



# Global species hotspots and COI barcoding cold spots of marine Gastropoda

Andrew F. Torres<sup>1,2</sup> · Owen S. Wangensteen<sup>3</sup> · Willem Renema<sup>2,4</sup> · Christopher P. Meyer<sup>5</sup> · Ian Kendrick C. Fontanilla<sup>6</sup> · Jonathan A. Todd<sup>1</sup>

Received: 13 November 2023 / Revised: 7 June 2024 / Accepted: 20 June 2024  
© The Author(s) 2024

## Abstract

With the on-going efforts in digitising museum collections, increased participation of citizen scientists, and greater accessibility to research data, accurately determining global patterns of diversity has become more achievable. Here, we used occurrence records from the Global Biodiversity Information Facility, with annotation of authoritative taxonomy, to evaluate the taxonomic richness of marine gastropods and to identify global species hotspots for this group. We also reviewed the availability of genetic resources within hotspots to detect potentially important regions where reference sequences for identifying these organisms are wanting. We find 33,268 unique and valid species under 3291 genera belonging to 380 gastropod families that have been recorded from 1662 to 2023. Globally, only 12.1% of reported species are linked to a COI barcode, whilst 27.1% of the families are represented with complete mitogenomes. Georeferenced records show the wide albeit disparate distribution of observations and species counts. The compiled dataset, published by organisations that are restricted to 55 countries, reveal a seeming lack of local submissions, but nonetheless demonstrates the growing contribution of citizen science platforms. We present 28 marine provinces across the globe as putative gastropod species hotspots and call for further work and stronger involvement, particularly within COI barcoding cold spots, to address the observed genetic reference inequity. Lastly, we highlight the important and relevant role of open and inclusive science to biodiversity monitoring and research.

**Keywords** Marine gastropods · Species occurrence · Global diversity · COI · Open science · Citizen science

## Introduction

Gastropods are molluscs of incredible diversity both in form and in function (Bouchet et al. 2017; Ponder et al. 2019). Likely stemming from univalved molluscs of the Ediacaran—early Cambrian, current gastropods comprise the snails, slugs, whelks, and limpets that are distinguished from all other molluscs by the occurrence of a single shell,

---

Communicated by Paolo G. Albano.

---

Extended author information available on the last page of the article

an operculum for most, and larval torsion at least once in their ontogeny (Aktipis et al. 2008; Parkhaev 2008). Their success in radiating into the marine, freshwater, and terrestrial ecosystems have made them ecologically important and cosmopolitan components in most trophic interactions as well as in nutrient and carbon cycles (Ponder et al. 2019). In the human context, gastropods have been a source of food, pharmaceutical compounds, and cultural identity but have also provided medical problems (e.g., as intermediate hosts to parasitic flatworms) and agricultural nuisance (Barker 2002; Olivera et al. 2014; Dang et al. 2015; Giannelli et al. 2016).

Despite their ubiquity, a definitive global assessment on the diversity of modern marine gastropods is still lacking, and with reason. The wide range of niches and environments that gastropods have been observed to occupy requires an equivalent range of taxonomic specialisation (and institutional support) to study them. It is also hypothesised that the majority of gastropod diversity would come from the generally unsampled and lesser-studied small and cryptic (i.e., hidden) molluscs, making them even harder to discover (Bouchet et al. 2002). But it is perhaps the sheer number of species within this group, backdropped by their complicated taxonomic history and the perennial reorganisation of their phylogeny, that contribute to the difficulty of determining how many species of marine gastropods there actually are. Current estimates of total gastropod diversity reach 150,000 species (Ponder and Lindberg 2008; Appeltans et al. 2012). Up to 45,000 of these are theorised to be undescribed species that are stored in specimen collections, whilst 60,000 are morphospecies yet to be sampled and discovered. Around 63,000 extant gastropod species have been named and described (Aktipis et al. 2008; Bouchet et al. 2017), and these numbers continue to grow. Advancements in imaging and sequencing technologies have accelerated describing new species and revealing non-monophyletic taxa and cryptic species (Meyer 2003; Giribet 2008; Duda et al. 2008; Puillandre et al. 2011, 2014, 2015; Golding et al. 2014; Zapata et al. 2014; Varney et al. 2021; Kantor et al. 2022). At the same time, the increased number of workers and research projects conducted in historically under-studied regions and water depths as well as the use of various sampling techniques have also steadily increased the rates of species discovery (Bouchet et al. 2002, 2016, 2023; Appeltans et al. 2012; Thaler and Amon 2019; Cunha et al. 2023).

Nevertheless, empirical information on the current overall diversity of any taxon is important for understanding their ecology and evolution and can facilitate developing strategies for their conservation and management. With the on-going efforts in digitising natural history collections (Yesson et al. 2007; Nelson and Ellis 2019) and the increased participation of citizen scientists (Dickinson et al. 2012; Kosmala et al. 2016; Bouchet et al. 2016), collaboration, and accessibility to research data (Costello 2009; Bingham et al. 2017; Escribano et al. 2018), accurately determining global patterns of biodiversity has become more achievable. For this, the Global Biodiversity Information Facility (GBIF) can be an invaluable source of such information. GBIF is an international network and data facility that is dedicated to aggregating and distributing open-access, standardised biodiversity data. It houses centuries worth of specimens-based and observations-based occurrence records compiled from museums, governmental environmental monitoring surveys, and research-grade community science platforms, among others. Although not without issues on data quality and taxonomic and regional biases, GBIF has been at the forefront among data infrastructures for primary biodiversity metadata, which can still be proven powerful and essential (Edwards 2004; Yesson et al. 2007; Maldonado et al. 2015; Kosmala et al. 2016; Nelson and Ellis 2019), especially for cosmopolitan, extensively collected, and diverse taxa such as marine gastropods.

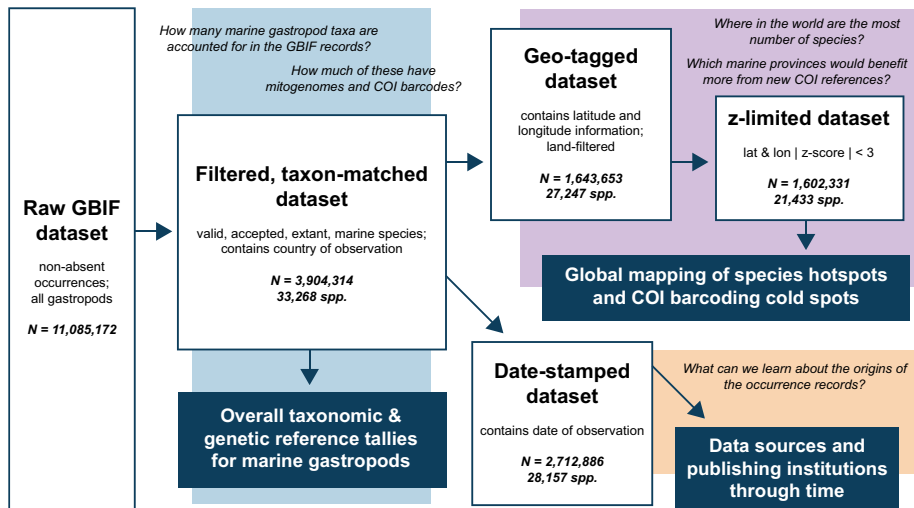
As essential as determining the taxonomic diversity of a group is knowing the status of available reference sequences that are used for genetic identification and characterisation of species and genetic diversity. Despite the expanding applications of DNA barcoding and metabarcoding to ecological, environmental, and health research, genetic methods for taxon identification are constrained by the availability of reference gene sequences (i.e., barcodes) representing the taxa to be identified. Lack of reference sequences, in metabarcoding studies for example, results in the use of broadly assigned molecular operational taxonomic units (MOTUs) which could limit the derivation of ecological insights (Blaxter et al. 2005; Schmidt et al. 2015; Múrria et al. 2020). Determining species-rich regions and taxonomic groups deficient in genetic resources, particularly of complete mitochondrial genomes and cytochrome c oxidase I (COX1 or COI) gene sequences, could be helpful in prioritising areas and taxa for the augmentation of reference sequence databases. In this study, we used occurrence records from GBIF, with annotation of authoritative taxonomy, to evaluate the taxonomic richness of marine gastropods and to identify putative global species hotspots for this group. Additionally, the availability of mitogenomes and COI barcodes of species was reviewed to assess the state of genetic resources for these marine molluscs and detect potentially important regions where reference sequences for identifying these organisms are wanting.

## Methods

A dataset compiling all recorded, present (i.e., non-absent) occurrences of gastropods ( $N = 11,085,172$ ) was downloaded from GBIF (GBIF.org 2023) (19 Sep 2023, <https://doi.org/10.15468/dl.qgsn2b>) and analysed in R v4.2.2 (R Core Team 2022). The dataset was filtered to contain records that have species identifications at the least. Subspecies, form, and variety identifications were included in the dataset but were analysed only at the species level. The dataset was further filtered to have country assignments, irrespective of whether these were originally in the record or estimated by GBIF based on geographic coordinates provided by the publishing organisation (Fig. 1).

All occurrence events were taxon-matched via the LifeWatch Species Information Backbone ([www.LifeWatch.eu](http://www.LifeWatch.eu)) and were annotated with authoritative taxonomy based on the World Register of Marine Species (WoRMS) with the R package 'worrms' v0.4.3 (Chamberlain and Vanhoorne 2023). Fuzzy matches were manually corrected based on verbatim scientific names and authorities. Ambiguous scientific names with no attached authors were not considered for taxon matching and were removed. Any record found unmatched with a valid Aphia ID was likewise removed. Events based on fossil specimens in GBIF and records of species flagged as extinct in WoRMS were filtered out. To retain exclusively marine taxa, records containing species categorised in WoRMS to inhabit freshwater and terrestrial environments were excluded from downstream tallying. Records containing species categorised to inhabit brackish environments were retained only if they were flagged to be marine as well. The total diversity of marine gastropods at different taxonomic levels was evaluated by tallying distinct and accepted families, genera, and species (Table 1).

To examine the availability of genetic reference sequences for the resulting taxa, we mined publicly accessible online databases for complete mitochondrial genome sequences and curated COI barcodes. Metadata of all available full gastropod mitogenomes were downloaded from NCBI (query: "Gastropoda[Organism] AND mitochondrion, complete genome[Title]", 05 October 2023). Mitogenomes that are not yet integrated into the NCBI



**Fig. 1** Flowchart of the data filtering steps and the main research questions of this study

RefSeq database (i.e., from INSDC GenBank) were included in this study. COI barcode information was downloaded from the Barcode of Life Data System (BOLD) using the R package ‘bold’ v1.3.0 (Dubois and Chamberlain 2023). Percentages of families, genera, and species that are represented by at least one full mitogenome or COI barcode in the filtered, taxon-matched dataset were calculated under each gastropod order.

Records with geographic coordinates were subset for global mapping. Data points holding exactly similar values for latitudinal and longitudinal coordinates were assumed erroneous and removed ( $n=165$ ). Occurrences found in landlocked countries were seen to be clerical errors in georeferencing (e.g., GBIF country estimation from wrong geographic coordinates) and were removed from mapping. A map was produced and further land masking was performed based on world country polygons from Natural Earth using the R package ‘rnaturalearth’ v0.3.4 (Massicotte et al. 2023). Lastly, to exclude records that may potentially have erroneous specimen origins from the metadata, we removed records that are found significantly outside a species’ average geographic distribution by calculating each occurrence’s latitudinal and longitudinal absolute z-scores based on species means and standard deviations, and setting a threshold of  $z\text{-score}=3$  in both latitude and longitude. Resulting data points with  $z\text{-scores}<3$  ( $n=1,602,331$ ) were binned into  $5^\circ \times 5^\circ$  grid cells and summarised by the number of unique species present in each cell. Regions considered to be “species hotspots” (i.e., grid cells that hold higher numbers of reported species) were highlighted in the map by heuristically setting a minimum threshold of 500 species based on quantile values at Q90 (400) and Q95 (667). Found species hotspots were clustered based on the marine realms and provinces sensu Spalding et al. (2007). The barcoding references per region was also assessed by mapping percentages of species with COI from georeferenced occurrences in each grid cell.

Finally, temporal patterns in occurrence events were explored in a subset comprising records that have event dates (i.e., date-stamped dataset;  $n=2,712,886$ ). Here, data points were binned into five-year intervals. Changes in the relative contribution of the different data publishing organisations and their associated countries through time were explored. Each publishing organisation was classified into 12 categories (Table 2) sensu Groom et al.

**Table 1** Global taxonomic richness, mitogenome availability (NCBI and GenBank), and COI barcoding coverage (BOLD) of marine gastropods, covering 234 countries and territories, based on GBIF occurrence data (n = 3,904,314 records; 1662–2023)

Subclass/order	Family			Genus			Species			
	N <sub>GBIF</sub>	mt gen	%	N <sub>GBIF</sub>	COI	%	N <sub>WORMS</sub>	N <sub>GBIF</sub>	COI	%
Caenogastropoda										
Littorinimorpha	67	20	29.9	666	251	37.7	7058	5922	731	12.3
Neogastropoda	62	23	37.1	1144	379	33.1	15,608	13,014	1398	10.7
[unassigned] Caenogastropoda <sup>a</sup>	28	7	25.0	249	52	20.9	3438	2989	135	4.5
Heterobranchia										
Aplysiida	2	1	50.0	11	9	81.8	89	79	29	36.7
Cephalaspidea	24	2	8.3	93	42	45.2	847	703	134	19.1
Ellobiida	3	2	66.7	24	11	45.8	214	137	19	13.9
Nudibranchia	71	19	26.8	261	176	67.4	2616	2100	628	29.9
Pleurobranchida	2	1	50.0	10	6	60.0	92	79	32	40.5
Pteropoda	16	–	–	33	24	72.7	148	116	48	41.4
Runcinida	2	1	50.0	8	3	37.5	80	27	8	29.6
Siphonariida	1	1	100.0	9	2	22.2	108	95	19	20.0
Systellommatophora	1	1	100.0	6	6	100.0	66	36	26	72.2
Umbraculida	2	–	–	3	1	33.3	10	10	2	20.0
“Lower Heterobranchia” <sup>b</sup>	36	7	19.4	266	57	21.4	4629	3675	179	4.9
Neomphaliones										
Cocculinida	3	1	33.3	10	2	20.0	63	56	5	8.9
Neomphalida	4	2	50.0	25	12	48.0	56	46	15	32.6
Neritimorpha										
Cycloneritida	4	1	25.0	20	5	25.0	242	144	58	40.3
Patellogastropoda	9	2	22.2	37	29	78.4	353	327	168	51.4
Vetigastropoda										
Lepetellida	16	4	25.0	86	23	26.7	966	838	101	12.1
Pleurotomariida	1	1	100.0	5	4	80.0	36	35	10	28.6

**Table 1** (continued)

Subclass/order	Family		Genus		Species			
	N <sub>GBIF</sub>	mt gen %	N <sub>GBIF</sub>	COI	N <sub>WoRMS</sub>	N <sub>GBIF</sub>	COI	%
Seguenziida	11	1 9.1	77	9	708	594	12	2.0
Trochida	14	6 42.9	247	83	2564	2245	254	11.3
[unassigned] Vetigastropoda <sup>a</sup>	1	-	1	-	1	1	-	-
Gastropoda total	380	103 27.1	3291	1186	39,992	33,268	4011	12.1

<sup>a</sup>Temporary name as per MolluscaBase eds (2023) Accessed at <https://www.marinespecies.org> on 2023-10-19

<sup>b</sup>Infraclass

**Table 2** Relative contributions of the different publishing organisation categories to species occurrence records of marine gastropods in GBIF

Organisation category	Category description (adapted from Groom et al. 2017)	Occurrence records	
		N	%
Museum	Collections of preserved biodiversity specimens	2,304,528	59.03
Citizen Science Project	Projects specifically using volunteers to collect observation biodiversity data	572,370	14.66
Information Facility or Data Centre	Facilities specifically created for data aggregation and curation	329,345	8.44
Research Institution	Organisations with a specific focus on biodiversity research (includes international research projects and collaborative networks)	199,051	5.10
Governmental Institution	State-funded organisations primarily providing services to and from the state	189,067	4.84
Society	Associations of individuals, often containing a proportion of amateurs	140,450	3.60
Educational Institution	Universities, colleges, and schools	70,206	1.80
Foundation	Charitable NGO organizations	61,939	1.59
Consultancy Firm <sup>a</sup>	Organisations consulted for environmental impacts assessment and monitoring	32,770	0.84
Data Publisher	Organisations devoted to online publication and preservation of scientific data	3042	0.08
Parks Authority or Nature Reserve	Organisations responsible for the conservation and management of national parks and wildlife reserves	1371	0.04
Commercial Organisation	Organisations whose primary focus is the generation of wealth	175	< 0.01

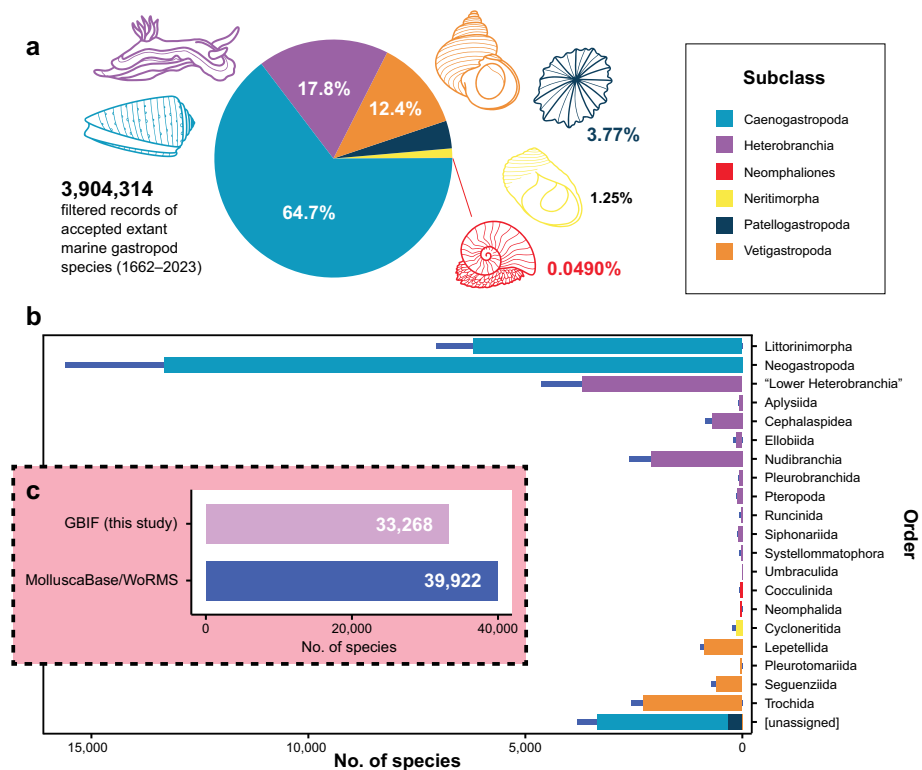
<sup>a</sup> Additional category created for this dataset

(2017). The relative contributions of each organisation category to the dataset were likewise explored.

## Results

### Systematic division of marine gastropod taxa

A filtered dataset of 3,904,314 occurrence records was obtained from GBIF, all taxon-matched and annotated with valid species names and Aphia IDs from WoRMS. The dataset consists of specimens collected or observations made from March 1662 until September 2023. Six gastropod subclasses were represented with widely varied relative occurrences (Fig. 2a). Caenogastropoda makes up the bulk of the dataset and accounts for 64.7% of the records, followed by Heterobranchia at 17.8%. Vetigastropoda accounts for 12.4% of the records and the rest are shared by Patellogastropoda (3.7%), Neritimorpha (1.2%), and Neomphaliones (<0.1%).



**Fig. 2** Systematic division of marine gastropods found in the GBIF dataset. **a** Relative abundance of occurrence records by gastropod subclass. **b** Number of species within each gastropod order found in the GBIF dataset. Thin bars represent the number of gastropod species documented in WoRMS/MolluscaBase. **c** Comparison of current totals of accepted extant marine species from WoRMS and this study's dataset from GBIF



At least 33,268 unique and valid marine species (33,987 when including subspecies, varieties, and forms) under 3291 genera belonging to 380 gastropod families were reported in the dataset (Table 1, Online Resource 1). About 64.5% of the listed species are Caenogastropoda, over half of which belong to the order Neogastropoda (Fig. 2b). Neogastropoda, also the most speciose of the orders in the dataset, accounts for 13,014 of the species. It is followed by Littorinimorpha (5922), Trochida (2245), and Nudibranchia (2100). These four most speciose orders are also the most genus-rich and account for 70% of all reported genera. Nudibranchia has the greatest number of reported families with 71, followed by Littorinimorpha, 67, and Neogastropoda with 62 families.

The total number of species represented in the GBIF dataset covers 83.3% of the current total number of accepted extant marine gastropod species curated in WoRMS ( $n=39,992$ ; Fig. 2c). By taxonomic order, the relative abundance of species in the dataset are found to generally correlate with WoRMS totals (Fig. 2b, Table 1).

### Mitogenome and COI barcode availability

We obtained a total of 1055 full mitochondrial genomes of Gastropoda from NCBI RefSeq and INSDC (GenBank) which represents 292 genera of marine, brackish, freshwater, and terrestrial gastropods. Out of the 380 reported marine families in the GBIF dataset, 103 (27.1%) have at least one species with a sequenced full mitogenome (Table 1). Fifty (48.5%) of the mitogenome-represented families are Caenogastropoda, whilst 35 (33.9%) of these are Heterobranchia. All six gastropod subclasses have at least one full mitogenome sequenced, although the Heterobranchia orders Pteropoda and Umbraculida are found to have none as of writing. Among the 33,268 reported species in the dataset, only 4011 (12.1%) have been barcoded according to BOLD (Table 1).

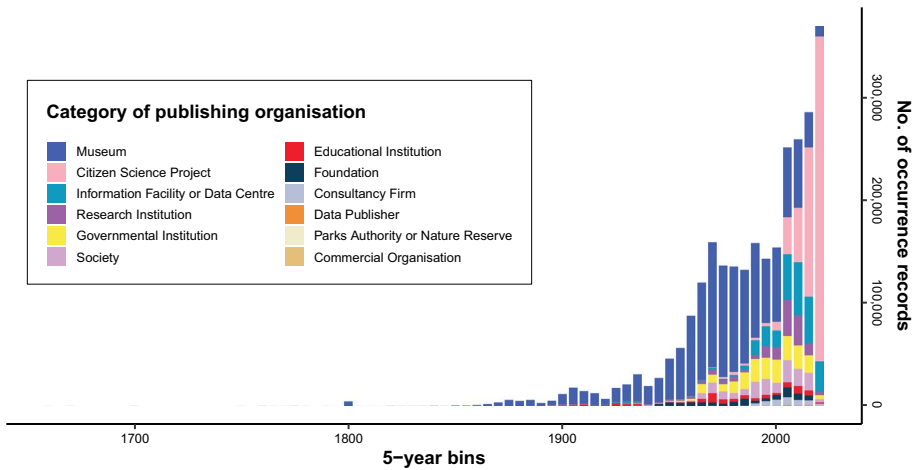
### Data sources and publishing institutions

Museums supply almost 60% of species occurrences in the filtered, taxon-matched dataset (Table 2). Until the 2010s, museums had been the major component of the recorded species occurrences (Fig. 3). However, in the last decade, GBIF submissions by citizen science platforms (e.g., iNaturalist, Seasearch) increased significantly and became the dominant source of marine gastropod observations during the period (Fig. 3, Online Resource 2).

The geo-tagged, z-limited dataset, whilst with unverified country assignments, comprises at least 199 observation countries and territories (Online Resource 3), the numbers of occurrence records in which range from two in Monaco up to 272,225 in Australia. It incorporates 1430 different datasets that were submitted by 334 publishing organisations which are registered in one of 55 countries or are produced by international initiatives (e.g., Conservation of Arctic Flora and Fauna, International Barcode of Life Consortium) (Online Resource 2).

### Reported species hotspots and COI barcoding cold spots

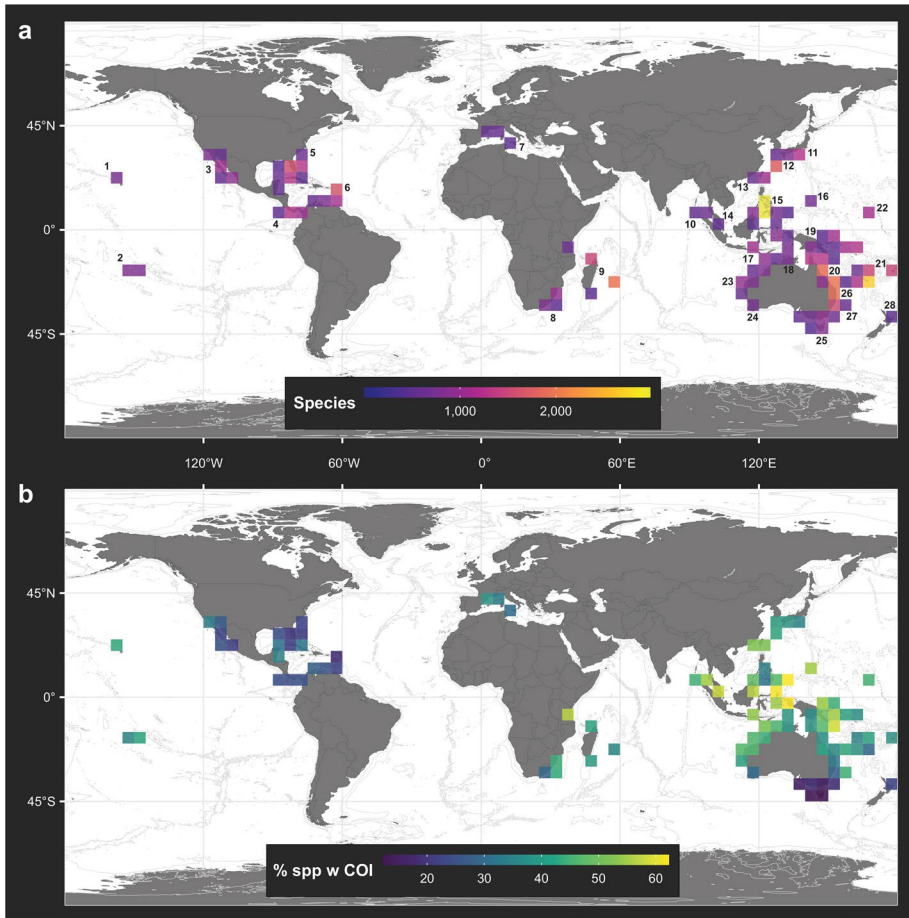
Retaining only georeferenced, land-filtered, non-outlying observations yields 21,433 distinct accepted species that are found to be widely distributed albeit disproportionate across all oceans (Online Resource 4). When mapped, non-empty grid cells are each found to contain up to 60,280 observations, with up to 2981 distinct species.



**Fig. 3** Relative contribution of the 12 categories of all publishing organisations to species occurrences recorded from 1662 to 2023, demonstrating the shift from museum-based collections towards citizen science observations

Species hotspots, defined in this study as  $5^{\circ} \times 5^{\circ}$  grid cells containing at least 500 species, were found in 93 cells and assigned into 28 marine provinces across nine marine realms (Fig. 4a, Table 3, Online Resource 5). Disparity in species numbers is observed among species hotspots, between provinces, and between realms. The greatest numbers of reported species are found in the two marine provinces that comprise the Coral Triangle, the islands of New Caledonia in the Tropical Southwestern Pacific, the Caribbean Islands of the Tropical Northwestern Atlantic, and around Madagascar and Réunion Island of the Western Indian Ocean. Reported species numbers are also high in the Tropical Eastern Pacific, in the southern end of the Kuroshio, and in the tropical to temperate range of the Eastern Australian Shelf.

Gaps in COI barcode availability were also revealed for the reported species within each cell, within hotspots, and their corresponding provinces and realms. Whilst the less species-rich, non-hotspot grid cells are fairly covered (Online Resource 6), COI/species ratios show a general shortage in barcode sequences for genetic identification of marine gastropod species. COI barcode coverage ranges from 12.3 to 62.3% ( $M = 38.0\%$ ,  $SD = 12.2\%$ ) among hotspot cells, which have 505–2981 reported species each (Fig. 4b, Online Resource 5). This general deficit in COI coverage and disparity between regions are echoed at the broader scale of marine provinces (Table 3). Hotspot provinces cumulatively hold 523–4464 species each and have COI coverage of 16.3–58.3% ( $M = 35.1\%$ ,  $SD = 10.7\%$ ). Meanwhile, disparity in COI coverage was also revealed among the more speciose provinces. A notable although counterintuitive example would be the Southeast Australian Shelf and the tropical and subtropical provinces in the western Atlantic, with 16.3–21.2% coverage, which are less likely to have a COI barcode for a given sampled species, as opposed to the even more species-rich provinces of Central Indo-Pacific, which are shown to have better COI coverage (27.9–58.3%).



**Fig. 4** Maps showing the distribution of the **a** 93 species hotspots cells based on georeferenced occurrences (numbered 1–28 by marine province) and **b** percentages of COI barcoding coverage for the recorded species within each 5°×5° species hotspot cell

## Discussion

### Distribution of marine species richness and barcoding coverage

This study provides the first baseline information on the global taxonomic richness of marine gastropods based on GBIF species occurrence records, totalling up to 33,268 valid species. Species hotspots based on the reduced z-limited dataset (comprising 21,433 species) were found to be unevenly spread throughout the globe, with 54 out of the 93 hotspots unsurprisingly concentrated between Central Indo-Pacific and Temperate Australasia whilst the rest of the provinces present 1–9 hotspot cells each. To a certain extent, these results could be reflective of actual patterns of biodiversity distribution. In general, diversity tends to increase towards the lower latitudes with various hypotheses posed (e.g., mid-domain effect, species-energy hypothesis, effect of climate stability or harshness on species persistence, faster rates of microevolution in the tropics, etc.), although why both terrestrial

**Table 3** Cumulative number of reported occurrences, total valid species, and species with COI barcodes in the 93 grid cells considered as species hotspots and grouped by marine province

Marine Realm	Marine Province	N		COI	%COI
		Obs	Spp		
Eastern Indo-Pacific	Hawai'i	9897	794	339	42.7
Temperate Northern Pacific	Southeast Polynesia	12,557	1132	395	34.9
	Warm Temperate Northeast Pacific	58,775	1754	430	24.5
Tropical Eastern Pacific	Tropical East Pacific	30,128	2385	527	22.1
Temperate Northern Atlantic	Warm Temperate Northwest Atlantic	8050	934	198	21.2
Tropical Atlantic	Tropical North-western Atlantic	147,237	2979	495	16.6
Temperate Northern Atlantic	Mediterranean Sea	29,297	897	296	33.0
Temperate Southern Africa	Agulhas	16,627	1574	539	34.2
Western Indo-Pacific	Western Indian Ocean	26,521	2672	882	33.0
Temperate Northern Pacific	Andaman	3823	1018	489	48.0
	Warm Temperate Northwest Pacific	26,985	1771	601	33.9
Central Indo-Pacific	South Kuroshio	13,846	2141	890	41.6
	South China Sea	2235	692	361	52.2
	Sunda Shelf	3085	547	319	58.3
	Western Coral Triangle	77,326	4464	1304	29.2

Table 3 (continued)

Marine Realm	Marine Province	N		COI		%COI
		Obs	Spp	Obs	COI	
16	Tropical North-western Pacific	8500	959	523		54.5
17	Northwest Australian Shelf	21,093	1695	714		42.1
18	Sahul Shelf	15,704	1766	592		33.5
19	Eastern Coral Triangle	31,396	2866	992		34.6
20	Northeast Australian Shelf	53,067	2656	884		33.3
21	Tropical South-western Pacific	59,591	3803	1059		27.8
22	Marshall, Gilbert, and Ellis Islands	5740	1127	508		45.1
23	West Central Australian Shelf	2875	672	275		40.9
24	Southwest Australian Shelf	8521	766	230		30.0
25	Southeast Australian Shelf	84,231	1530	250		16.3
26	East Central Australian Shelf	56,105	2319	692		29.8
27	Lord Howe and Norfolk Islands	3825	635	282		44.4
28	Northern New Zealand	4235	523	122		23.3

and aquatic biodiversity peak nearer the equator continues to be one of the major questions in biogeography. In the marine realm, this gradient is long-observed to peak longitudinally at the Coral Triangle, an Indo-Pacific region comprising six archipelagic nations (Hoeksema 2007). The origin of this megadiverse region has been associated with several, possibly synergistic hypotheses: as a (1) centre of speciation, (2) centre of geographic overlap of Indian and Pacific Ocean fauna, (3) centre of accumulation of expanding geographic ranges of species, or as a (4) centre of surviving old lineages (Bellwood and Meyer 2009; Barber and Meyer 2015).

The geo-tagged, z-limited dataset, however, also exhibited high disparity in the number of records distributed throughout the globe. In a broader context, in addition to showing the geographic distribution of gastropod diversity in nature, the dataset more clearly reveals the uneven levels of sampling effort made throughout history, which may have confoundingly shaped the biodiversity patterns perceived in this study. For example, as it stands, there are only nine occurrence records of marine gastropods found for Cameroon: eight of these were submitted by the Muséum National d'Histoire Naturelle (MNHN, Paris, France) pointing to occurrences of the borsoniid *Genota mitriformis* (W. Wood, 1828) and one from the zoological collections of Universität Ulm (Baden-Württemberg, Germany) for one horaclavid *Micropleurotoma melvilli* (Sykes, 1906). Given the country's 400-km coastline and knowledge of species range and ecology, it is highly implausible that Cameroon has just nine records of marine gastropods. Consequently, this means that the dataset in its current form, despite the millions of submitted records it contains, could have still been restrictive for certain countries and regions, like Cameroon, because of underreporting or an overall undersampling.

This bias in sampling effort is then further highlighted by the patchy geographic distribution not only of the overall counts of occurrence records themselves, but also of the observed number of species. With the degradation of marine habitats, backdropped by the continuously increasing ocean temperatures due to anthropogenic climate change (Hoegh-Guldberg et al. 2007; Peñaflores et al. 2009; Lough et al. 2018), it has become more urgent for the world to characterise baseline marine biodiversity for conservation in coral reefs, which, in turn, has stimulated more research in the tropical regions (Myers et al. 2000; Roberts et al. 2002). Exemplifying this are the brightest cells within the tropics, which may be traced to intensive sampling activities conducted in New Caledonia and the Philippines (Bouchet et al. 2002, 2009) where special emphasis was placed on various complementing sampling approaches to document the species richness of benthic molluscs. The work of Bouchet et al. (2002) of the MNHN detected 2738 molluscan species in New Caledonia (2187 of which were gastropods and around 90% of which were considered micromolluscs or molluscs sized 0.4–40.9 mm). Their results reflect both the potential yield from maximising sampling effort and the extent of undiscovered diversity especially within under-sampled environments and taxa (or size classes). This highly focused work can account for the steepness of some diversity gradients found in our results. Fittingly, it is observed that the MNHN group has majorly contributed to New Caledonia (68.0% of the georeferenced records) and the overall Tropical Southwestern Pacific (57.9%), the Philippines (21.7%), as well as in several other hotspot provinces where they have been involved like the Mediterranean Sea (50.4%), Western Indian Ocean (47.8%), Southeast Polynesia (33.6%), and in the Eastern Coral Triangle (23.8%). A few other hotspots could also be well-linked to where research funding is available (e.g., Japan, Australia), or to where research has been prioritised (e.g., critically important but threatened regions like the Coral Triangle), or to old natural history collections linked with colonialism (e.g., the Caribbean, Australia). That there is a deficit of locally generated occurrence records or biodiversity research in general

in previously colonised or economically poor countries has been previously observed (Fontanilla et al. 2014; Titley et al. 2017; Berba and Matias 2022) and could very well speak for itself, which only underscores the need for further, geographically systematic taxon sampling in these areas. Low-latitude regions outside the Coral Triangle or the Central Indo-Pacific may still be posed to house high levels of gastropod biodiversity and should be explored and investigated.

The important work ahead is further made evident by the limited availability of reference sequences that have been generated so far for marine gastropods. As high-throughput sequencing technologies have grown to be more efficient and affordable, so does de novo sequencing of genomes. Comparative phylogenomic analyses have all become important approaches in studying the biology and ecology of organisms. Full mitochondrial genomes have become useful data in illuminating unique genomic architecture (Knudsen et al. 2006; Grande et al. 2008; Rawlings et al. 2010; Sun et al. 2018; Ghiselli et al. 2021), in providing evolutionary insights on morphology and adaptation (Medina et al. 2011; Osca et al. 2014; Du et al. 2020), and in supporting or rejecting long-troubled phylogenetic hypotheses from within families across most subclasses to between broader taxonomic groups (Grande et al. 2002; Cunha et al. 2009; Arquez et al. 2014; Uribe et al. 2016, 2017a, b, 2019; Jiang et al. 2019; Varney et al. 2021; Sanders et al. 2021). However, current research tends to sample and sequence already extensively studied, usually economically important or public health-related taxa, or the easier-to-access freshwater and terrestrial taxa, or the congeners and conspecifics that warrant taxonomic delineation (Lopes-Lima et al. 2021). As a result, even with the huge amount of research, only a little over a quarter (27.1%) of the marine gastropod families listed in this study are represented with at least one mitogenome. As of writing, less diverse but equally important orders consisting of the pteropod and umbraculid gastropods are yet to be represented in NCBI RefSeq or in GenBank.

In terms of the availability of reference sequences for COI barcoding, whilst it is acknowledged that molluscs have become one of the most barcoded phyla among non-chordate metazoans (Kvist 2013; Mugnai et al. 2021), the enormous number of species within this group means an equivalent effort, if not much more, is needed to generate a comprehensive reference database. Since the inception of DNA barcoding as a tool for genetic identification of species (Hebert et al. 2003), it appears that only 12.1% of the total number of species of marine gastropods accounted in this study (36.0% at the genus level) have been coupled with a COI barcode. Geographically, this disparity in barcoding coverage has been made evident by the putative COI cold spots (e.g., species hotspots in the Tropical Atlantic, Tropical Eastern Pacific, and Southern Australia; Fig. 4b) where sometimes over 3600 species can be found but of which only about 17% are barcoded (Table 3). To address these barcoding resource inequities, we therefore continue to call for further taxon sampling, sequencing work, and prioritisation within the putative COI cold spot provinces. Moving forward, we also advocate for increased funding for and stronger involvement of local institutions within these regions by building capacity and actively following through open, inclusive and non-parachute collaborative biodiversity research (*sensu* Stefanoudis et al. 2021).

## Open and inclusive science can pave the way for biodiversity research

The opening up of science and allowing access to research data have provided the opportunity to review global patterns of biodiversity. Open science allows for transparency and reproducibility and stimulates collaboration. Through this, the creation of many free and

universally available biodiversity databases (e.g., GBIF), data standards (e.g., Darwin Core; Wieczorek et al. 2012), tissue and DNA repositories (e.g., Global Genome Biodiversity Network; Droege et al. 2014), and open-source tools (e.g., taxize; Chamberlain and Szöcs 2013) has been advanced. With the development of publicly accessible species occurrence repositories, data from centuries worth of biodiversity collections and research have been and are continually being made available. These can be rich sources of biological insights that will be useful for making proper ecological inferences and conservation strategies. However, despite the substantial taxonomic richness that our dataset accounts in this study, it proves to inadequately represent the diversity of marine gastropods in regions where records are lacking. Thus, the promise of an accurate and definitive global evaluation of marine gastropod diversity, though not fully realised herein, is still on its way to fulfilment.

With the continuous curation being done through WoRMS and MolluscaBase, the study was able to minimise the occurrence of outdated taxonomy uploaded in the GBIF dataset, which is most helpful and critical for taxonomically dynamic groups of organisms such as gastropods. As we also foresee an explosive increase in species occurrence records derived from metabarcoding environmental samples, the maintenance, expansion, and further integration and cross-referencing between these different open technologies will be crucial. These data points will come from slurries derived from plankton tows, settlement plates or gut (i.e., trophic) analyses, as well as the less invasive or destructive filtered water samples. The development and extension of GBIF's capacity to track and version sequence variants (Abarenkov et al. 2023) will rapidly expand occurrence records beyond the current one to one specimen:record model. The ability to put species names to these sequence records will be significantly handicapped without augmented voucher-based reference libraries. Tools and benchmarks such as those developed herein can help steer expeditionary efforts and funding resources strategically to close our knowledge gaps for documenting life on the planet.

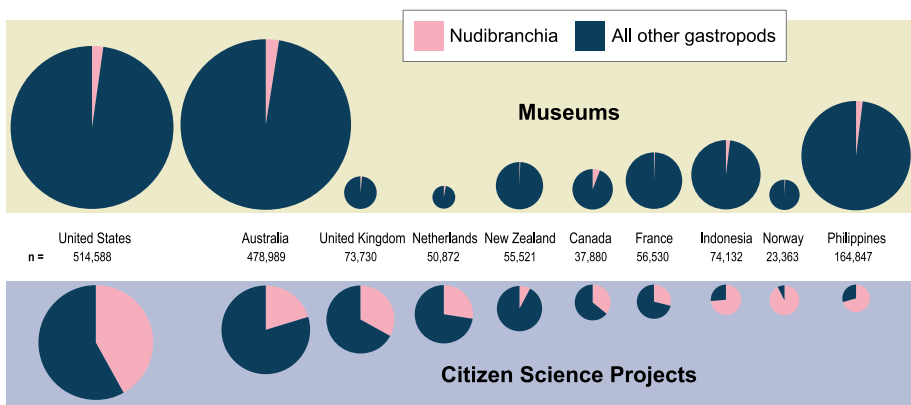
The analysis also reveals a noticeable shift in the categories of uploading institutions from museums and natural history collections towards volunteer-based organisations and biodiversity data centres that monitor and collate data in recent years. We attribute this to two main reasons—the rising involvement of citizen science and the massive shift away from specimen-based sampling in biodiversity studies—which have been some of science's responses considering the current biodiversity crisis (Troudet et al. 2018; Byrne 2023). Observation-based sampling, as opposed to collecting and keeping actual animal specimens, is viewed as a less destructive and logistically easier way of gathering, storing, and sharing diversity data and has been favoured in recent years (although counterarguments signifying the importance and continued relevance of specimen-based sampling are also discussed by Gropp (2018) and Nachman et al. 2023). Observation-based occurrences have also been the capital of participatory citizen science projects such as iNaturalist, which allows for anyone with a camera to upload an observation of an organism and have a community of avid naturalists and scientists improve the record by vouching for or amending its taxon identification. The inclusion in GBIF of research-grade contributions resulting from this process (i.e., species occurrences that are well-photographed, taxonomically vouched, and georeferenced) has indeed supplemented species coverages, particularly for gastropods, as we later found in this study. Gathering the ten countries with the highest numbers of citizen science contributions (i.e., USA, Australia, UK, Netherlands, New Zealand, Canada, France, Indonesia, Norway, Philippines) shows an increase of uploaded underwater observations in such platforms, presumably as a consequence of the recent accessibility of underwater photography among recreational and professional SCUBA divers and the



general interest for marine biota content in social media (Retka et al. 2019; Ruiz-Frau et al. 2020; Roberts et al. 2023). Nudibranchs in particular have been a favourite among underwater photographers due to the diversity of their colours, textures, and anatomy (Witabora and Homan 2021) and have had significant presence in social media platforms such as Instagram (Hoffman et al. 2022). Their rise in popularity may have been pushed further by the mainstreaming of science communication especially on biodiversity, ecology, and conservation (Burns et al. 2003; López-Goñi and Sánchez-Angulo 2018; Lamb et al. 2018; Heathcote 2021; Habibi and Salim 2021) alongside recent internet memes and videos on nudibranchs (e.g., Ze Frank 2020). We surmise that all these may have helped boost the species occurrence records of non-shelled gastropods like nudibranchs particularly in citizen science platforms, which would otherwise have been inadequately represented within museum collections that are predominated by shells (Fig. 5).

For a clearer picture of gastropod or molluscan taxonomy, diversity, and distribution, we believe that a more widely open and inclusive science will be key. It is estimated that 40% of the total morphospecies of gastropods are yet to be discovered and sampled from the wild. It has also been shown that the waning numbers of taxonomists and that the difficulties of sampling gastropods in more niche environments have become the main limiting factors in the discovery and description of more new species (Bouchet et al. 2016). In recent years, aside from the submission of species occurrence reports, the involvement of “amateur” naturalists and citizen scientists in biodiversity monitoring and research have contributed greatly to molluscan taxonomy through new species descriptions. In the study of Bouchet et al. (2016), 57% of new species descriptions from 2000 to 2014 have been first-authored by citizen scientists. This suggests that an even more active collaboration of academics and non-academics in malacology and taxonomy could only be beneficial to molluscan science, as has already been demonstrated by citizen science-led studies on biodiversity monitoring and discovery (Sneha Chandran et al. 2017; Smith and Davis 2019; Chow et al. 2022) and on geographic range expansion of both native and non-indigenous taxa (Nimbs and Smith 2018; Kleitou et al. 2019; Smith and Nimbs 2022).

The ubiquity and diversity of gastropods have made them ecologically, economically, and culturally important components of the ecosphere. Their evolutionary history, the richness of their taxa, and their skeletal records have all placed them in a unique position



**Fig. 5** Relative occurrence of nudibranch and non-nudibranch gastropod observations made in the ten countries that have the most citizen science contributions

whose potential can be made further useful in biomonitoring and research in the context of biodiversity loss amidst rapidly changing marine environments. By pinning potentially important regions of high species richness and barcode deficiency, this study hopes to guide future work and research agenda to advance more taxon sampling, further support for barcoding efforts, and the participation and empowerment of local institutions and literally anyone who can help.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10531-024-02896-9>.

**Acknowledgements** This study received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813360, through the Innovative Training Network *4D-REEF*, and from the University of Amsterdam. We thank Patrick R. Pata for his support in the analysis and for creating the gastropod iconography used in this paper.

**Author contributions** A.T., O.W., W.R., C.M., and J.T. contributed to the study conception and design. J.T., O.W., and W.R. provided supervision. Data collection, formal analysis, and visualisation were performed by A.T. The original draft of the manuscript was written by A.T. All authors commented on previous versions of the manuscript and have read and approved the final manuscript.

**Funding** This study received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 813360, through the Innovative Training Network *4D-REEF*, and from the University of Amsterdam.

**Data availability** The filtered datasets analysed and R script used in this study are openly available through the Natural History Museum's Data Portal (<https://doi.org/10.5519/vfaeq4e5>). The unprocessed dataset of all gastropod species occurrences can be downloaded from GBIF (<https://doi.org/10.15468/dl.qgsn2b>).

## Declarations

**Competing interests** The authors have no competing interests to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abarenkov K, Andersson A, Bissett A et al (2023) Publishing DNA-derived data through biodiversity data platforms. GBIF Secretariat, Copenhagen. <https://doi.org/10.35035/DOC-VF1A-NR22>
- Aktipis SW, Giribet G, Lindberg DR, Ponder WF (2008) Gastropoda: an overview and analysis. In: Ponder W, Lindberg DRR (eds) *Phylogeny and evolution of the mollusca*. University of California Press, Berkeley, pp 201–237
- Appeltans W, Ahyong ST, Anderson G et al (2012) The magnitude of global marine species diversity. *Curr Biol* 22:2189–2202. <https://doi.org/10.1016/j.cub.2012.09.036>
- Arquez M, Colgan D, Castro LR (2014) Sequence and comparison of mitochondrial genomes in the genus *Nerita* (Gastropoda: Neritimorpha: Neritidae) and phylogenetic considerations among gastropods. *Mar Genomics* 15:45–54. <https://doi.org/10.1016/j.margen.2014.04.007>







- Barber PH, Meyer CP (2015) Pluralism explains diversity in the coral triangle. In: Mora C (ed) Ecology of fishes on coral reefs. Cambridge University Press, Cambridge
- Barker GM (2002) Molluscs as crop pests. CABI, Wallingford
- Bellwood DR, Meyer CP (2009) Searching for heat in a marine biodiversity hotspot. *J Biogeogr* 36:569–576. <https://doi.org/10.1111/j.1365-2699.2008.02029.x>
- Berba CMP, Matias AMA (2022) State of biodiversity documentation in the Philippines: metadata gaps, taxonomic biases, and spatial biases in the DNA barcode data of animal and plant taxa in the context of species occurrence data. *PeerJ* 10:e13146. <https://doi.org/10.7717/peerj.13146>
- Bingham H, Doudin M, Weatherdon L et al (2017) The biodiversity informatics landscape: elements, connections and opportunities. *Res Ideas Outcomes* 3:e14059
- Blaxter M, Mann J, Chapman T et al (2005) Defining operational taxonomic units using DNA barcode data. *Philos Trans R Soc Lond B* 360:1935–1943. <https://doi.org/10.1098/rstb.2005.1725>
- Bouchet P, Lozouet P, Maestrati P, Heros V (2002) Assessing the magnitude of species richness in tropical marine environments: exceptionally high numbers of molluscs at a New Caledonia site. *Biol J Linn Soc* 75:421–436. <https://doi.org/10.1046/j.1095-8312.2002.00052.x>
- Bouchet P, Ng PK, Largo D, Tan SH (2009) PANGLAO 2004: Investigations of the marine species richness in the Philippines. *Raffles Bull Zool Supp* 20:1–19
- Bouchet P, Bary S, Virginie H, Marani G (2016) How many species of molluscs are there in the world's oceans, and who is going to describe them? *Mem Mus Natl Hist Nat* 208:9–24
- Bouchet P, Rocroi J-P, Hausdorf B et al (2017) Revised classification, nomenclator and typification of gastropod and Monoplacophoran families. *Malacologia* 61:1–526. <https://doi.org/10.4002/040.061.0201>
- Bouchet P, Decock W, Lonneville B et al (2023) Marine biodiversity discovery: the metrics of new species descriptions. *Front Mar Sci*. <https://doi.org/10.3389/fmars.2023.929989>
- Burns TW, O'Connor DJ, Stocklmayer SM (2003) Science communication: a contemporary definition. *Public Underst Sci* 12:183–202. <https://doi.org/10.1177/09636625030122004>
- Byrne AQ (2023) Reimagining the future of natural history museums with compassionate collection. *PLoS Biol* 21:e3002101. <https://doi.org/10.1371/journal.pbio.3002101>
- Chamberlain SA, Szöcs E (2013) taxize: taxonomic search and retrieval in R. *F1000Res* 2:191. <https://doi.org/10.12688/f1000research.2-191.v2>
- Chamberlain S, Vanhoorne B (2023) worms: World Register of marine species (WoRMS) Client
- Chow LH, Yu VPF, Kho ZY et al (2022) An updated checklist of sea slugs (Gastropoda, Heterobranchia) from Hong Kong supported by citizen science. *Zool Stud* 61:e52. <https://doi.org/10.6620/ZS.2022.61-52>
- Costello MJ (2009) Motivating online publication of data. *Bioscience* 59:418–427. <https://doi.org/10.1525/bio.2009.59.5.9>
- Cunha RL, Grande C, Zardoya R (2009) Neogastropod phylogenetic relationships based on entire mitochondrial genomes. *BMC Evol Biol* 9:210. <https://doi.org/10.1186/1471-2148-9-210>
- Cunha TJ, Fernández-Simón J, Petrula M et al (2023) Photographic checklist, DNA barcoding, and new species of sea slugs and snails from the Faafu atoll, Maldives (Gastropoda: Heterobranchia and Vetigastropoda). *Diversity* 15:219. <https://doi.org/10.3390/d15020219>
- Dang VT, Benkendorff K, Green T, Speck P (2015) Marine snails and slugs: a great place to look for antiviral drugs. *J Virol* 89:8114–8118. <https://doi.org/10.1128/jvi.00287-15>
- Dickinson JL, Shirk J, Bonter D et al (2012) The current state of citizen science as a tool for ecological research and public engagement. *Front Ecol Environ* 10:291–297. <https://doi.org/10.1890/110236>
- Droege G, Barker K, Astrin JJ et al (2014) The global genome biodiversity network (GGBN) data portal. *Nucleic Acids Res* 42:D607–D612. <https://doi.org/10.1093/nar/gkt928>
- Du L, Cai S, Liu J et al (2020) The complete mitochondrial genome of a cold seep gastropod *Phymorhynchus buccinoides* (Neogastropoda: Conoidea: Raphitomidae). *PLoS ONE* 15:e0242541. <https://doi.org/10.1371/journal.pone.0242541>
- Dubois S, Chamberlain S (2023) bold: interface to bold systems API
- Duda TF, Bolin MB, Meyer CP, Kohn AJ (2008) Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. *Mol Phylogenet Evol* 49:867–876. <https://doi.org/10.1016/j.ympev.2008.08.009>
- Edwards JL (2004) Research and societal benefits of the global biodiversity information facility. *Bioscience* 54:485–486. [https://doi.org/10.1641/0006-3568\(2004\)054\[0486:RASBOT\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0486:RASBOT]2.0.CO;2)
- Escribano N, Galicia D, Ariño AH (2018) The tragedy of the biodiversity data commons: a data impediment creeping nigher? *Database* 2018:1–6. <https://doi.org/10.1093/database/bay033>
- Fontanilla IKC, Torres AF, Cañasa JADG et al (2014) State of animal DNA barcoding in the Philippines: a review of COI sequencing of Philippine native fauna. *Philipp Sci Lett* 7:104–137

- GBIF.org (2023) GBIF occurrence download. <https://doi.org/10.15468/dl.qgsn2b>
- Ghiselli F, Gomes-dos-Santos A, Adema CM et al (2021) Molluscan mitochondrial genomes break the rules. *Philos Trans R Soc Lond B* 376:20200159. <https://doi.org/10.1098/rstb.2020.0159>
- Giannelli A, Cantacessi C, Colella V et al (2016) Gastropod-borne helminths: a look at the snail-parasite interplay. *Trends Parasitol* 32:255–264. <https://doi.org/10.1016/j.pt.2015.12.002>
- Giribet G (2008) Assembling the lophotrochozoan (=spiralian) tree of life. *Philos Trans R Soc Lond B* 363:1513–1522. <https://doi.org/10.1098/rstb.2007.2241>
- Golding RE, Bieler R, Rawlings TA, Collins TM (2014) Deconstructing *Dendropoma*: a systematic revision of a World-Wide Worm-Snail Group, with descriptions of new genera (Caenogastropoda: Vermetidae). *Malacologia* 57:1–97. <https://doi.org/10.4002/040.057.0103>
- Grande C, Templado J, Lucas Cervera J, Zardoya R (2002) The complete mitochondrial genome of the nudibranch *Roboastra europaea* (Mollusca: Gastropoda) supports the monophyly of opisthobranchs. *Mol Biol Evol* 19:1672–1685. <https://doi.org/10.1093/oxfordjournals.molbev.a003990>
- Grande C, Templado J, Zardoya R (2008) Evolution of gastropod mitochondrial genome arrangements. *BMC Evol Biol* 8:61. <https://doi.org/10.1186/1471-2148-8-61>
- Groom Q, Weatherdon L, Geijzendorffer IR (2017) Is citizen science an open science in the case of biodiversity observations? *J Appl Ecol* 54:612–617. <https://doi.org/10.1111/1365-2664.12767>
- Gropp RE (2018) Specimens, collections, and tools for future biodiversity-related research. *BioScience* 68:3–4. <https://doi.org/10.1093/biosci/bix155>
- Habibi SA, Salim L (2021) Static vs. dynamic methods of delivery for science communication: a critical analysis of user engagement with science on social media. *PLoS ONE* 16:e0248507. <https://doi.org/10.1371/journal.pone.0248507>
- Heathcote G (2021) Animals of Instagram: taxonomic bias in science communication online. *JCOM J Sci Commun*. <https://doi.org/10.22323/2.20040210>
- Hebert PD, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond B* 270:313–321
- Hoegh-Guldberg O, Mumby PJ, Hooten AJ et al (2007) Coral reefs under rapid climate change and ocean acidification. *Science* 318:1737–1742. <https://doi.org/10.1126/science.1152509>
- Hoeksema BW (2007) Delineation of the Indo-Malayan Centre of Maximum Marine Biodiversity: the coral triangle. In: Renema W (ed) *Biogeography, time, and place: distributions, barriers, and islands*. Springer, Dordrecht, pp 117–178
- Hoffman N, Christa G, Bleidibél S (2022) Using citizen science in the biodiversity research of Nudibranchia. In: *Spixiana Supplement 30 A - World Congress of Malacology Munich 2022*. Bavarian State Collection of Zoology, Munich
- Jiang D, Zheng X, Zeng X et al (2019) The complete mitochondrial genome of *Harpago chiragra* and *Lambis lambis* (Gastropoda: Stromboidea): implications on the Littorinimorpha phylogeny. *Sci Rep* 9:17683. <https://doi.org/10.1038/s41598-019-54141-x>
- Kantor YI, Fedosov AE, Kosyan AR et al (2022) Molecular phylogeny and revised classification of the Buccinoidea (Neogastropoda). *Zool J Linn Soc* 194:789–857. <https://doi.org/10.1093/zoolinnean/zlab031>
- Kleitou P, Giovas I, Wolf W, Crocetta F (2019) On the importance of citizen-science: the first record of *Goniobranchus obsoletus* (Rüppell & Leuckart, 1830) from Cyprus (Mollusca: Gastropoda: Nudibranchia). *Bioinvasions Rec* 8:252–257. <https://doi.org/10.3391/bir.2019.8.2.06>
- Knudsen B, Kohn AB, Nahir B et al (2006) Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: conservation of the gene order in Euthyneura. *Mol Phylogenet Evol* 38:459–469. <https://doi.org/10.1016/j.ympev.2005.08.017>
- Kosmala M, Wiggins A, Swanson A, Simmons B (2016) Assessing data quality in citizen science. *Front Ecol Environ* 14:551–560. <https://doi.org/10.1002/fee.1436>
- Kvist S (2013) Barcoding in the dark? A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Mol Phylogenet Evol* 69:39–45. <https://doi.org/10.1016/j.ympev.2013.05.012>
- Lamb CT, Gilbert SL, Ford AT (2018) Tweet success? Scientific communication correlates with increased citations in ecology and conservation. *PeerJ* 6:e4564. <https://doi.org/10.7717/peerj.4564>
- Lopes-Lima M, Riccardi N, Urbanska M et al (2021) Major shortfalls impairing knowledge and conservation of freshwater molluscs. *Hydrobiologia* 848:2831–2867. <https://doi.org/10.1007/s10750-021-04622-w>
- López-Goñi I, Sánchez-Angulo M (2018) Social networks as a tool for science communication and public engagement: focus on Twitter. *FEMS Microbiol Lett* 365:fnx246. <https://doi.org/10.1093/femsle/fnx246>
- Lough JM, Anderson KD, Hughes TP (2018) Increasing thermal stress for tropical coral reefs: 1871–2017. *Sci Rep* 8:6079. <https://doi.org/10.1038/s41598-018-24530-9>

- Maldonado C, Molina CI, Zizka A et al (2015) Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? *Glob Ecol Biogeogr* 24:973–984. <https://doi.org/10.1111/geb.12326>
- Massicotte P, South A, Hufkens K (2023) *natural earth*: World Map Data from Natural Earth
- Medina M, Lal S, Vallès Y et al (2011) Crawling through time: transition of snails to slugs dating back to the Paleozoic, based on mitochondrial phylogenomics. *Mar Genomics* 4:51–59. <https://doi.org/10.1016/j.margen.2010.12.006>
- Meyer CP (2003) Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biol J Linn Soc Lond* 79:401–459. <https://doi.org/10.1046/j.1095-8312.2003.00197.x>
- Mugnai F, Megléc E, Abbiati M et al (2021) Are well-studied marine biodiversity hotspots still blackspots for animal barcoding? *Glob Ecol Conserv* 32:e01909. <https://doi.org/10.1016/j.gecco.2021.e01909>
- Múrria C, Väisänen LOS, Somma S et al (2020) Towards an Iberian DNA barcode reference library of freshwater macroinvertebrates and fishes. *Linn* 39:73–92
- Myers N, Mittermeier RA, Mittermeier CG et al (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858
- Nachman MW, Beckman EJ, Bowie RC, et al (2023) Specimen collection is essential for modern science. *PLoS Biol* 21:e3002318. <https://doi.org/10.1371/journal.pbio.3002318>
- Nelson G, Ellis S (2019) The history and impact of digitization and digital data mobilization on biodiversity research. *Philos Trans R Soc Lond B* 374:20170391. <https://doi.org/10.1098/rstb.2017.0391>
- Nimbs MJ, Smith SDA (2018) Beyond Capricornia: tropical sea slugs (Gastropoda, Heterobranchia) extend their distributions into the Tasman Sea. *Diversity (base)* 10:99. <https://doi.org/10.3390/d10030099>
- Olivera BM, Showers Corneli P, Watkins M, Fedosov A (2014) Biodiversity of cone snails and other venomous marine gastropods: evolutionary success through neuropharmacology. *Annu Rev Anim Biosci* 2:487–513. <https://doi.org/10.1146/annurev-animal-022513-114124>
- Osca D, Templado J, Zardoya R (2014) The mitochondrial genome of *Ifremeria nautilei* and the phylogenetic position of the enigmatic deep-sea Abyssochrysoidea (Mollusca: Gastropoda). *Gene* 547:257–266. <https://doi.org/10.1016/j.gene.2014.06.040>
- Parkhaev PY (2008) The early Cambrian radiation of Mollusca. *Phylogeny Evol Mollusca* 33:69
- Peñaflo EL, Skirving WJ, Strong AE et al (2009) Sea-surface temperature and thermal stress in the Coral Triangle over the past two decades. *Coral Reefs* 28:841–850. <https://doi.org/10.1007/s00338-009-0522-8>
- Ponder WF, Lindberg DR (2008) *Phylogeny and evolution of the Mollusca*. University of California Press, Berkeley
- Ponder WF, Lindberg DR, Ponder JM (2019) *Biology and evolution of the Mollusca*, vol 1. CRC Press, Boca Raton
- Puillandre N, Kantor YuI, Sysoev A et al (2011) The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). *J Molluscan Stud* 77:259–272. <https://doi.org/10.1093/mollus/eyr015>
- Puillandre N, Bouchet P, Duda TF et al (2014) 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, Duda TF, Meyer C et al (2015) One, four or 100 genera? A new classification of the cone snails. *J Molluscan Stud* 81:1–23. <https://doi.org/10.1093/mollus/eyu055>
- R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Rawlings TA, MacInnis MJ, Bieler R et al (2010) Sessile snails, dynamic genomes: gene rearrangements within the mitochondrial genome of a family of caenogastropod molluscs. *BMC Genomics* 11:440. <https://doi.org/10.1186/1471-2164-11-440>
- Retka J, Jepson P, Ladle RJ et al (2019) Assessing cultural ecosystem services of a large marine protected area through social media photographs. *Ocean Coast Manag* 176:40–48. <https://doi.org/10.1016/j.ocecoaman.2019.04.018>
- Roberts CM, McClean CJ, Veron JEN et al (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295:1280–1284. <https://doi.org/10.1126/science.1067728>
- Roberts CJ, Vergés A, Poore AGB (2023) A new resource for monitoring reef ecosystems: the background of recreational diver photographs contains valuable habitat data. *J Appl Ecol*. <https://doi.org/10.1111/1365-2664.14518>
- Ruiz-Frau A, Ospina-Alvarez A, Villasante S et al (2020) Using graph theory and social media data to assess cultural ecosystem services in coastal areas: Method development and application. *Ecosyst Serv* 45:101176. <https://doi.org/10.1016/j.ecoser.2020.101176>

- Sanders MT, Merle D, Laurin M et al (2021) Raising names from the dead: a time-calibrated phylogeny of frog shells (Bursidae, Tonnoidea, Gastropoda) using mitogenomic data. *Mol Phylogenet Evol* 156:107040. <https://doi.org/10.1016/j.ympev.2020.107040>
- Schmidt TSB, Matias Rodrigues JF, von Mering C (2015) Limits to robustness and reproducibility in the demarcation of operational taxonomic units. *Environ Microbiol* 17:1689–1706. <https://doi.org/10.1111/1462-2920.12610>
- Smith SDA, Nimbs MJ (2022) Citizen scientists record significant range extensions for tropical sea slug species in subtropical Eastern Australia. *Diversity (basel)* 14:244. <https://doi.org/10.3390/d14040244>
- Smith SDA, TomR D (2019) Slugging it out for science: volunteers provide valuable data on the diversity and distribution of heterobranch sea slugs. *Molluscan Res* 39:214–223. <https://doi.org/10.1080/13235818.2019.1594600>
- Sneha Chandran BK, Shrinivaasu S, Ravinesh R et al (2017) Opisthobranch (Mollusca: Gastropoda) fauna of Kerala, India: a citizen science initiative. *J Mar Bio Assoc India* 59:49–58. <https://doi.org/10.6024/jmbai.2017.59.1.1970-08>
- Spalding MD, Fox HE, Allen GR et al (2007) Marine ecoregions of the World: a bioregionalization of coastal and shelf areas. *Bioscience* 57:573–583. <https://doi.org/10.1641/B570707>
- Stefanoudis PV, Licuanan WY, Morrison TH et al (2021) Turning the tide of parachute science. *Curr Biol* 31:R184–R185. <https://doi.org/10.1016/j.cub.2021.01.029>
- Sun S, Li Q, Kong L, Yu H (2018) Multiple reversals of strand asymmetry in molluscs mitochondrial genomes, and consequences for phylogenetic inferences. *Mol Phylogenet Evol* 118:222–231. <https://doi.org/10.1016/j.ympev.2017.10.009>
- Thaler AD, Amon D (2019) 262 Voyages Beneath the Sea: a global assessment of macro- and megafaunal biodiversity and research effort at deep-sea hydrothermal vents. *PeerJ* 7:e7397. <https://doi.org/10.7717/peerj.7397>
- Titley MA, Snaddon JL, Turner EC (2017) Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. *PLoS ONE* 12:e0189577. <https://doi.org/10.1371/journal.pone.0189577>
- Troutet J, Vignes-Lebbe R, Grandcolas P, Legendre F (2018) The increasing disconnection of primary biodiversity data from specimens: how does it happen and how to handle it? *Syst Biol* 67:1110–1119. <https://doi.org/10.1093/sysbio/syy044>
- Uribe JE, Colgan D, Castro LR et al (2016) Phylogenetic relationships among superfamilies of Neritimorpha (Mollusca: Gastropoda). *Mol Phylogenet Evol* 104:21–31. <https://doi.org/10.1016/j.ympev.2016.07.021>
- Uribe JE, Puillandre N, Zardoya R (2017a) 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>
- Uribe JE, Williams ST, Templado J et al (2017b) Denser mitogenomic sampling improves resolution of the phylogeny of the superfamily Trochoidea (Gastropoda: Vetigastropoda). *J Molluscan Stud* 83:111–118. <https://doi.org/10.1093/mollus/eyw049>
- Uribe JE, Irisarri I, Templado J, Zardoya R (2019) New Patellogastropod mitogenomes help counteracting long-branch attraction in the deep phylogeny of gastropod mollusks. *Mol Phylogenet Evol* 133:12–23. <https://doi.org/10.1016/j.ympev.2018.12.019>
- Varney RM, Brenzinger B, Malaquias MAE et al (2021) Assessment of mitochondrial genomes for heterobranch gastropod phylogenetics. *BMC Ecol Evol* 21:6. <https://doi.org/10.1186/s12862-020-01728-y>
- Wieczorek J, Bloom D, Guralnick R et al (2012) Darwin core: an evolving community-developed biodiversity data standard. *PLoS ONE* 7:e29715. <https://doi.org/10.1371/journal.pone.0029715>
- Witabora J, Homan DK (2021) Visual exploration through visual communication design work creation as an introduction attempt of marine biodiversity phase 2—Nudibranchs. *IOP Conf Ser Earth Environ Sci* 794:012108. <https://doi.org/10.1088/1755-1315/794/1/012108>
- Yesson C, Brewer PW, Sutton T et al (2007) How global is the global biodiversity information facility? *PLoS ONE* 2:e1124. <https://doi.org/10.1371/journal.pone.0001124>
- Zapata F, Wilson NG, Howison M et al (2014) Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc R Soc Lond B* 281:20141739. <https://doi.org/10.1098/rspb.2014.1739>
- Ze Frank (2020) True Facts: Freaky Nudibranchs. <https://youtu.be/F7V8DRfZBQI>. Accessed 19 Oct 2023

## Authors and Affiliations

Andrew F. Torres<sup>1,2</sup>  · Owen S. Wangensteen<sup>3</sup>  · Willem Renema<sup>2,4</sup>  ·  
Christopher P. Meyer<sup>5</sup>  · Ian Kendrick C. Fontanilla<sup>6</sup>  · Jonathan A. Todd<sup>1</sup> 

✉ Andrew F. Torres  
a.torres@nhm.ac.uk

<sup>1</sup> Natural History Museum, London, UK

<sup>2</sup> Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands

<sup>3</sup> Department of Evolutionary Biology, Ecology and Environmental Sciences and Biodiversity Research Institute (IRBIO), University of Barcelona, Barcelona, Spain

<sup>4</sup> Naturalis Biodiversity Center, Leiden, Netherlands

<sup>5</sup> Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

<sup>6</sup> Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines