hughes et al., 2018, 2021; Rincon-Sandoval et al., 2019; Wang et al., 2022; González-Delgado et al., 2024).

When using target capture sequencing, one crucial consideration is the taxa of interest, with the aim of selecting target loci with appropriate evolutionary rates and capturing conserved loci within the study group that contain sufficient genetic variation to infer evolutionary relationships (Andermann et al., 2020). Moreover, the greater the evolutionary divergence between the taxa of interest and the probe bait set, the lower the capture success and ultimately the number of recovered loci. Some studies suggested that this bait-to-target DNA sequence distance threshold at which capture success significantly declines could be around 10% (Bragg et al., 2016), although other research sets this threshold at 15–20% (Mayer et al., 2016) or up to 39% (Li et al., 2013). To mitigate this distance and effectively sequence a divergent group of taxa, the probe bait set is designed to include baits from multiple reference species for the same target locus, capturing in-group nucleotide sequence divergence. However, an increase in the number of baits could lead to an increase in the final cost of the bait set (Mayer et al., 2016). Consequently, the choice of the distance threshold from the reference species for the bait set design, and the bait set from the target species, becomes crucial for cost effective and successful target capture sequencing.

A particularly interesting system to apply NGS techniques to is Nudibranchia (Mollusca: Gastropoda: Heterobranchia). Although these charismatic organisms are important for numerous fields of science, including chemical ecology and evolution (Cimino and Ghiselin, 1999; da Cruz et al., 2011; Dean and Prinsep, 2017), neuroscience and behavioral evolution (Lillvis et al., 2012; Katz, 2016; Prieto-Baños and Layton, 2025), toxicology and pharmacology (Fuhrman et al., 1979; Ruiz-Torres et al., 2021), etc., their backbone phylogenetic relationships remain poorly resolved, hampering further research into the evolution of key traits. Various attempts to resolve the position of key groups using morphology and Sanger sequencing data have failed due to rampant homoplasy (Valdés, 2004) and gene saturation (Johnson and Gosliner, 2012; Hallas et al., 2017), respectively. At the same time, NGS techniques have been shown to be successful in resolving phylogenetic relationships of large nudibranch clades (Goodheart et al., 2017; Pabst and Kocot, 2018; Layton et al., 2020; Paz-Sedano et al., 2024). Although the taxonomic status of Nudibranchia has recently been challenged by Korshunova et al. (2025), who proposed restricting the name Nudibranchia exclusively to the suborder Cladobranchia, and using the order name Doridida for the suborder Doridina. This proposal was based on the classification of suborders by Blainville (1814), which has been extensively discussed in the literature (e.g., Férussac,1822; Alder J., Hancock A. (1845)-1855; Bergh, 1890, 1902, 1906; Odhner, 1934; Tardy, 1970) and ultimately rejected. In this paper, we follow the long-established and widely accepted classification of Nudibranchia of Cuvier (1817), which is supported by numerous morphological and molecular studies (Wägele and Willan, 2000; Zapata et al., 2014; Mahguib and Valdés, 2015), and retain Cladobranchia and Doridina as suborders of the order Nudibranchia.

Currently, two exon capture bait sets are available for the order Nudibranchia (Layton et al., 2020; Maroni et al., 2026), both originally designed for use in a subset of nudibranchs called dorids. Both approaches used a taxonomically diverse set of transcriptomes to find homologous genes, selecting them from transcriptomes of Doridina taxa, including species of Bathydoridiidae, Doridiidae, Actinocyclidae and Chromodorididae. Moreover, genes selected from eupulmonates in Teasdale et al. (2016) but present in dorids were also included in these bait sets (Layton et al., 2020; Maroni et al., 2026). However, the final set of targets differed because of their informativeness with respect to the taxonomic group in question (genus *Chromodoris* Alder & Hancock, 1855 vs. suborder Doridina) (Layton et al., 2020; Maroni et al.,2026). For example, Layton et al. (2020) identified homologous genes from 15 nudibranch transcriptomes but selected those that were maximally informative for species delimitation in *Chromodoris*. After searching for homologous genes, informative genes for discriminating between *Chromodoris westraliensis* (O'Donoghue, 1924) and *C. magnifica* (Quoy & Gaimard, 1832) were chosen for synthesis and the final set of sequences for bait design were from the species *C. westraliensis*. Maroni et al. (2026) used several Doridina transcriptomes for the design of the probes of this new bait set, combined with the transcriptome set from Layton et al. (2020), ultimately synthesizing baits that included multiple species of Doridina that target the same gene. A total of 304 genes were selected as phylogenetically informative for dorids (Maroni et al., 2026).

The evolutionary success of nudibranchs has been linked to the evolution of chemical defenses (Cimino and Ghiselin, 1999; da Cruz et al., 2011; Wilson et al., 2013; Winters et al., 2018), feeding diversification (Faucci et al., 2006; Goodheart et al., 2017) and/or ecological specialization (Ekimova et al., 2019). These factors could constitute key drivers of macroevolutionary patterns and cladogenesis of nudibranchs, explaining their remarkable diversity as the result of adaptive radiations shaped by ecological interactions. However, the lack of resolution in phylogenetic analyses makes the study of this line of inquiry challenging. The design of a bait set targeting more divergent taxa has the potential to address key questions about the evolution of nudibranchs, as has been demonstrated by previous studies using different approaches based on NGS. In Cladobranchia, transcriptome sequencing has allowed a greater understanding of the evolution of their cnidosacs, dietary patterns or swimming behaviour (Goodheart et al., 2015, 2017, 2018; Goodheart and Wägele, 2020; Karmeinski et al., 2021). In Doridina, target sequencing has been used for systematic studies, to obtain well-supported phylogenetic trees and to study evolutionary relationships among species (Layton et al., 2020; Paz-Sedano et al., 2024).

In the present study, we evaluate the possibility of using target capture sequencing across divergent taxa (i.e., all nudibranchs) for which a bait set was not intended, in order to assess the boundaries of the bait-to-target DNA sequence distance threshold for successful sequencing and exon assembly. Knowledge of this added value will be useful in the design of new bait sets optimized for capturing divergent species while minimizing redundancy in baits across relevant genetic distances. Moreover, it enables the evaluation of the effectiveness of bait sets previously designed for different taxonomic groups prior to conducting a new study. Using Nudibranchia as a case study, we use an exon capture bait set originally designed for subgroups of the Doridina to assess the effectiveness of the bait set across nudibranchs as a whole.

2. Material and methods

2.1. Taxon sampling

Taxon sampling included 79 newly sequenced specimens, belonging to both nudibranch suborders and representing 12 superfamilies, ensuring broad taxonomic coverage across the order. Specimens used in this study were loaned from the collection of the California State Polytechnic University, Pomona (CPIC), as well as collected during the expedition to Koumac, New Caledonia in 2018, organized by the Muséum national d'Histoire naturelle, Paris, France (MNHN). All collected specimens were individually photographed, labeled, preserved in 95% ethanol, and deposited at the MNHN (Table S1).

The species *Ardeadoris egretta* Rudman, 1984*, Chromodoris magnifica, Chromodoris westraliensis, Doriprismatica atromarginata* (Cuvier, 1804) and *Goniobranchus fidelis* (Kelaart, 1858) were included from available data from NCBI Genbank (Table S1), increasing taxon sampling using species previously sequenced with the chromodorid bait set (Layton et al., 2020).

2.2. DNA extraction, and sequencing

DNA of specimens was extracted from foot muscle tissue using E.Z.N.

A® Mollusc DNA Kit (Omega Bio-Tek), following the manufacturer's protocol with the modification of extending the lysis step to an overnight incubation. Quality control of the DNA was carried out using a NanoDrop™ spectrophotometer (Thermo Scientific, model 1000) to ensure quality and quantity of DNA extracted. Extracted DNA samples were sent to Arbor Biosciences (Ann Arbor, MI, USA) for target enrichment and high-throughput sequencing using a custom bait set 40 K-60 K kit (design ID D10333NUD2, Arbor Biosciences). Sequencing was performed on the Illumina NovaSeq 6000 platform, and demultiplexed FASTQ files were provided as output.

2.3. Sequence processing

Raw reads were trimmed using Trimmomatic v0.39 (Bolger et al., 2014), removing adapters and low-quality reads with a quality score below 15 in a 4-pb sliding window, and a minimum length of 26 bases. Clean reads were analyzed using HybPiper v2.2.0 (Johnson et al., 2016). The reads were mapped to the target bait set with nucleotide sequences using BLASTx (Camacho et al., 2009), sorted by gene, each gene was extracted, and assembled de novo using SPAdes v4.0.0 (Bankevich et al., 2012). Gene sequences were retrieved against the targeted exons using Exonerate (Slater and Birney, 2005), and statistics of the gene recovery efficiency were obtained (supplementary material Table S2). The genes obtained were filtered using the phylTER (Comte et al., 2023) tool in R v4.3.2. (R Core Team 2024). Phylter uses independent gene trees as input to detect gene sequences with low phylogenetic signal, genes with saturated signal leading to long branches, paralogous genes, genes associated with synteny breaks, and other sequences that are dubious in gene phylogenies (Comte et al., 2023), also filtering out possible contamination. To do this, retrieved genes were first aligned using MAFFT v7.526 (Katoh and Standley, 2013) and Maximum Likelihood (ML) analyses were performed IQ-TREE 2 (Minh et al. 2020) for each individual gene, using MoldelFinder (Kalyaanamoorthy et al., 2017) to find the best-fitting model for each gene. Outliers were identified using phylTER (Comte et al., 2023) with default parameters (supplementary material Table S3). Results associated with detected outliers were then visualized, comparing before and after filtering (Supplementary Figs. S1-S4), including how many species per gene were considered outliers, how many genes per species were considered outliers, dispersion of gene per species association, and optimization scores during optimization. Outliers were removed from initial retrieved gene sets, eliminating genes designated for specific taxa, and cleaned extracted genes were used for further analyses.

2.4. Phylogenetic analysis

Genes were aligned using MAFFT v7.526 (Katoh and Standley, 2013) and trimmed to remove poorly aligned regions using trimAL v1.5.rev0, using default parameters (Capella-Gutiérrez et al., 2009). IQ-TREE 2 (Minh et al. 2020) was used for concatenation of individual gene alignments. Matrices were filtered based on occupancy values, including in the matrix only those genes for which at least 25%, 50% and 75% of the taxa contained that gene. Summary statistics for the different matrices were calculated using the Alignment_assessment tool from (Portik et al., 2016). The full dataset and matrix with 25% occupancy were the same, as no genes were removed during filtering. Maximum Likelihood (ML) analyses in IQ-TREE 2 (Minh et al. 2020) were then performed with 1,500 ultrafast bootstrap (UFBoot) replicates, using MoldelFinder (Kalyaanamoorthy et al., 2017) to find the best-fitting model for each concatenated matrix and matrix subset. Nodes with UFBoot support less than 50% were collapsed. A branch was considered supported when UFBoot was >= 95%. Trees obtained were visualized using FigTree v1.4.3 (Rambaut, 2016). Because all trees showed the same topology, only the result from the 75% occupancy matrix was edited in Inkscape and is shown in the results.

2.5. Evaluation of broad effectiveness of taxon restricted datasets

To observe if the capture efficiency differs among taxa, descriptive statistics were performed including bar diagrams using the ggplot2 package (Wickham, 2016) in R v4.3.2. (R Core Team 2024), exploring whether specificity (percentage of reads that map to targets) and sensitivity (percentage of genes captured) vary depending on the taxa. The specificity is obtained in Hybpiper stats, expressed as PctOnTarget (supplementary Table S1). Because sequences previously available and newly sequenced species were obtained under different sequencing parameters, two specificity tests were carried out. Since specificity is influenced by both the sequencing approach and the number of initial reads, the sequences available in GenBank represent outliers in the specificity analyses. Results obtained from specificity analysis including all sequences are shown in supplementary file S1. A second specificity test excluded species previously available from GenBank. Regarding sensitivity, we evaluated differences in the percentage of captured genes among taxa, calculated from the number of genes extracted for each species after removing outliers divided between 304 genes included in the bait set. Additionally, we examined gene presence/absence in Doridina and Cladobranchia to determine whether the bait set preferentially targets genes conserved across Eupulmonata, potentially influencing its broad applicability. The length of the recovered gene in proportion to the gene in the bait set was also analyzed to identify potential differences between suborders. For this analysis, Hybpiper stats gives information about the length of each extracted gene for each species, and the mean length of a target gene, being the average over each source (species in the bait) for that gene. The percentage of length is the comparison of the extracted gene with the mean length, so sometimes it can be greater than 100%.

We investigated the hypotheses of *i*) whether capture specificity is related to the bait-to-target DNA sequence distance, with the number of reads mapping to targets decreasing with an increase in the bait-to-target sequence distances; *ii*) whether capture sensitivity is correlated with bait-to-target DNA distance. It was hypothesized that the greater the distance, the lower efficacy of the dataset and thus lower number of recovered genes. Furthermore, it was expected that these captured genes would be shorter in length with higher bait-to-target DNA distance. To first calculate genetic distances, each of the target genes extracted for each species was aligned with the gene in the bait set used for its recovery using MAFFT v7.526 (Katoh and Standley, 2013). The bait set includes sequences for the same gene from different species. However, the specific sequence in the bait set used for recovery of target genes could be seen in the Exonerate output during the extraction and assembly of exons in Hybpiper. Uncorrected pairwise distances were calculated, excluding gaps and missing data, and the average distance of all genes for a species was established as the genetic distance of the species with the bait set (supplementary Table S4).

The normal distribution of the variables, percentage of exons recovered for each species, percentage of reads mapped to target regions and bait-to-target DNA distances values were tested using the Shapiro-Wilk test (Shapiro and Wilk, 1965). Because none of the parameters follow a normal distribution, the two hypotheses above were first tested using the non-parametric Kendall rank correlation coefficient test (Kendall, 1938) performed in R v4.3.2. (R Core Team 2024). In addition, we used Davies test to further test for non-linearity between bait-to-target DNA distances and sensitivity. Because there was strong evidence for a change in slope between parameters, piecewise regression analysis (Toms and Lesperance, 2003) was conducted for this second hypothesis, to identify the distance threshold beyond which the percentage of genes extracted decreased. The analysis was performed using the segmented package (Fasola et al. 2018) in R v4.3.2. (R Core Team 2024), with bait-to-target DNA distance as the segmented covariate, estimating the threshold where the slope of the relationship changes. We also extracted a confidence interval using likelihood-based methods for the threshold, using the confint() function in the segmented package in R

v4.3.2. (R Core Team 2024). In addition, to confirm that the observed correlations are not inflated by non-independence among taxa Phylogenetic Generalized Least Squares (PGLS) analyses were performed in R, for specificity *vs.* p-distances as well as sensitivity *vs*. p-distances. Because phylogenetic correction did not affect the results, PGLS outputs are shown in Supplementary material File S1. AI technology ChatGPT was used in code construction for calculating genetic distances and correlation analyses.

3. Results

3.1. Exon capture sequencing and data matrices

A total of 76 nudibranch species spanning 35 families and three outgroup taxa belonging to Heterobranchia were newly sequenced. The average library insert size was consistent across the samples, around ~ 600 nt. After trimming adapters, the number of reads for the newly sequenced data ranged from 4,888,502 for *Berthellina delicata* (Pease, 1861) (an outgroup) to 24,039,128 for *Asteronotus cespitosus* (van Hasselt, 1824) (supplementary material Table S1). Chromodorid species previously sequenced in another study ranged from 66,404,420 for *Ardeadoris egretta* to 91,266,094 for *Chromodoris westraliensis* (supplementary Table S1). Between 157 and 303 genes were extracted from the newly sequenced species (supplementary material Table S1). The number of genes with a percentage length greater than 75% against the exon capture dataset (GenesAt75pc from HybPiper) varied from 5 in *Curnon granulosa* (Vayssière, 1906) to 268 genes in *Ceratosoma trilobatum* (J. E. Gray, 1850) for the newly sequenced data, and 263 in *Ardeadoris egretta* to 299 in *Chromodoris westraliensis* for previously available taxa (Supplementary material Table S2).

Outlier analyses identified 124 genes, each corresponding to a single gene from a specific species. These results included one complete outlier gene, but no complete outlier species (Supplementary material Table S3). The matrices gained 2.4% of concordance after outliers removal, filtering 0.55% of the data. A visualization of these results is provided in Supplementary material Figs. S1–S4.

After filtering outlier genes, the concatenated matrices with all genes and all species (full matrix), as well as the 25% occupancy matrix, included all 303 genes of the original bait set, with 45.7% informative sites and 48.0% gaps across alignment. The matrices with 50% and 75% occupancy retained 298 and 248 genes, respectively. In both matrices, 45.8% of sites were informative, while gaps accounted for 47.7% and 44.8% of sites across the alignments (Table 1).

3.2. Effectiveness of dorid exon capture bait set for broader phylogeny

Results on the specificity and sensitivity of the bait set showed variation in the efficiency of exon capture depending on the taxon and its phylogenetic affiliation. A lower specificity was observed in Cladobranchia, with the percentage of reads that map to targeted regions being between 9.4% and 45.2% in Doridina, 1.6%-21.5% in Cladobranchia and 5.5%-9.8% for outgroups (Supplementary material Table S1, Fig. 1).

The percentage of genes extracted after filtering outliers ranged from 91.8% to 99.7% in species belonging to Doridina, 51.6%-81.5% in Cladobranchia and 70.4%-72.4% in the outgroups (Fig. 2a). The gene presence/absence plot showed a subset of more conserved genes, recovered across all taxa; some genes with few absences in Cladobranchia/outgroup taxa; and a group of genes predominantly represented in Doridina, but rarely recovered in Cladobranchia and outgroups (Fig. 2b). In addition, we observed that the eupulmonate genes were distributed across the range of completeness, with some exhibiting high conservation and others being more frequently recovered in dorid taxa (Fig. 2b).

Furthermore, we observed that the extracted gene sequences were in general shorter than the original bait set, as reflected in the recovered gene length, expressed as a percentage of the target gene length against the gene in the bait set (Fig. 3). Moreover, cladobranchs and dorids showed opposite trends. In Cladobranchia, a higher percentage of genes was recovered at shorter lengths, with the number of extracted genes decreasing as the recovered gene length relative to the bait set increased. The same occurred with the outgroups. In the case of Doridina, the highest percentage of genes extracted almost the entire sequence relative to the bait set (Fig. 3).

The non-parametric Kendall rank test showed a linear negative correlation between the percent of reads that map to target regions and the bait-to-target DNA distance, with a correlation tau of −0.55 (z = -7.15, p-value = 8.63e-13) (Fig. 4a). The cladobranch species as well as outgroups exhibited the greatest distance and the lowest specificity. The sensitivity analysis also showed a negative correlation between the bait-to-target DNA distance and the percentage of genes extracted (tau = -0.56, z = -7.46, p-value = 8.5e-14). The Davies' test indicated a highly significant change in slope in the relationships between bait-to-target DNA distance and sensitivity, with a p-value of 2.2e-16, fitting a segmented regression model with a threshold at a mean distance of 0.19 ('best' at = 0.19, n.points = 10) or ~ 19%. Results of piercewise regression analysis are shown in Fig. 4b, where a linear trend was observed up to a distance threshold where the percentage of genes extracted drops sharply, with the threshold at 0.185 or 18.5% of bait-to-target DNA distance (R-squared: 0.79), with a 95% confidence interval ranging from 0.18 to 0.19. The taxa with the smallest distance belonged to the Chromodoridoidea superfamily. In general, the number of genes extracted from Doridina species was higher than cladobranchs and the outgroups. A cluster of taxa with greater bait-to-target DNA distance and a lower percentage of genes corresponded to Cladobranchia and outgroup taxa (Fig. 4b). Furthermore, genes of Doridina species generally showed lower genetic distances and recovered a higher percentage of gene length, while Cladobranchia tended to recover a lower percentage of gene length (Fig. 4c). In both suborders, a decline in the proportion of gene length recovery was observed beyond a genetic distance of 15–25% (Fig. 4c), obtaining shorter genes. The gene with the greatest genetic distance with respect to the bait set was found in the species *Antiopella barbarensis* (J. G. Cooper, 1863), with a distance of 44.5% (gene XP_005096019), with the extracted gene covering only 33.2% of the original bait sequence.

3.3. Phylogeny of Nudibranchia

The results obtained from maximum likelihood phylogenetic analyses (Fig. 5) showed maximum statistical support for most of

Table 1
Summary Statistics for the Different Occupancy Matrices. Stats Include Data of Total Genes, Sequence Length, Percentage of Informative Sites, Alignment base pairs, Proportion of N's and '?' and Proportion of Gaps Across Alignment.

% occupancy	Total genes	Sequence length	Percentage of informative sites	Alignment bp	Proportion of N's and?'s	Proportion of gaps across alignment
Nudibranchs_all and 25	303	359,272	45.7%	30,178,848	0.0%	48.0%
Nudibranchs_50	298	355,681	45.8%	29,877,204	0.0%	47.7%
Nudibranchs_75	248	305,295	45.8%	25,644,780	0.0%	44.8%

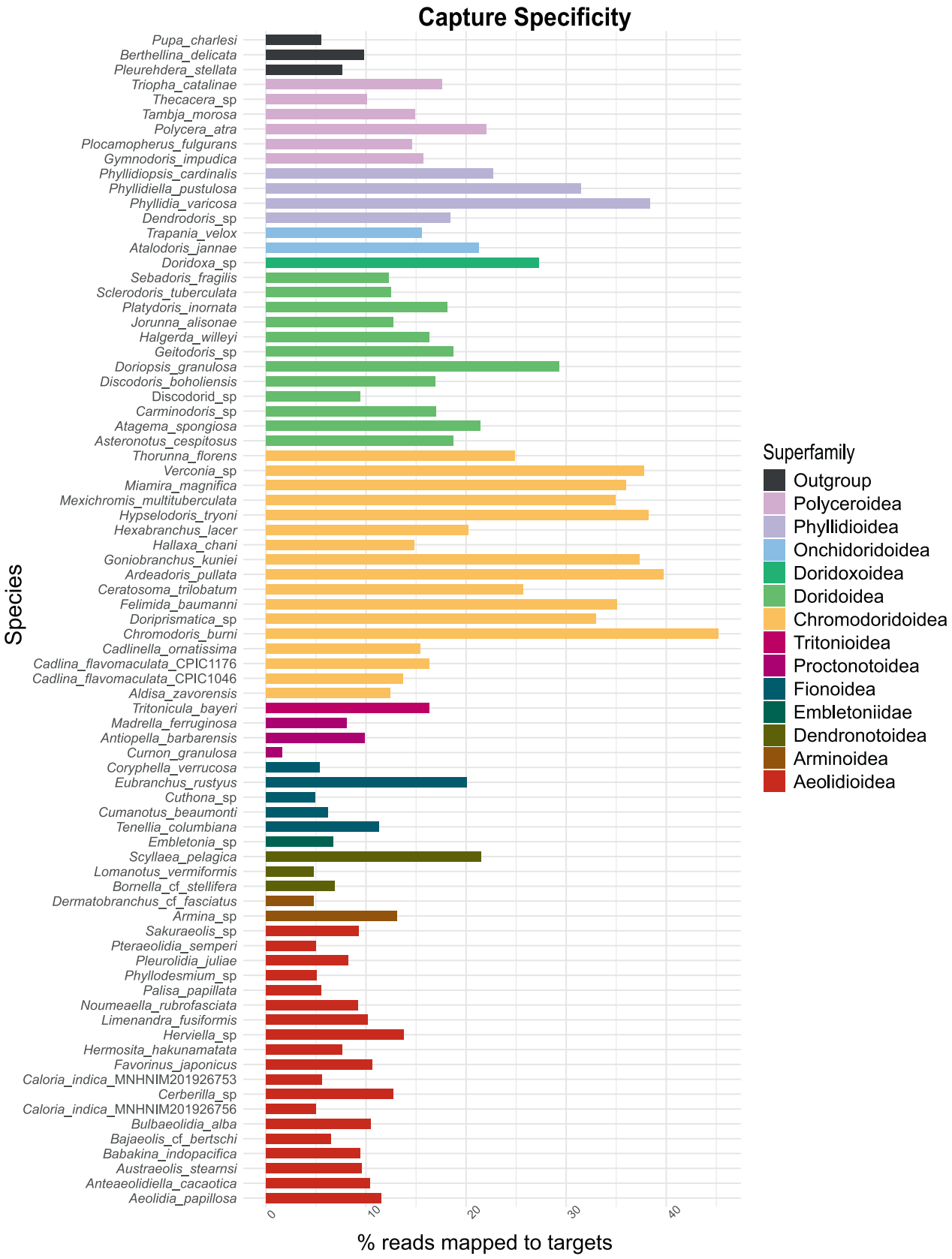

Fig. 1. Bar diagram of specificity, or the percentage of reads that map to target regions, in the different taxa.

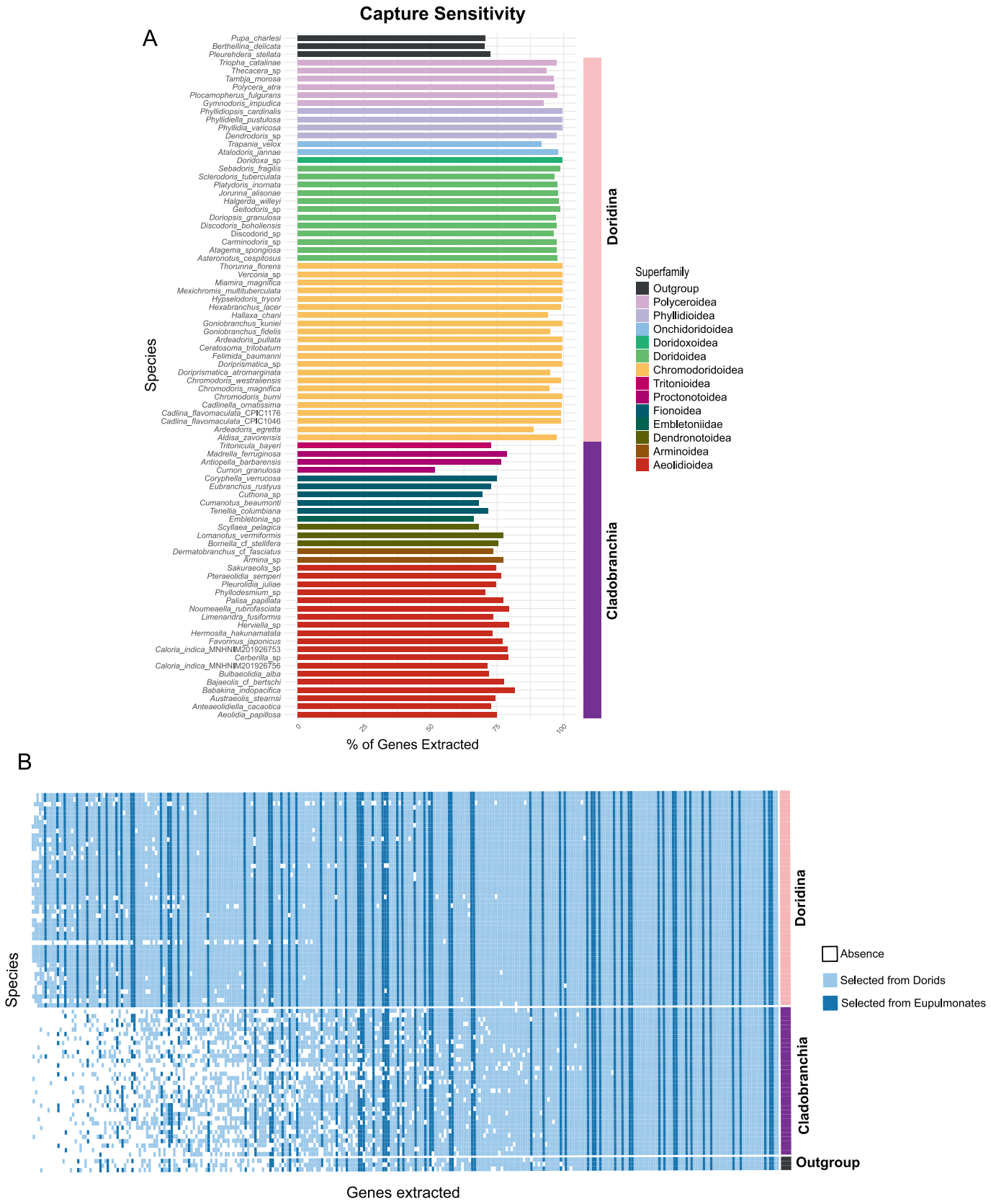

Fig. 2. (A) Bar diagrams of capture sensitivity, or percentage of exons that were extracted, in the different taxa; (B) Heatmap of genes extracted in the different taxa, indicating absence (white), presence selected in nudibranchs (light blue), presence of genes from eupulmonates in Teasdale et al. (2016) (dark blue).

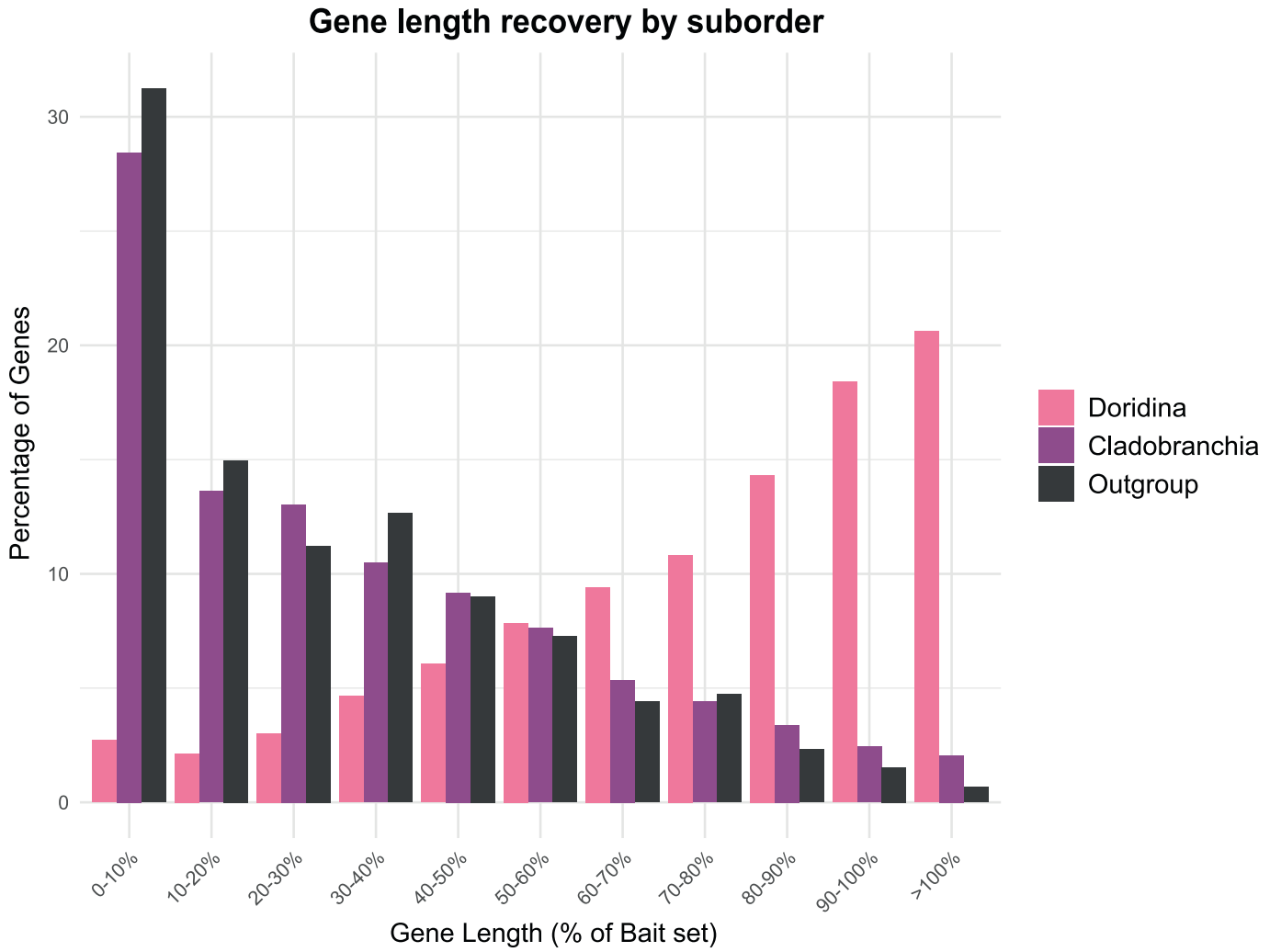

Fig. 3. Percentage of genes recovered at different proportions of sequence length relative to the bait set, classified by suborder.

phylogenetic relationships at all levels, suborder, superfamily, family and genus. Most of the genera were represented by a single species. Detailed results obtained from the phylogenetic relationships are found in the Supplementary material File S2.

The results obtained for the suborder Cladobranchia mostly agreed with previous studies (Goodheart et al., 2015, 2017, 2018; Karmeinski et al., 2021). Dendronotoidea was the sister superfamily to the remaining suborder (Fig. 5). A second well supported clade clustered the superfamilies Tritonioidea and Arminoidea as sister to the remaining cladobranchs. The superfamily Proctonotoidea joined the enigmatic Embletoniidae, but this relationship was not well supported. This clade was sister to Aeolidina (the aeolids). Another well supported clade clustered the superfamilies Fionoidea + Aeolidioidea. These last two were shown to be non-monophyletic (Fig. 5).

Within the suborder Doridina, Chromodoridoidea and Onchidoridoidea were not monophyletic, with the families Actinocyclidae (Chromodoridoidea) and Goniodorididae (Onchidoridoidea) forming a clade that was sister to the remaining dorid nudibranchs (Fig. 5). Regarding the superfamily Doridoidea, Discodorididae was recovered as paraphyletic due to the position of Doridiidae (Fig. 5). Doridoxoidea was recovered as well within Doridina and sister to the clade containing Chromodoridoidea + Polyceroidea + Phyllidioidea. The families Cadlinellidae, Hexabranchidae, and Cadlinidae, which are traditionally classified within Chromodoridoidea, were not recovered within Chromodoridoidea in our phylogenetic analyses, with Hexabranchidae recovered as more closely related to Polyceroidea, and Cadlinidae recovered as more closely related to the monophyletic superfamily Phyllidioidea (Fig. 5).

4. Discussion

Target enrichment (such as exon capture) allows for the sequencing of phylogenetically informative target genes in a group of interest without the need for whole genome sequencing or the collection of fresh tissues for transcriptomes. However, differences in exon capture efficiency have been associated with the genetic distance between the bait set sequence(s) and the corresponding target gene in each taxon, which is related to the evolutionary divergence from the group for which the bait set was designed. This distance is of interest both when designing the bait set, where greater divergence may require a broader and more variable set of baits, and for sequencing a new taxonomic group of interest using an already designed bait set. Despite several previous studies agreeing that the capture success decreases with the increase of genetic bait-to-target DNA distance, the distance threshold has varied depending on the study, measured in turn under different criteria (Bragg et al., 2016; Portik et al., 2016; Bartoš et al., 2023; Lemarcis et al., 2025). For example, this threshold distance for capture success has changed from 10% in skinks (Bragg et al., 2016) when measuring the sequencing depth of coverage, 15–20% in apoid wasps (Mayer et al., 2016) using a normalized base-coverage depth of contigs divided by the total number of sequenced nucleotides, and up to 39% in gnathostome vertebrates (Li et al., 2013) when considering the number of target

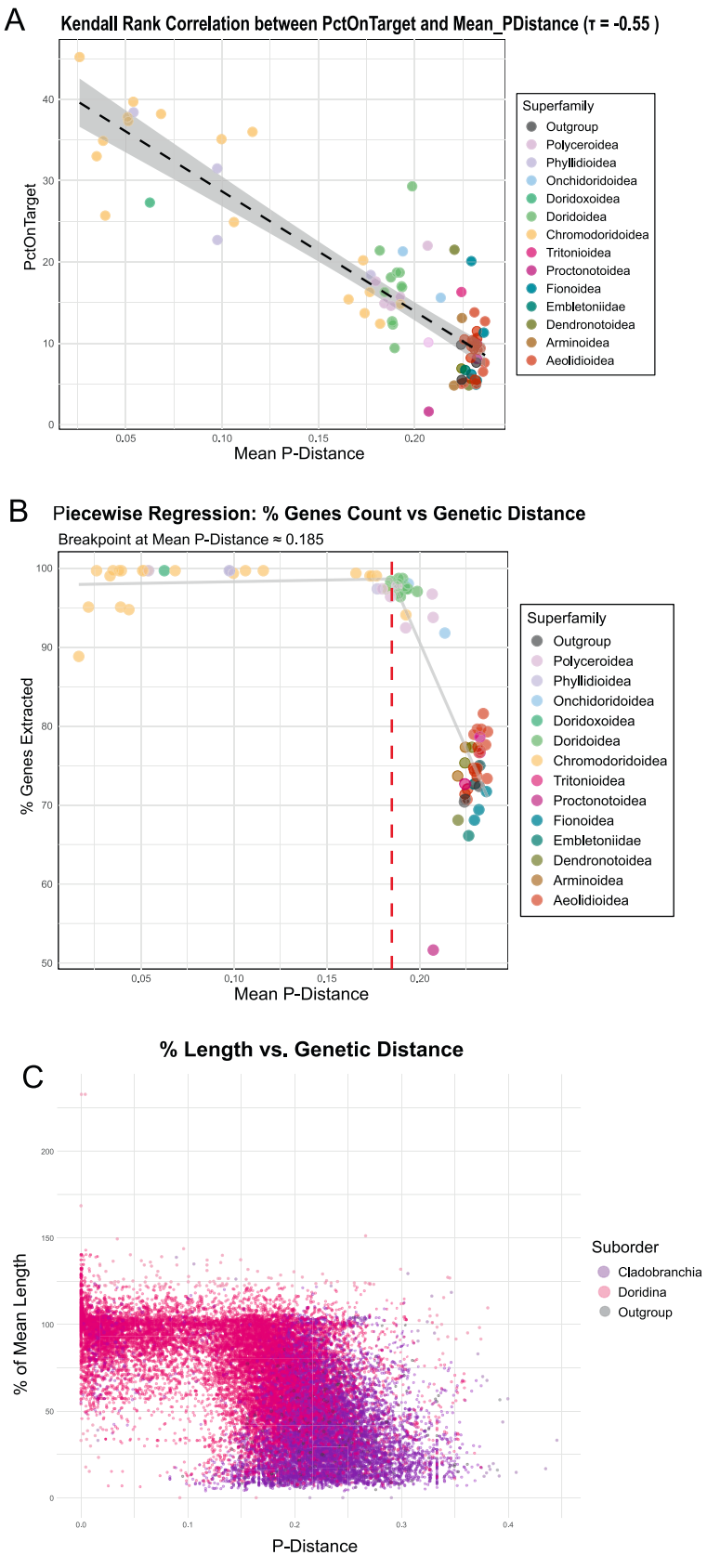

Fig. 4. (A) Kendall rank correlation test of capture specificity (PctOnTarget) and bait-to-target DNA distance. (B) Piece regression analysis of capture efficiency, showing the correlation between percentage of gene extracted in relation with bait-to-target DNA distance; (C) plot comparing the genetic distance of each gene for each species, and the percentage of length for that gene.

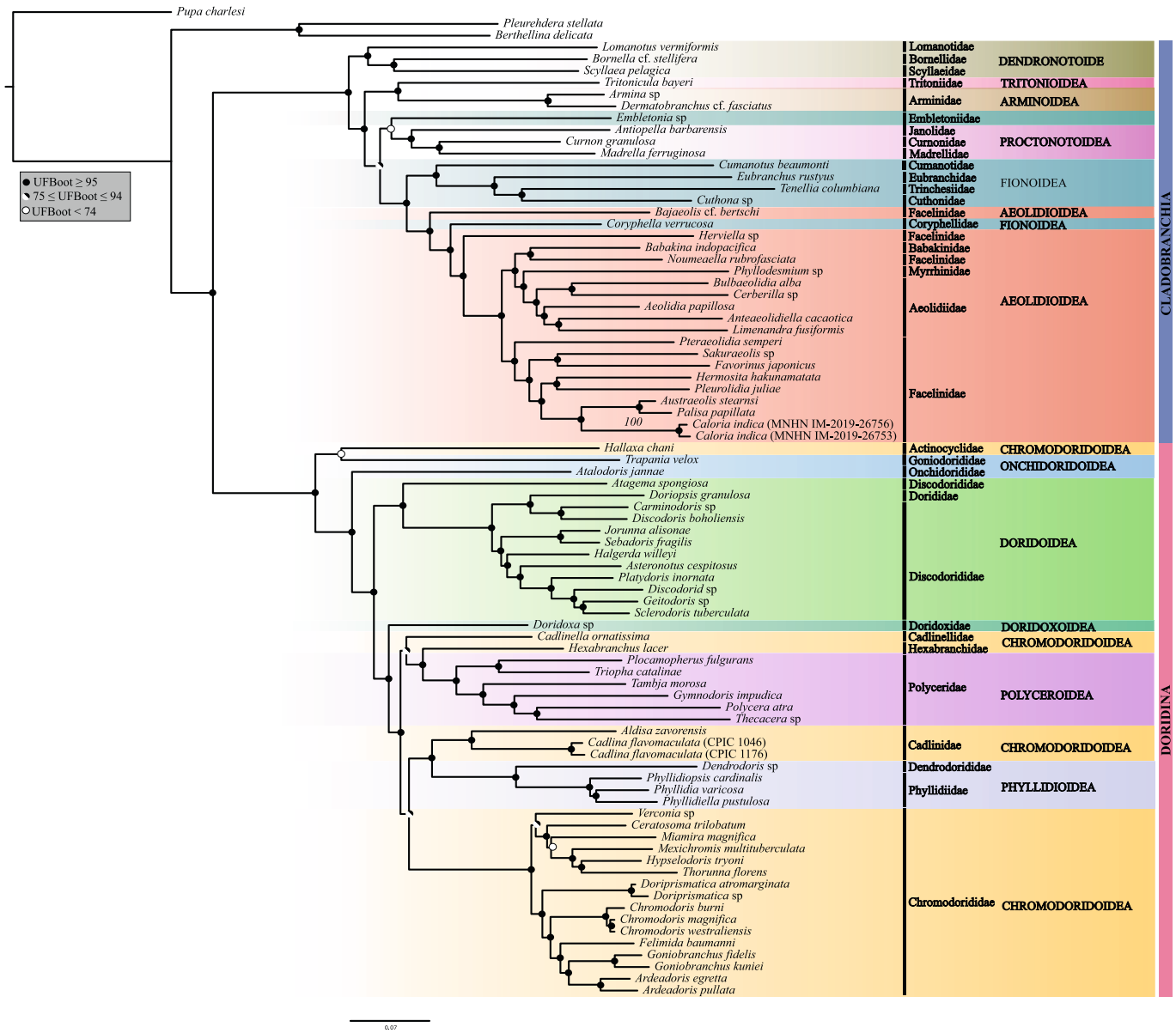

Fig. 5. Phylogenetic analyses of Nudibranchia based on 75% occupancy matrix, including 248 genes, using IQtree. Scale bar represents substitutions per site.

sequences captured. In gastropod molluscs, this threshold was previously investigated in Neogastropoda exons. Lemarcis et al. (2025) related the number of captured loci against the average p-distances. However, a clear threshold of genetic distances above which the exon capture efficiency significantly decreases was not identified, and the authors instead observed a gradual decrease in capture efficiency (Lemarcis et al., 2025). This may be because the authors optimized target capture by incorporating consensus sequences that reduce the genetic distance to the most divergent taxon pairs. As a result, all ingroup samples in the study showed genetic distances below 10% (Lemarcis et al., 2025). These consensus sequences were generated using ancestral sequence reconstruction, which infers past character states from modern sequences to represent the diversity of the input taxa (Dickson et al., 2021). Consequently, the resulting probe baits have a lower genetic distance to the target taxa, improving capture efficiency.

Differences in threshold values reported across studies may also depend on taxonomic sampling and bait design, as these can be influenced by the phylogenetic distribution of the sampled taxa. For example, if taxa more closely related to those included in the bait set are preferentially sampled, smaller genetic distances would be observed, potentially leading to lower threshold values. In such cases, a more gradual decline in capture efficiency would be expected, as reported by Lemarcis et al. (2025). Conversely, when more divergent taxa are included, higher threshold distances may be obtained, although capture efficiency is expected to decrease.

In the present study, we included taxa from all superfamilies within the order, allowing us to identify a bait-to-target DNA distance threshold of 18%, below which a high percentage of genes were still recovered, albeit with shorter sequence lengths compared to the original bait set. As a case study, we show here that an exon capture bait set designed for the nudibranch suborder Doridina can be effective for sequencing more distant species across the order Nudibranchia, and the number of sequences obtained are still useful for phylogenetic reconstruction.

4.1. Exon capture sequencing and data matrices

We observed that the total number of extracted genes varied among Doridina, Cladobranchia and outgroup taxa, with sequence capture for this bait set more effective in Doridina, for which 303 genes were extracted in most species, while Cladobranchia taxa ranged from 157 to

248 genes captured, and 214–220 genes captured for outgroups. This is consistent with the intended use of this bait set across taxa within Doridina.

In addition, species sequenced with a different bait set designed for chromodorids (Layton et al., 2020) were included to test for compatibility across the different bait sets. It was expected that most of the genes included in the Doridina bait set (Maroni et al., 2026) could be extracted from species sequenced using the chromodorid bait set. The Doridina bait set was designed using the same methodology as Layton et al. (2020), selecting 304 genes from Teasdale et al. (2016), Layton et al (2020), and dorid transcriptome datasets, and identifying intron/exon boundaries using *Aplysia californica* J. G. Cooper, 1863 genome across dorid transcriptomes. However, the Doridina bait set included new transcriptomes from two *Doris* Linnaeus, 1758 species, and the final set included baits from the different species with transcriptomes available (Maroni et al., 2026). Almost all target genes were extracted for previously available chromodorid taxa (chromodorid bait set) using the Doridina bait set. This compatibility between bait sets can also be observed in the genes present at different numbers of percent length (Supplementary material Table S2). Samples sequenced in (Layton et al., 2020) resulted in a range of 270–301 extracted genes (88 to 99%) with a percent length greater than 75% with respect to the target genes.

Our results showed a threshold distance of 18% for nudibranchs, after which capture efficiency drops sharply. The specificity decreased to around 10%, capture sensitivity was lower and the length of the genes were shorter. With a similar mean p-distance, cladobranch taxa had a lower percentage of genes captured compared to some dorids. The presence-absence matrix showed a group of genes more commonly extracted in Doridina, but rarely in cladobranch taxa. This difference in number of extracted genes when having a similar distance may be influenced by the presence of genes more specific to the Doridina suborder. Furthermore, it was not observed that the genes extracted in Cladobranchia could be largely influenced by genes conserved from eupulmonates, because these genes are distributed between those commonly extracted in both suborders, and those extracted mostly in Doridina. Although efficiency correlation analyses may be influenced by the presence of Doridina-specific genes, causing the lack of mapping of the reads (decreasing specificity), and extraction of the genes (decreasing sensitivity) against the same distance, these genes are not exclusive to this suborder; those same genes were recovered from a few cladobranch taxa, so genetic variability in these genes continues to affect their capture. Furthermore, genes recovered in Cladobranchia were generally shorter than the corresponding targets in the bait set, suggesting that although the genes were present, the reads did not fully map or assemble the complete gene. This indicates that sequence divergence has negatively impacted capture efficiency.

It should be noted that our results show a cladobranch cluster with similar average bait-to-target DNA distance as well as percentage of extracted genes for Cladobranchia, except for a lower number for *Curnon granulosa*. However, in the case of Doridina suborder, we find greater variability in the bait-to-target DNA distance, with particularly low average distance in some species belonging to Chromodoridoidea. This is likely because the original bait set design used transcriptomes from species also included in our analysis. Therefore, the distance between the bait set and target gene extracted is zero, since they correspond to the same species. This occurs in the species *Chromodoris magnifica, Ardeadoris egretta, Chromodoris westraliensis, Goniobranchus fidelis*, and *Doriprismatica atromarginata*.

In addition, the effectiveness of target capture sequencing may be influenced by other different parameters, such as the type of probes, hybridization parameters, and sample characteristics. Target enrichment utilizes DNA or RNA baits that can be immobilized on arrays or beats (Mamanova et al. 2010; Paijmans et al. 2016). Each type of probe has different advantages; RNA probes have shown a greater affinity for targeted genes than DNA probes, with RNA-DNA hybridization being more efficient, exhibiting stronger binding power. DNA probes are simpler and more cost-effective to synthesize, and DNA-DNA hybridization performs better in GC regions (Zhou et al. 2021; Sun et al. 2025). In addition, the probe length in the library may also affect the effectiveness of the library. When the target gene and probe gene sequences are highly variable, shorter probes reduce the number of variable positions in each probe, increasing sequencing efficiency. Longer probes ensure that the majority of probes are unique, but if the sequences are variable, hybridization fails and the efficiency decreases (Haslam et al. 2008). During the hybridization process, the number of sequenced genes can be influenced by temperature, although this in turn changes with sequence length and the nature of the sample. For example, Paijmans et al. (2016) observed that a lower hybridization temperature favors target capture sequencing for fresh material, while degraded samples obtained best results with a higher hybridization temperature.

4.2. Phylogenetic Implications for Nudibranchia

Suborder Cladobranchia.— Our analysis included six of the seven superfamilies in Cladobranchia (Aeolidioidea, Arminoidea, Dendronotoidea, Doridoxoidea, Fionoidea, Proctonotoidea, and Tritonioidea), as well as the family Embletoniidae. Results obtained matched those from previous studies (Goodheart et al., 2018; Karmeinski et al., 2021), supporting the monophyly of Dendronotoidea, Tritonioidea, Arminoidea, and Proctonotoidea, but Aeolidioidea and Fionoidea as paraphyletic. Dendronotoidea was recovered as the sister superfamily of all remaining cladobranchs, and Tritonioidea and Arminoidea clustered with high support. This shows that the former Dendronotacea (Dendronotoidea + Tritonioidea) and Arminacea (Proctonotoidea + Arminoidea) were again not supported as natural groups (Goodheart et al., 2018; Karmeinski et al., 2021).

The family Flabellinidae has faced a systematic problem, including paraphyletic lineages and ambiguities in the generic placement of several taxa due to unclear diagnostic morphological characters. Consequently, flabellinid taxa had been merged under the genus *Flabellina* Gray, 1833 (Ekimova et al., 2022). However, as a result of more recent morphological and molecular analyses, the highly diverse family Flabellinidae has undergone significant taxonomic reclassification, including the division of several other families such as Paracoryphellidae (Fionoidea), Flabellinopsidae (Aeolidioidea), Coryphellidae (Fionoidea), Unidentiidae (Fionoidea), Samlidae (Fionoidea) and Apataidae (Fionoidea) (Korshunova et al., 2017). However, these reclassifications show a lack of support at the superfamily level (Korshunova et al., 2017).

In the phylogeny by Korshunova et al. (2017), Coryphellidae (Fionoidea) was clustered with Flabellinidae (considered Aeolidioidea), both of which were closely related to Paracoryphellidae (Fionoidea); all three clades were grouped with Flabellinopsidae (Aeolidioidea). In other studies (Goodheart et al., 2018; Karmeinski et al., 2021), the monophyly of Aeolidioidea was disrupted by the inclusion of the genus *Flabellina* (family Flabellinidae, Fionoidea) within this clade, and in our results, the family Coryphellidae (Fionoidea) was recovered within Aeolidioidea. However, in 2022, Ekimova et al. (2022) showed possible oversplitting problems in Flabellinidae s.l.. In their study, the authors synonymized all Coryphellidae genera under the genus *Coryphella* J. E. Gray, 1850. This action could affect higher taxonomic ranks, such as rendering families and superfamilies as not monophyletic. Further studies focused on this group are needed to clarify the taxonomy of this complex group of nudibranchs.

Our results are also consistent with previous studies in the uncertain position of the enigmatic Embletoniidae. This taxon is critical for understanding the evolution of nematocyst sequestration, which is the capture and use of stinging organelles from the cnidarian prey of some nudibranchs (Goodheart et al., 2018; Karmeinski et al., 2021). Within Cladobranchia, nematocyst sequestration occurs in taxa belonging to Embletoniidae, Fionoidea, Aeolidioidea and Hancockiidae (superfamily Dendronotoidea). Previous studies identified two evolutionary origins

for this defence, one at the base of Aeolidida (Fionoidea + Aeolidioidea) and one in *Hancockia* Gosse, 1877 (Goodheart et al., 2018; Karmeinski et al., 2021). However, the position of Embletoniidae has long been ambiguous, with low support in previous analyses (Goodheart et al., 2018; Karmeinski et al., 2021) as well. In our study, *Embletonia* Alder & Hancock, 1851 was recovered as sister to Proctonotoidea, which would suggest a third origin, but this relationship did not have strong support.

Suborder Doridina.— One of the most compelling results of this study is the maximally supported placement of the superfamily Doridoxoidea within the suborder Doridina. This is inconsistent with current hypotheses, which place this clade within the suborder Cladobranchia. The taxonomic position of this group of polar nudibranchs has been hotly debated in the literature due to its peculiar anatomy along with the presence of anatomical symplesiomorphies shared with *Bathydoris* Bergh, 1884, another deep-sea/polar nudibranch. Originally *Doridoxa* Bergh, 1899 was considered as link between two nudibranch suborders (Bergh, 1899), then as the sister taxon to *Bathydoris* – the two sole members of the Gnathodoridoidea (Schmekel, 1985), and more recently as sister to the rest of the Doridina, including *Bathydoris* (Salvini-Plawen, 1990; Schrödl et al., 2001; Martynov, 2011; Mahguib and Valdés, 2015; Karmeinski et al., 2021). Contradicting all previously proposed hypotheses, our results strongly support the placement of Doridoxoidea within Doridina, closely related to the superfamily Doridoidea. Despite the peculiar lateral position of the anus, which suggested a close relationship with Cladobranchia, Doridoxoidea has a dorid body shape and blood gland (Bergh, 1899; Schrödl et al., 2001). The migration of the gill and anus from the dorsum to the side of the body, and the presence of a divided digestive gland, are not exceptional within Doridina (Domínguez et al., 2007; Martynov, 2011). The ontogeny study of the anus and gill arrangement in dorids showed that a differential growth of notal lobes during postmetamorphic juveniles causes the displacement of the anus to a posterior position, being a right lateral position of the anus a plesiomorphic characteristic (Martynov, 2011). Therefore, the shared characteristics of the digestive system between Doridoxoidea and Cladobranchia, as well as the lateral position of the gills, could be a reversion to a plesiomorphic character of Nudibranchia.

Our phylogenomic analyses supported the monophyly of three (Doridoidea, Polyceroidea and Phyllidioidea) of the five superfamilies into which Doridina is currently divided. The superfamily Phyllidioidea includes the families Dendrodorididae with Phyllidiidae, which was initially debated morphologically (Brunckhorst, 1993; Valdés and Gosliner, 1999) but was later confirmed in studies using Sanger sequencing (e.g. Valdés, 2003). However, in some studies this clade lacked support (Hallas et al., 2017; Korshunova et al., 2020). In addition, phylogenetic inference obtained from Sanger sequence data has shown Actinocyclidae as sister group to the rest dorids (Hallas et al., 2017; Korshunova et al., 2020). Our results agreed, showing with high support for Actinocyclidae + Goniodorididae as sister groups of remaining nudibranchs in Doridina.

The monophyly of the traditional superfamily Chromodoridoidea sensu Rudman (Rudman, 1984) was not supported in our analysis, but Chromodorididae is monophyletic when both *Cadlina* Bergh, 1879 and *Cadlinella* Thiele, 1931 are excluded. The separation of the temperate/cold-water taxon *Cadlina* from other chromodorids (which are mainly tropical) had been previously supported in molecular phylogenetic studies, resulting in the re-arrangement of *Cadlina* and *Aldisa* Bergh, 1878 in family-level taxon Cadlinidae (Johnson, 2011; Hallas et al., 2017; Korshunova et al., 2020; Prieto-Baños and Layton, 2025). *Cadlinella*'s affinity within Chromodorididae has long been questioned (Wilson and Healy, 2002). We recovered a well-supported clade in which Hexabranchidae and Polyceridae are closely associated, consistent with relationships inferred in previous molecular phylogenies (Hallas et al., 2017; Korshunova et al., 2020). The superfamily Doridoidea was also well-supported in our analyses. However, the families of this group of nudibranchs did not constitute natural groups. In our results, the dorid genus *Doriopsis* Pease, 1860 was nested inside the remaining Discodorididae genera, rendering it paraphyletic. In previous studies, the monophyly of Discodorididae was also compromised, with the genus *Aphelodoris* Bergh, 1879 joining as sister to the discodoridid genus *Atagema* Gray, 1850 (Fernández-Vilert et al., 2025).

5. Conclusion

Target capture sequencing is a widespread and efficient technique for studying the evolutionary history of species. In this work, we observed that the distance between the bait set and the target species is important for the success of sequencing and extraction of genes, being a key consideration when designing new bait sets or applying designed ones for different data. The number of extracted genes can be similar up to a distance threshold of 18%, but after that, the amount of reads that map to those genes decreases exponentially, and the extracted genes are shorter in length. However, target capture remains effective at this level of genetic divergence, at least for recovering sufficient data for phylogenetic inference. We sequenced 12 of the 13 superfamilies into which Nudibranchia is currently divided, including taxa belonging to both suborders Doridina and Cladobranchia, using a bait set designed for Doridina. Our results showed the efficiency of the bait set capturing genes across Nudibranchia, allowing the performance of well-supported phylogeny for the order.

6. Benefit-Sharing Statement

A research collaboration was developed with scientists from the countries providing genetic samples, all collaborators are included as co-authors, the results of research have been shared with the provider communities and the broader scientific community (see above).

7. Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used AI ChatGPT in order to generate the code for efficiency analysis, including obtaining genetic distances and statistical analysis and graphs. After using this tool/service, the authors reviewed and edited the content as needed and took full responsibility for the content of the published article.

CRediT authorship contribution statement

Sofía Paz-Sedano: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Ángel Valdés:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. **Carla C Stout:** Writing – review & editing, Methodology, Funding acquisition. **Kendall Feliciano:** Methodology. **Sandra Muro:** Methodology. **Nerida G. Wilson:** Writing – review & editing, Methodology. **Kara K.S. Layton:** Writing – review & editing, Methodology. **Jessica A. Goodheart:** Writing – review & editing, Writing – original draft, Supervision, Resources, Conceptualization.

Funding

This work was supported by NSF DEB-1711359 awarded to CCS. The specimens from New Caledonia were collected during the "Our Planet Reviewed" – New Caledonia expeditions (2016–2019), a joint project of MNHN and Conservatoire d'Espaces Naturels (CEN) de Nouvelle-Calédonie, funded mainly by the Gouvernement de la Nouvelle-Calédonie, Province Nord, Agence Française de la Biodiversité (AFB), the Lounsbery Foundation, and Office des Postes et Télécommunications (OPT). We also thank the Richard Gilder Graduate School for the Gerster Scholar Postdoctoral Research Fellowship to SPS and the high-performance computing cluster at the American Museum of Natural History.

Acknowledgements

The expeditions operated under a permit issued by Direction du Développement Economique et de l'Environnement (DDEE) of Province Nord. We are especially grateful to Philippe Bouchet for the invitation to participate in the expedition and to Sébastien Faninoz, Virginie Héros, Philippe Maestrati, Pierre Lozuet, Alain Daoulas, Jean-François Hervé, Yves Thévenet, and many others for their help and support. We also thank Sajesh Singh for his support using the computing cluster at AMNH, and members of Goodheart Lab for their feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2026.108630. It includes the following data: Figure S1. Number of species in each gene, showing how many per gene are outliers. Figure S2. Number of genes where each species is found, showing how many are outliers. Figure S3. Dispersion of data before and after outlier removal. One dot represents one gene x species association. Figure S4. Optimization scores during optimization. Table S1. Exon capture statistics. Species included in the taxon sampling, museum voucher, SRA - NCBI accession number, number of clean reads after trimming sequences, reads mapped, Percentage of reads that mapped to sequences in the target file (Pct On Target), total of bases recovered, genes mapped, genes with contigs, total of genes with sequences extracted and number of genes included in 75% occupancy matrix. Abbreviations: CPIC – Cal Poly Pomona Invertebrate Collection; MNHN – Muséum National d'Histoire Naturelle, Paris, France; WAM – Western Australian Museum, Perth, Australia. Table S2. Summary statistics of target enrichment and gene recovery efficiency after extraction and assembly of genes using hybpiper. Table S3. Outlier genes per species obtained after phylTER. Table S4. Bait-to-target DNA distances obtained for each species and each gene. File S1. Phylogenetic Generalized Least Squares (PGLS) analyses for specificity vs. p-distances as well as sensitivity vs. p-distances. File S2. Further results including specificity analyses including chromodorid species previously available, Kendall rank correlation test of capture sensitivity and number of reads in the analysis as well as more detailed phylogenetic results.

Data availability

All the original data necessary to reproduce the analyses reported in this study can be accessed through the Dryad link: https://datadryad.org/share/_Bfru0JyjnzYbU-VYi5VSxo0uqdAxeEv09yyUO38BPQ. Scripts used for the analyses are deposited in Zenodo https://doi.org/10.5281/zenodo.16929937 as well as github https://github.com/Sofiaps4/dorid-exons. The raw sequences are available in SRA-NCBI under the project Accession PRJNA1283123.

References

Alder J., Hancock A. 1845-1855. A monograph of the British nudibranchiate Mollusca. Parts 1-7. Ray Society.

Andermann, T., Torres Jiménez, M.F., Matos-Maraví, P., Batista, R., Blanco-Pastor, J.L., Gustafsson, A.L.S., Kistler, L., Liberal, I.M., Oxelman, B., Bacon, C.D., Antonelli, A., 2020. A guide to carrying out a phylogenomic target sequence capture project. Front. Genet. 10, 502347.

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 (5), 455–477.

Bartoš, O., Bohlen, J., Šlechtová, V.B., Kočí, J., Röslein, J., Janko, K., 2023. Sequence capture: obsolete or irreplaceable? a thorough validation across phylogenetic distances and its applicability to hybrids and allopolyploids. Mol. Ecol. Resour. 23, 1348–1360.

Bergh R. 1899. Nudibranchiate Gasteropoda. Dan Ingolf Exped. 2:1–49, pls. 1–5.

Bergh, R., 1890. Die cladohepatischen Nudibranchien. Zool. Jahrb. Abt. Syst. Geog. Biol. Tiere 5 (1), 1–75.

Bergh R. 1902. The Danish Expedition to Siam 1899-1900. Results of the zoological collections made by Dr. Th. Mortensen. I. Gasteropoda Opisthobranchiata. Kongel. Danske Vidensk. Selsk. Skr., Naturvidensk. Math. Afd, 12(2): 153–218.

Bergh, R., 1906. Uber clado- und holohepatische nudibranchiate Gastropoden. Zool. Jahrb. Abt. Syst. Geog. Biol. Tiere 23 (6), 739–742.

Blainville M.H. 1814. Mémoire sur la classification méthodique des animaux mollusques, et établissement d'une nouvelle considération pour y parvenir. Bulletin des Sciences par la Société Philomatique de Paris, Zoologie:175–80.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Bragg, J.G., Potter, S., Bi, K., Moritz, C., 2016. Exon capture phylogenomics: efficacy across scales of divergence. Mol. Ecol. Resour. 16, 1059–1068.

Brunckhorst, D.J., 1993. The systematics and phylogeny of Phyllidiid Nudibranchs (Doridoidea). Rec. Aust. Mus Suppl 16, 1–107.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 1–9.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Cheon, S., Zhang, J., Park, C., 2020. Is Phylotranscriptomics as reliable as phylogenomics? Mol. Biol. Evol. 37, 3672–3683.

Cimino, G., Ghiselin, M.T., 1999. Chemical defense and evolutionary trends in biosynthetic capacity among dorid nudibranchs (Mollusca: Gastropoda: Opisthobranchia). Chemoecology 9, 187–207.

Comte, A., Tricou, T., Tannier, E., Joseph, J., Siberchicot, A., Penel, S., Allio, R., Delsuc, F., Dray, S., de Vienne, D.M., 2023. PhylteR: efficient identification of outlier sequences in phylogenomic datasets. Mol. Biol. Evol. 40 (11), msad234.

Cowman, P.F., Quattrini, A.M., Bridge, T.C.L., Watkins-Colwell, G.J., Fadli, N., Grinblat, M., Roberts, T.E., McFadden, C.S., Miller, D.J., Baird, A.H., 2020. An enhanced target-enrichment bait set for Hexacorallia provides phylogenomic resolution of the staghorn corals (Acroporidae) and close relatives. Mol. Biol. Evol. 153, 106944.

Cuvier, G. 1817. Le règne animal distribué d'après son organisation, pour servir de base a l'Histoire Naturelle des Animaux et de l'Introduction a l'Anatomie Comparee volumen 2. Deterville.

da Cruz, J.F., Gaspar, H., Calado, G., 2011. Turning the game around: toxicity in a nudibranch-sponge predator–prey association. Chemoecology 22, 47–53.

Dean, L.J., Prinsep, M.R., 2017. The chemistry and chemical ecology of nudibranchs. Nat. Prod. Rep. 34, 1359–1390.

Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6, 361–375.

Dickson, Z.W., Hackenberger, D., Kuch, M., Marzok, A., Banerjee, A., Rossi, L., Klowak, J. A., Fox-Robichaud, A., Mossmann, K., Miller, M.S., Surette, M.G., Golding, G.B., Poinar, H., 2021. Probe design for simultaneous, targeted capture of diverse metagenomic targets. Cell Rep. Methods 1 (6).

Domínguez, M., Quintas, P., Troncoso, J.S., 2007. Phyllidiidae (Opisthobranchia: Nudibranchia) from Papua New Guinea with the description of a new species of Phyllidiella. American Malacological Bulletin 22 (1), 89–117.

Ekimova, I., Valdés, Á., Chichvarkhin, A., Antokhina, T., Lindsay, T., Schepetov, D., 2019. Diet-driven ecological radiation and allopatric speciation result in high species diversity in a temperate-cold water marine genus Dendronotus (Gastropoda: Nudibranchia). Mol. Phylogenet. Evol. 141, 106609.

Ekimova, I., Valdés, Á., Malaquias, M.A.E., Rauch, C., Chichvarkhin, A., Mikhlina, A., Antokhina, T., Chichvarkhina, O., Schepetov, D., 2022. High-level taxonomic splitting in allopatric taxa causes confusion downstream: a revision of the nudibranch family Coryphellidae. Zool. J. Linn. Soc. 196, 215–249.

Erickson, K.L., Pentico, A., Quattrini, A.M., McFadden, C.S., 2021. New approaches to species delimitation and population structure of anthozoans: two case studies of octocorals using ultraconserved elements and exons. Mol. Ecol. Resour. 21, 78–92.

Fasola, S., Muggeo, V.M., Kuchenhoff, H., 2018. A heuristic, iterative algorithm for change-point detection in abrupt change models. Comput. Stat. 33 (2), 997–1015.

Faucci, A., Toonen, R.J., Hadfield, M.G., 2006. Host shift and speciation in a coral-feeding nudibranch. Proc. R. Soc. B 274 (1606), 111–119.

Fernández-Vilert, R., Arnedo, M.A., Salvador, X., Valdés, Á., Schrödl, M., Moles, J., 2025. Shining disco: shedding light into the systematics of the family Discodorididae (Gastropoda: Nudibranchia). Zool. J. Linn. Soc. 203, zlae170.

Férussac, A. 1822. Tableaux Systématiques des Animaux mollusques. Surivie d'un Prodrome Générale, I-XLVII: 1–110.

Fuhrman, F.A., Fuhrman, G.J., DeRiemer, K., 1979. Toxicity and pharmacology of extracts from dorid nudibranches. Biol. Bull. 156, 289–299.

González-Delgado, S., Rodríguez-Flores, P.C., Giribet, G., 2024. Testing ultraconserved elements (UCEs) for phylogenetic inference across bivalves (Mollusca: Bivalvia). Mol. Phylogenet. Evol. 198, 108129.

Goodheart, J.A., Bazinet, A.L., Collins, A.G., Cummings, M.P., 2015. Relationships within Cladobranchia (Gastropoda: Nudibranchia) based on RNA-Seq data: an initial investigation. R. Soc. Open Sci. 2 (9), 150196.

Goodheart, J.A., Bazinet, A.L., Valdés, Á., Collins, A.G., Cummings, M.P., 2017. Prey preference follows phylogeny: evolutionary dietary patterns within the marine gastropod group Cladobranchia (Gastropoda: Heterobranchia: Nudibranchia). BMC Evol. Biol. 17, 1–14.

Goodheart, J.A., Bleidißel, S., Schillo, D., Strong, E.E., Ayres, D.L., Preisfeld, A., Collins, A.G., Cummings, M.P., Wägele, H., 2018. Comparative morphology and evolution of the cnidosac in Cladobranchia (Gastropoda: Heterobranchia: Nudibranchia). Front. Zool. 15, 1–18.

Goodheart, J.A., Wägele, H., 2020. Phylogenomic analysis and morphological data suggest left-right swimming behavior evolved prior to the origin of the pelagic Phylliroidae (Gastropoda: Nudibranchia). Orgn. Divers. Evol. 20, 657–667.
Hallas, J.M., Chichvarkhin, A., Gosliner, T.M., 2017. Aligning evidence: concerns regarding multiple sequence alignments in estimating the phylogeny of the Nudibranchia suborder Doridina. R. Soc. Open Sci. 4 (10), 171095.
Haslam, N.J., Whiteford, N.E., Weber, G., Prügel-Bennett, A., Essex, J.W., Neylon, C., 2008. Optimal probe length varies for targets with high sequence variation: implications for probe library design for resequencing highly variable genes. PLoS One 3 (6), e2500.
Hughes, L.C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C.C., Thompson, A.W., Arcila, D., Betancur-R, R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhou, Z., Huang, H., Chen, S., Venkatesh, B., Shi, Q., 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proc. Natl. Acad. Sci. 115, 6249–6254.
Hughes, L.C., Ortí, G., Saad, H., Li, C., White, W.T., Baldwin, C.C., Crandall, K.A., Arcila, D., Betancur-R, R., 2021. Exon probe sets and bioinformatics pipelines for all levels of fish phylogenomics. Mol. Ecol. Resour. 21, 816–833.
Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. C., Wickett, N.J., 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Appl. Plant. Sci. 4.
Johnson, R.F., 2011. Breaking family ties: taxon sampling and molecular phylogeny of chromodorid nudibranchs (Mollusca, Gastropoda). Zool. Scr. 40, 137–157.
Johnson, R.F., Gosliner, T.M., 2012. Traditional taxonomic groupings mask evolutionary history: a molecular phylogeny and new classification of the chromodorid nudibranchs. PLoS One 7, e33479.
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.
Karmeinski, D., Meusemann, K., Goodheart, J.A., Schroedl, M., Martynov, A., Korshunova, T., Wägele, H., Donath, A., 2021. Transcriptomics provides a robust framework for the relationships of the major clades of cladobranch sea slugs (Mollusca, Gastropoda, Heterobranchia), but fails to resolve the position of the enigmatic genus Embletonia. BMC Ecol. Evol. 21, 1–17.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Katz, P.S., 2016. Phylogenetic plasticity in the evolution of molluscan neural circuits. Curr. Opin. Neurobiol. 41, 8–16.
Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30, 81–93.
Korshunova, T., Fletcher, K., Picton, B., Lundin, K., Kashio, S., Sanamyan, N., Sanamyan, K., Padula, V., Schrödl, M., Martynov, A., 2020. The Emperor's *Cadlina*, hidden diversity and gill cavity evolution: new insights for the taxonomy and phylogeny of dorid nudibranchs (Mollusca: Gastropoda). Zool. J. Linn. Soc. 189, 762–827.
Korshunova, T., Martynov, A., Bakken, T., Evertsen, J., Fletcher, K., Mudianta, I.W., Saito, H., Lundin, K., Schrödl, M., Picton, B., 2017. Polyphyly of the traditional family Flabellinidae affects a major group of Nudibranchia: aeolidacean taxonomic reassessment with descriptions of several new families, genera, and species (Mollusca, Gastropoda). Zookeys 717, 1–139.
Korshunova, T., Fletcher, K., Martynov, A., 2025. The endless forms are the most differentiated—how taxonomic pseudo-optimization masked natural diversity and evolution: the nudibranch case. Zool. J. Linn. Soc 204 (4), zlaf057.
Kulski, J., 2016. Next-generation sequencing—an overview of the history, tools, and "Omic" applications. Next generation sequencing-advances, applications and challenges 3–60.
Layton, K.K.S., Carvajal, J.I., Wilson, N.G., 2020. Mimicry and mitonuclear discordance in nudibranchs: new insights from exon capture phylogenomics. Ecol. Evol. 10, 11966–11982.
Lemarcis, T., Blin, A., Cariou, M., Derzelle, A., Farhat, S., Fedosov, A., Zaharias, P., Zuccon, D., Puillandre, N., 2025. Too far from relatives? impact of the genetic distance on the success of exon capture in phylogenomics. Mol. Ecol. Resour. 25, e14064.
Li, C., Hofreiter, M., Straube, N., Corrigan, S., Naylor, G.J.P., 2013. Capturing protein-coding genes across highly divergent species. Biotechniques 54, 321–326.
Lillvis, J.L., Gunaratne, C.A., Katz, P.S., 2012. Neurochemical and neuroanatomical identification of central pattern generator neuron homologues in Nudipleura molluscs. PLoS One 7, e31737.
Lozano-Fernandez, J., 2022. A Practical Guide to Design and Assess a Phylogenomic Study. Genome. Biol. Evol. 14.
Mahguib, J., Valdés, Á., 2015. Molecular investigation of the phylogenetic position of the polar nudibranch *Doridoxa* (Mollusca, Gastropoda, Heterobranchia). Polar Biol. 38, 1369–1377.
Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods 7, 111–118.
Maroni, P.J., Carvajal, J.I., Baker, B.J., Layton, K.K.S., Wilson, N.G., 2026. Testing a mitochondrial species radiation with exon capture data – the Antarctic sea slug Doris 'kerguelenensis. Molecular Phylogenetics and Evolution. Accepted.
Martynov, A.V., 2011. From "tree-thinking" to "cycle-thinking": ontogenetic systematics of nudibranch molluscs. Thalassas 27, 193–224.
Mayer, C., Sann, M., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R.S., Petersen, M., Meusemann, K., Liere, K., Wägele, J.-W., Misof, B., Bleidorn, C., Ohl, M., Niehuis, O., 2016. BaitFisher: a software package for multispecies target DNA enrichment probe design. Mol. Biol. Evol. 33, 1875–1886.
McCormack, J.E., Tsai, W.L.E., Faircloth, B.C., 2016. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16, 1189–1203.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37 (5), 1530–1534.
Odhner, N.H., 1934. The Nudibranchiata. British Museum (Terra Nova) Expedition, 1910. Nat. Hist. Rep. 7 (5), 229–310.
Pabst, E.A., Kocot, K.M., 2018. Phylogenomics confirms monophyly of Nudipleura (Gastropoda: Heterobranchia). J. Moll. Stud. 84, 259–265.
Paijmans, J.L., Fickel, J., Courtiol, A., Hofreiter, M., Förster, D.W., 2016. Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. Mol. Ecol. Resourc. 16 (1), 42–55.
Paz-Sedano, S., Moles, J., Smirnoff, D., Gosliner, T.M., Pola, M., 2024. A combined phylogenetic strategy illuminates the evolution of Goniodorididae nudibranchs (Mollusca, Gastropoda, Heterobranchia). Mol. Phylogenet. Evol. 192, 107990.
Portik, D.M., Smith, L.L., Bi, K., 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). Mol. Ecol. Resour. 16, 1069–1083.
Prieto-Baños, S., Layton, K.K.S., 2025. Tracing the evolution of key traits in dorid nudibranchs. PLoS One 20, e0317704.
R Core Team, 2024. _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
Rambaut A. 2016. FigTree. 1.4. 3. Website https://tree.bio.ed.ac.uk/software/figtree/.
Rincon-Sandoval, M., Betancur-R, R., Maldonado-Ocampo, J.A., 2019. Comparative phylogeography of trans-andean freshwater fishes based on genome-wide nuclear and mitochondrial markers. Mol. Ecol. 28, 1096–1115.
Rudman, W.B., 1984. The Chromodorididae (Opisthobranchia: Mollusca) of the Indo-West Pacific: a review of the genera. Zool. J. Linn. Soc. 81, 115–273.
Ruiz-Torres, V., Forsythe, N., Pérez-Sánchez, A., Van Schaeybroeck, S., Barrajón-Catalán, E., Micol, V., 2021. A Nudibranch Marine Extract Selectively Chemosensitizes Colorectal Cancer Cells by Inducing ROS-Mediated Endoplasmic Reticulum stress. Front. Pharmacol. 12, 625946.
Salvini-Plawen, L.V., 1990. Origin, phylogeny and classification of the phylum mollusca. Iberus 9, 1–34.
Schmekel, L., 1985. Aspects of Evolution within the opisthobranchs. In: Trueman, R. (Ed.), The Mollusca Vol. 10. Evolution. Academic Press, New York, pp. 221–267.
Schrödl, M., Wägele, H., Willan, R.C., 2001. Taxonomic redescription of the doridoxidae (gastropoda: opisthobranchia), an enigmatic family of deep water nudibranchs, with discussion of basal nudibranch phylogeny. Zool. Anz. 240, 83–97.
Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples)†. Biometrika 52, 591–611.
Slater, G.S.C., Birney, E., 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinform. 6, 1–11.
Sun, T., Li, S., Liu, Y., Zhou, K., Wang, J., Lei, Z., Guo, X., Xing, J., Guo, W., 2025. Comparative evaluation of DNA and RNA probes for capture-based mitochondrial DNA next-generation sequencing. BMC Biol. 23 (1), 248.
Tardy, J., 1970. Contribution a l'etude des metamorphoses chez les nudibranches. Ann. Sci. Nat. Zool. 12, 299–370.
Teasdale, L.C., Köhler, F., Murray, K.D., O'Hara, T., Moussalli, A., 2016. Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. Mol. Ecol. Resour. 16 (5), 1107–1123.
Toms, J.D., Lesperance, M.L., 2003. Piecewise Regression: a Tool for Identifying Ecological Thresholds. Ecology 84, 2034–2041.
Valdés, A., 2003. Preliminary molecular phylogeny of the radula-less dorids (gastropoda: opisthobranchia), based on 16S MTDNA sequence data. J. Moll. Stud. 69, 75–80.
Valdés, A., 2004. Phylogeography and phyloecology of dorid nudibranchs (Mollusca, Gastropoda). Biol. J. Linn. Soc. 83, 551–559.
Valdés, A., Gosliner, T.M., 1999. Phylogeny of the radula-less dorids (Mollusca, Nudibranchia), with the description of a new genus and a new family. Zoo. Scr. 28, 315–360.
Wägele, H., Willan, R.C., 2000. Phylogeny of the Nudibranchia. Zool. J. Linn. Soc. 130 (1), 83–181.
Wang, Q., Purrafee, D.L., Huang, J., Kumar, S.K., Kevrekidis, C., Reichenbacher, B., Reza, E.H., Straube, N., Moritz, T., Li, C., 2022. Molecular phylogenetics of the Clupeiformes based on exon-capture data and a new classification of the order. Mol. Phylogenet. Evol. 175, 107590.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York, ISBN 978-3-319-24277-4.
Wilson, N.G., Healy, J.M., 2002. Is *Cadlinella ornatissima* a chromodorid? sperm ultrastructure in an enigmatic nudibranch (Opisthobranchia, Mollusca). Invertebr. Reprod. Dev. 42, 179–188.
Wilson, N.G., Alan, M.J., Baker, B.J., 2013. A species flock driven by predation? secondary metabolites support diversification of slugs in Antarctica. PLoS One 8, e80277.
Winters, A.E., Wilson, N.G., van den Berg, C.P., How, M.J., Endler, J.A., Justin, M.N., White, A.M., Garson, M.J., Cheney, K.L., 2018. Toxicity and taste: unequal chemical defences in a mimicry ring. Proc. R. Soc. B 285 (1880), 20180457.
Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C.S., Jörger, K.M., Schrödl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc. R. Soc. Lond. B 281, 20141739.
Zhou, J., Zhang, M., Li, X., Wang, Z., Pan, D., Shi, Y., 2021. Performance comparison of four types of target enrichment baits for exome DNA sequencing. Hereditas 158 (1), 10.