Data in Brief 66 (2026) 112766

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

An expert-curated dataset of glass sponges (Porifera, Hexactinellida) distribution across the Atlantic and Arctic Oceans

Celso Domingos[a,b,c,*], Ana Sofia Soares[a,b], Andreu Santín[a,c], Joana R. Xavier[a,c,d]

[a] CIIMAR/CIMAR LA – Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Matosinhos, Portugal
[b] Faculty of Science, University of Porto, Porto, Portugal
[c] IUCN Sponge Specialist Group, IUCN Species Survival Commission, Gland, Switzerland
[d] Department of Natural History, University Museum of Bergen, University of Bergen, Norway

ARTICLE INFO

Article history:
Received 13 May 2025
Revised 11 March 2026
Accepted 8 April 2026
Available online 15 April 2026

Dataset link: An expert-curated dataset of glass sponges (Porifera, Hexactinellida) distribution across the Atlantic and Arctic Oceans. (Reference data)

Keywords:
Marine biodiversity
Species distribution
Deep-sea ecosystems
Occurrence records
Biogeographic information

ABSTRACT

Glass sponges (Porifera, Hexactinellida) are key habitat-forming organisms in deep-sea ecosystems, yet their diversity and distribution remains poorly understood. This work presents the first comprehensive and expert-curated dataset of Hexactinellida distribution records compiled from published literature and key biodiversity data repositories, including the Ocean Biodiversity Information System (OBIS), the Global Biodiversity Information Facility (GBIF), NOAA's National Database for Deep-sea Corals and Sponges (NOAA) and the Vulnerable Marine Ecosystems database of the International Council for the Exploration of the Sea (ICES VME database). The dataset was carefully curated and standardized following Darwin Core Standards to ensure consistency, accuracy and reliability in future research applications. The initial compilation included 28563 records, which were subjected to a rigorous data-cleaning process to remove duplicates, fossil records, taxonomically uncertain entries, and records lacking geographic coordinates. The final dataset comprises 4136 unique records representing 153 species (ca. 21% of all known hexactinellid species worldwide) across the Atlantic and Arctic Oceans. Notably, the

* Corresponding author.
E-mail address: bio.celso.domingos@gmail.com (C. Domingos).

https://doi.org/10.1016/j.dib.2026.112766

dataset contains distribution records of 214 type specimens. Species occurrences are predominantly concentrated in the North Atlantic, with significant data gaps in the South Atlantic and particularly along the African coastline. Depth distribution analyses reveal that glass sponges are most commonly recorded between 800 and 1200 m, with peaks in species diversity occurring at depths of 500, 1500, 2000 and 2500 m. This dataset provides valuable baseline information to support conservation efforts, biogeographic modelling, and deep-sea ecological studies. By integrating expert-validated species records from multiple sources, it offers a crucial resource for advancing knowledge on the diversity and distribution of glass sponges and their role in marine ecosystems. The dataset is publicly available through Zenodo.

Specifications Table

Subject	Biology
Specific subject area	Glass sponge's species-level records from published literature and major online repositories.
Type of data	Table (.csv format) and multiple-tables (.xlsx format).
Data collection	The data for this study on glass sponges were collected following a Data Framework, comprising three main phases: **Literature Survey (LS)**: A systematic review of published studies to extract georeferenced species distribution records. **Database Compilation (DC)**: Aggregation of occurrence data from major online repositories, ensuring compliance with FAIR principles. **Data Harmonization (DH)**: Standardization and integration of data from different sources, aligning spatial and taxonomic information.
Data source location	The dataset of glass sponges' distribution was compiled from the following biodiversity information sources: (1) Ocean Biodiversity Information System (https://obis.org) (2) Global Biodiversity Information Facility (https://www.gbif.org) (3) National Database for Deep-Sea Corals and Sponges, National Oceanic and Atmospheric Administration (https://www.ncei.noaa.gov/maps/deep- sea-corals/mapSites.htm) (4) Vulnerable Marine Ecosystems database, International Council for the Exploration of the Sea (https://vme.ices.dk/download.aspx) (5) Comprehensive bibliographic survey based on the World Porifera Database (https://www.marinespecies.org/porifera/)
Data accessibility	Repository name: Zenodo. Data identification number: 10.5281/zenodo.17575711 Direct URL to data: https://doi.org/10.5281/zenodo.17575711

1. Value of the Data

- **Filling knowledge gaps** – The dataset provides essential information on the diversity and distribution of glass sponges across the Atlantic and Arctic Oceans, a group that is still understudied.
- **Expert curated** – The dataset was carefully compiled and validated by experts in sponge taxonomy, ensuring the accuracy and reliability of species identifications and the standardization of information used in future research.
- **Support conservation and environmental management** – By mapping the occurrence of these species, the data supports the identification of critical areas for conservation and guide protection strategies for these vulnerable ecosystems.

- **Basis for biogeographic modelling** – The integration of georeferenced records allows the development of predictive models that help to understand the environmental factors that influence the distribution of glass sponges.
- **Facilitation of future studies** – The dataset compiles information from multiple sources, including published literature and global databases, making it a valuable resource for researchers studying marine biodiversity and deep-sea ecology.

2. Background

The deep sea, one of the largest yet least explored ecosystems on Earth, harbors a remarkable diversity of life despite its extreme conditions [1]. Among the many organisms shaping these ecosystems, glass sponges (Class Hexactinellida Schmidt, 1870) play a crucial ecological role. These animals act as key habitat builders, providing shelter and nursery grounds for various species [2]. Currently, 710 valid species of glass sponges have been identified, representing approximately 7.28% of all known sponge diversity [3]. They occur in all oceans, from shallow waters to deep-sea habitats, but are predominantly found in deep waters, with only a few occurrences in shallow regions [4]. The lack of baseline data on sponge diversity and distribution hinders conservation efforts and ecosystem management strategies [5]. Without a clear understanding of their population trends and ecological functions, it is nearly impossible to assess the risks they face or implement protective measures. By integrating species-level records from published literature and major online repositories this dataset will provide a valuable resource for future research on biogeography, diversity and conservation of deep-sea sponge habitats.

3. Data Description

The dataset presents an extensive compilation of distribution records of sponges of the Class Hexactinellida from the Atlantic and Arctic Ocean [6]. These records come from scientific articles from a variety of research fields, such as taxonomy, ecology, phylogeny and systematics, and from the main online biodiversity data repositories such as OBIS, GBIF and NOAA. The dataset is provided in .csv format and columns are standardized with the data fields of Darwin Core Standards (Table 1), which is divided into three parts: taxonomy, collection and references. The taxonomic section provides a comprehensive listing of all names assigned to the taxon across various classification levels. It includes the authorship details, the unique Aphia ID, and the corresponding voucher code, ensuring accurate identification and reference. The collection section contains detailed information regarding the collection of the recorded specimen. This includes precise geographical coordinates, depth measurements, collection dates, and the specific locality where the specimen was found. These data points help contextualize the specimen's environmental background. The final section presents additional associated information relevant to the record. This includes bibliographic references, citations to related studies, and supporting documentation that establish the basis of the record. These references provide scientific context, ensuring that the record is well-documented and traceable for further research.

The dataset documenting the distribution records of the class Hexactinellida in the Atlantic Ocean and the Arctic Ocean is available in three document files.

The first file - [01_raw_data]: contains the raw, unprocessed distribution data for the class Hexactinellida as originally retrieved from four major marine biodiversity platforms: the Ocean Biodiversity Information System (OBIS), the Global Biodiversity Information Facility (GBIF), NOAA's National Database for Deep-sea Corals and Sponges (NOAA) and the Vulnerable Marine Ecosystems database, International Council for the Exploration of the Sea (ICES VME database). This file is organised into the following sections:

Table 1
Data fields of the dataset containing distribution records of the class Hexactinellida in the Atlantic Ocean and the Arctic Ocean.

	Label	Definition	Example
Taxonomic Information	phylum	The full scientific name of the phylum.	Porifera
	class	The full scientific name of the class.	Hexactinellida
	subclass	The full scientific name of the subclass.	Hexasterophora
	order	The full scientific name of the order.	Lyssacinosida
	family	The full scientific name of the family.	Rossellidae
	subfamily	The full scientific name of the subfamily.	Rossellinae
	genus	The full scientific name of the genus.	*Caulophacus*
	subgenus	The full scientific name of the subgenus.	*Caulophacus*
	specificEpithet	The full scientific name of the specific epithet.	*abyssalis*
	scientificNameAuthorship	Authorship information for the acceptedName	Tabachnick, 1990
	acceptedName	Accepted name of the taxon, retrieved from the World Register of Marine Species	*Caulophacus* (*Caulophacus*) *abyssalis* Tabachnick, 1990
	TaxonID	Identifier of the taxon, linked to the World Register of Marine Species (Aphia ID).	172049
	materialSampleID	A unique identifier for a material sample, such as a museum voucher code, specimen number, or catalog identifier.	IORAS 5/2/541
	typeStatus	The designation of the type of specimen associated with the taxon, such as holotype, paratype, lectotype, or syntype.	yes/no
	typeInStudyArea	The type locality of the species is in the Atlantic Ocean or in the Arctic Ocean.	yes/no
Sampling Information	decimalLatitude	Geographical latitude in decimal degrees of the record's location	-36.215
	decimalLongitude	Geographical longitude in decimal degrees of the record's location	-49.161
	depthInMeters	The depth in meters or the centroid depth of a range of depth below the local surface, in meters.	4625
	minimumDepthInMeters	The lesser depth of a range of depth below the local surface, in meters.	4620
	maximumDepthInMeters	The greater depth of a range of depth below the local surface, in meters.	4630

(*continued on next page*)

Table 1 (*continued*)

	Label	Definition	Example
	day	The integer day of the month on which the event occurred.	8
	month	The integer month on which the event occurred.	3
	year	The integer year on which the event occurred.	1963
	eventDate	The date-time or interval during which an event occurred.	08/03/1963
	waterBody	The name of the water body in which the location occurs. Limits of ocean and sea 4th edition	Indian Ocean; Baltic Sea; Lago Nahuel Huapi
	country	The name of the country or major administrative unit in which the location occurs.	Denmark; Colombia; España
	locality	The specific description of the place.	Bariloche, 25 km NNE via Ruta Nacional 40 (=Ruta 237)
Reference Information	basisOfRecord	The specific nature of the data record.	MaterialEntity, PreservedSpecimen, LivingSpecimen, HumanObservation, MachineObservation
	bibliographicCitation	Bibliographic reference of the record	Tabachnick, K.R. (1990). Hexactinellid sponges from south-east Atlantic Ocean. Trudÿ of the Institute of Oceanology Academy of Sciences of the USSR. 126: 67–73.
	associatedReferencesType	The type of reference associated with the occurrence.	Literature; online database; Museum Collection;
	NOTES	Any important notes regarding the record.	NA

- Original Data Summary Pages: dedicated pages for each of the four data sources (OBIS, GBIF, NOAA and ICES VME database), presenting the unmodified raw records obtained from each platform.
- Excluded Data pages: pages containing records removed during the initial screening phase. These include (i) duplicate records identified within the same platform, and (ii) records excluded for not meeting the study criteria, such as entries reported above the species taxonomic rank (e.g., genus or family), fossil records or records lacking essential information such as geographic coordinates. The excluded data are retained in this file for transparency and reproducibility.

The second file - [02_processed_data]: is structured into three distinct sections, each providing critical information about the data collection and organisation process:

- Records from Literature Sources – contains all distribution records obtained exclusively from scientific publications, i.e. books and research articles.
- Records from Online Databases – includes all records that were sourced directly from the four online repositories (OBIS, GBIF, NOAA and ICES VME). These correspond to the same original data sources presented in the "original data summary pages" of the first file ([01_raw_data]), but here the data have been standardized and harmonised for further processing.
- De-replication Phase Records – includes records that were identified and processed during the de-replication phase, ensuring that duplicate entries were removed, and that data accuracy and integrity were maintained.

The last file - [03_Hexactinellida_dataset]: contains the Final Consolidated List – This document presents the compilation of unique distribution records, integrating data from both published literature and online databases. It represents the refined dataset after processing and validation.

The initial dataset, compiled from literature sources and online databases, contained **28563 records of** glass sponges. After applying a series of data curation steps including the removal of: i) duplicate records; ii) entries with coordinates on land; iii) records lacking geographic coordinates; iv) fossil records; and v) taxonomically uncertain entries; the final dataset was reduced to **4136 records r**epresenting **153 species** across the Atlantic and Arctic Oceans (Fig. 1a), which represents 21% of all known hexactinellid species worldwide. The final dataset encompasses a comprehensive classification of deep-sea hexactinellid sponges, including the records of 214 specimens (holotype, paratype, etc.) considered to be type specimens (Fig. 1b). The dataset includes records of the order Amphidiscosida Schrammen, 1924 (subclass Amphidiscophora Schulze, 1886; Fig. 1c) and within the subclass Hexasterophora Schulze, 1886, records from the *incertae sedis* Hexactinellida family Dactylocalycidae Gray, 1867, five genera of uncertain affinity (Fig. 1d) and the orders Lyssacinosida (Fig. 1e, Sceptrulophora (Fig. 1d) and Lychniscosida (Fig. 1d).

Among these, Sceptrulophora is the order with the largest number of families, comprising six families: Farreidae Gray, 1872, Euretidae Zittel, 1877, Tretodictyidae Schulze, 1886, Auloplacidae Schrammen, 1912, Aphrocallistidae Gray, 1867, and Fieldingiidae Tabachnick & Janussen, 2004. The second largest order, Lyssacinosida, includes four families: Rossellidae Schulze, 1885, Euplectellidae Gray, 1867, Aulocalycidae Ijima, 1927 and Leucopsacidae Ijima, 1903. The remaining orders have two families each: Aulocystidae Sollas, 1887 and Diapleuridae Ijima, 1927 (Lychniscosida), Hyalonematidae Gray, 1857 and Pheronematidae Gray, 1870 (Amphidiscosida). Finally, Hexasterophora *incertae sedis* includes the family Dactylocalycidae as well as the genera *Diaretula* Schmidt, 1879, *Lithospongia* Duchassaing & Michelotti, 1864, *Myliusia* Gray, 1859, *Reiswiginella* Van Soest, 2024 and *Rhabdostauridium* Schmidt, 1880, which cannot be assigned to any of the currently established families (Fig. 2). Nonetheless, species-level data reveal a slightly different pattern. The order Lyssacinosida contains the highest number of species, featuring 77 species—nearly twice as many as the second-largest order, Sceptrulophora, which comprises 40 species. The order Amphidiscosida includes 15 species, while Hexasterophora *incertae sedis* accounts for

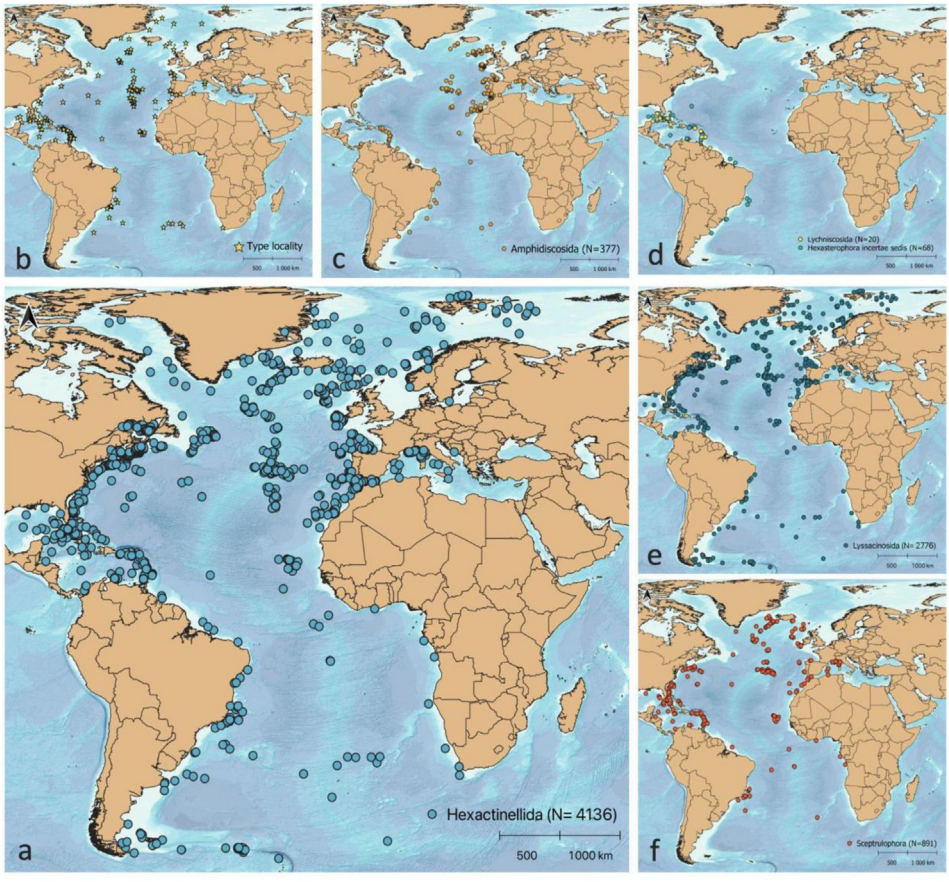

Fig. 1. Study Area. (a) Distribution of all glass sponges across the Atlantic and Arctic Oceans. (b) Recorded type specimens of glass sponges from these regions. (c) Recorded occurrences of sponges from the Order Amphidiscosida. (d) Recorded occurrences of sponges from the Order Lyssacinosida. (e) Recorded occurrences of sponges from the Order Sceptrulophora. (f) Recorded occurrences of sponges from the Order Lychniscosida and Hexasterophora *incertae sedis*.

13 species. The least diverse order is Lychniscosida, with only 4 known species, making it the smallest in terms of the number of species.

The distribution of glass sponge occurrences is highly concentrated in the North Atlantic. In particular, high densities of occurrences are reported along the Mid-Atlantic Ridge and the continental margins of Europe and North America and on the Caribbean Sea. In contrast, there is a significant data gap in the South Atlantic, with particularly limited records along the African margin (Fig. 1a). A striking characteristic of glass sponge records is the limited number of occurrences per species. A substantial 33.3% of all documented glass sponge species (51 spp.) in this region are known from a single recorded occurrence, usually resulting from their original description (Fig. 3). This trend becomes even more pronounced when considering species with 2 to 25 recorded occurrences, which together make up 52.9% of all species known for the area. Only a small fraction of species, 13.8%, corresponding to 21 species, have been documented more extensively, with more than 26 recorded occurrences.

Hexactinellid sponges exhibit a remarkably wide depth range, with recorded occurrences spanning from relatively shallow coastal waters to the deep abyssal plains. The shallowest known record belongs to *Oopsacas minuta* Topsent, 1927, found at 15 m at Živa Voda anchia-

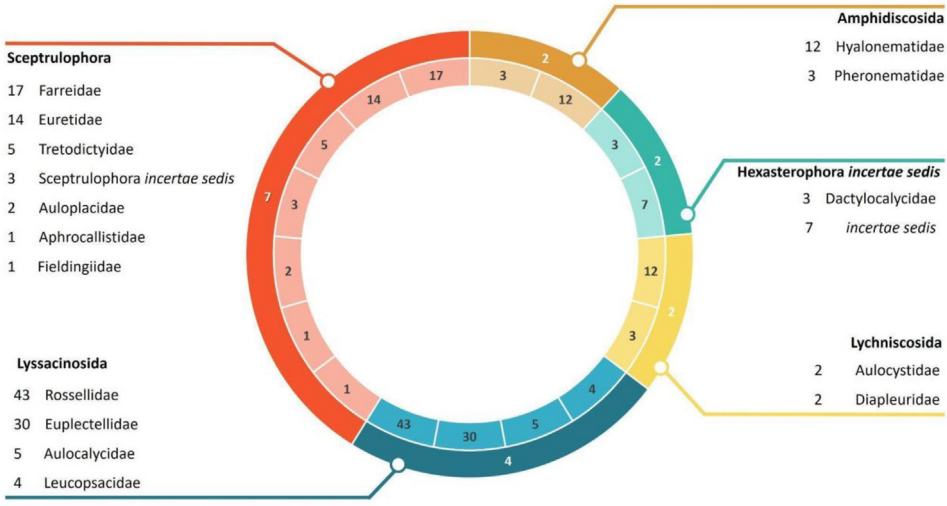

Fig. 2. Representation of the distribution of species among different taxonomic groups of the Class Hexactinellida. The chart illustrates the number of species within each family, categorized under the orders Sceptrulophora, Lychniscosida, Lyssacinosida, Amphidiscosida and Hexasterophora incertae sedis. The outer ring represents the orders, and the inner ring represents the families, with each segment indicating a specific family and numbers showing the species count per family.

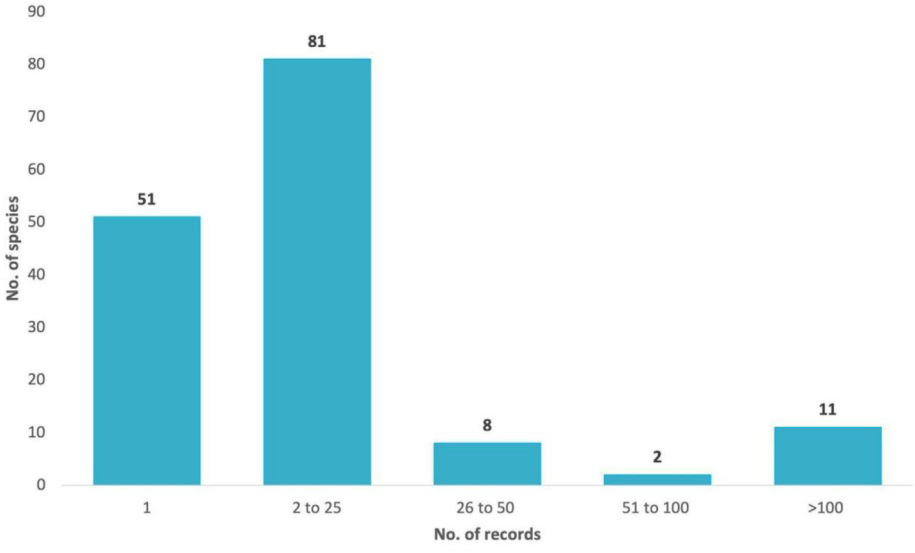

Fig. 3. Distribution of species by the number of records. The x-axis represents different ranges of records per species, while the y-axis shows the number of species in each category.

line cave on Hvar Island (Adriatic Sea, Croatia), while the deepest known occurrence is that of *Chonelasma ijimai* Topsent, 1901, recorded at depth of 5605 m in the tropical Areas Beyond National Jurisdiction [6]. Despite this broad range, glass sponge records are not evenly distributed across depth strata. The majority of occurrences (18,6%) are concentrated between 800 and 1200 m depth (Fig. 4a). When analyzing species diversity across depth gradients, a slightly

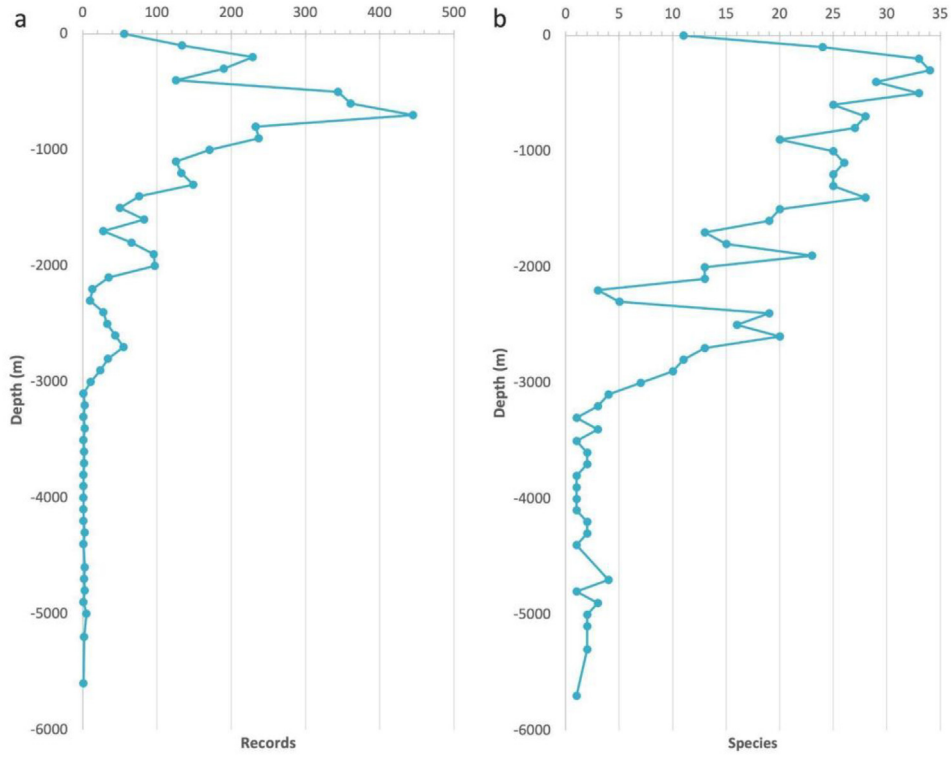

Fig. 4. Depth distribution of records and species. (a) Number of records across different depth ranges, showing the frequency of occurrences at various depths. (b) Number of species recorded at different depths, illustrating species richness along the depth gradient.

different pattern emerges. The highest number of species is observed at approximately 500 m, after which species diversity gradually declines as depth increases. However, there are also four distinct peaks in species diversity at depths of 500, 1500, 2000, and 2500 m (Fig. 4b).

4. Experimental Design, Materials and Methods

To establish a comprehensive dataset on the distribution of sponges belonging to the Class Hexactinellida across both the Atlantic and Arctic Oceans, we compiled data from multiple sources, including online databases (OBIS [7], GBIF [8], NOAA [9], and ICES VME database [10]) as well as published taxonomic literature, spanning over 152 years (1870 to 2022). The data collection process took place between October 2024 and January 2025 and focused on aggregating all available species-level records of Hexactinellida for the online databases and between January and March 2025 for the literature survey. Our methodology consisted of four main stages: (1) **Database Compilation**, which involved extracting occurrence data from established online repositories; (2) **Literature Survey**, where relevant scientific publications and reports were reviewed for distribution records; (3) **Data Harmonization**, during which we standardized taxonomic names, geospatial information, and metadata to ensure consistency.

Following standardization, the compiled data were organized and summarized in an Excel file [6]. The dataset includes three main categories of information: (1) **Taxonomic Information**, detailing the scientific classification of each species; (2) **Sampling Information**, which records

details such as date, location, and sampling method; and (3) **Reference Information**, citing the original data sources. A detailed explanation of these categories is provided in Table 1.

4.1. Literature survey

The initial step in our study involved compiling a comprehensive list of species that occur within the designated study area. To achieve this, we first examined all species classified under the class Hexactinellida that are recorded in the World Porifera Database (WPD) [3]. From this extensive dataset, we identified and extracted all species known to inhabit our specific study area.

Once we had this preliminary list, we proceeded to gather all available records of these species using three distinct approaches:

1. **Database Review** – We systematically collected all references listed in the WPD that contained relevant data on the identified species.
2. **Targeted Literature Search** – We conducted searches on Google Scholar, using species-specific keywords. These searches included the species' scientific names (e.g. "*Asconema setubalense*", "*Aphrocallistes beatrix*") combined with geographic terms (e.g., "Atlantic", "Arctic", "North Atlantic", "South Atlantic", "Arctic Ocean", "Mediterranean Sea") to locate additional studies documenting their occurrence.
3. **Broader Taxonomic Search** – To ensure we captured all relevant records, we performed additional searches on Google Scholar using broader taxonomic terms such as "Porifera", "Hexactinellida,", "glass sponges", "deep sea sponges" again paired with geographic terms to retrieve any overlooked references.

After identifying potential sources, we carefully assessed each article to confirm that it provided verifiable records — those supported by traceable evidence such as voucher-based taxonomic descriptions with reference to specimens deposited in museums or other recognised scientific institutions, or image-based identifications validated by taxonomists — of glass sponge occurrences within our study area. Records lacking sufficient taxonomic resolution, clear geographic information, or verifiable supporting evidence were excluded from the dataset (e.g., reports lacking precise locality information or records where species identification could not be confirmed). From these validated sources, we extracted all relevant data necessary to complete Table 1, ensuring that the dataset was as comprehensive and accurate as possible.

Following data extraction, we performed a taxonomic review to ensure the accuracy of species names and classification. Using the latest updates available in the WPD, we standardized species nomenclature and taxonomic groupings to reflect the most updated scientific consensus, including correcting misspellings and updating synonymized names. When species had known taxonomic synonyms, these were identified using the synonymy lists provided in the WPD and relevant taxonomic literature. Records reported under former or synonymized names were traced and incorporated into the dataset, and subsequently standardized under the currently accepted species name. This step was crucial for maintaining taxonomic consistency and ensuring that records published under historical nomenclature were not overlooked.

4.2. Database compilation

To gather data on the distribution of glass sponges from online databases, we selected the four most known publicly available repositories:

1. OBIS (Ocean Biodiversity Information System)
2. GBIF (Global Biodiversity Information Facility)
3. NOAA's National Database for Deep-sea Corals and Sponges

4. ICES (International Council for the Exploration of the Sea) Vulnerable Marine Ecosystems Database

These databases were selected for their extensive coverage, reliability, and accessibility of occurrence records for deep-sea organisms, including glass sponges. The data were manually extracted from each platform and subsequently processed in Microsoft Excel (Microsoft 365, v. 2602). After obtaining raw data from each source, the datasets were saved using the naming convention "[repository name] records_study area_Raw," as referenced in [6]. To maintain consistency in the analysis, relevant variables were pre-selected according to the criteria in Table 1, including species identification, geographic coordinates, depth. Any fields irrelevant to the study were systematically removed before further processing.

4.3. Data harmonization

To ensure consistency and reliability across the compiled dataset, we implemented a multi-step data harmonization process, which involved taxonomic validation, duplicate removal, and standardization of key metadata. The harmonization process consisted of the following steps:

4.4. Variable selection

1. A copy of the "Raw data" CSV was converted into an Excel file called "Data Treatment".
2. Relevant variables identified earlier (Table 1) were selected for further analysis, while all remaining fields were excluded.

4.5. De-replication process

4.5.1. First De-replication

1. **Taxonomic Cleaning**: Records at taxonomic rank above species level (e.g., genus or family) were removed and added to the file "[repository name]_data excluded" and flagged as "incomplete identification."
2. **Initial Duplicate Identification**: Duplicates were identified and removed based on the following conditions: records with the exact same coordinates (latitude and longitude), midpoint depth, and scientific name.
3. The eliminated records were moved to a separate file, labeled as "[repository name]_data excluded"

4.5.2. Second De-replication

1. First, subspecies were transformed into species-level records, and a note was added in the variable "notes" to indicate they were originally identified at the subspecies level.
2. A second de-replication was performed after the subspecies were standardized to species-level. This ensured that any previously missed duplicates, now consolidated under a common species name, were identified and removed.
3. These duplicates were extracted to the "[repository name]_data excluded" **Note**: The second de-replication was necessary due to the standardization of subspecies, which may have introduced additional duplicates that the first round missed due to extra taxonomic detail (subspecies).

4.5.3. Data curation

1. Taxonomic Verification: A list of all unique species was generated from the "scientific name" variable, and each species was verified using the WPD platform.

- Species names have been verified and corrected where necessary.
- If identified as **fossil species**, all records were eliminated, flagged as fossils, and moved to the "[repository name]_data excluded" file.
- The **Aphia ID** of each species was verified and corrected where necessary.
- A third de-replication was performed to remove any duplicate records that emerged after scientific name corrections.

4.5.4. Data corrections

1. **Date Verification**: Verified that dates followed the correct format (day/month/year).
2. **Depth Verification**: Checked minimum, maximum, and middle depth values to ensure consistency. In some cases (e.g., GBIF), only the middle depth was provided. This ensured consistent comparisons and helped identify previously undetected duplicates where different datasets had incomplete depth information (e.g., minimum only, maximum only, or a middle point).
3. Records found to be duplicated based on depth criteria were flagged and moved to the "[repository name]_data excluded" file and flagged as "duplicate"

4.6. Compilation of final dataset

4.6.1. Merging datasets

1. All individual datasets were merged, and a new variable, "source," was added to indicate the origin of each record (e.g., NOAA, OBIS, Literature).
2. **Final De-replication**: Duplicates were again identified and removed based on matching coordinates, middle depth, and scientific name.
3. The **number of duplicates eliminated** was recorded, and the removed records were moved to a new file "[repository name]_data excluded" file.
4. **Depth Verification**: Again, depth values were verified across species, and duplicates identified at this stage were flagged and extracted to the file "[repository name]_data excluded".
5. **Verification Against Literature**: Records from the literature were compared with other datasets. If records shared the same species, dates and depth but had slightly different coordinates (e.g., differing only in the last two digits), the literature record was retained and the other records extracted to the file "[repository name]_data excluded"*Data Publication*

The final dataset was compiled into three separate documents and made publicly available via Zenodo [6].

Limitations

One significant limitation of this study was the need to adjust taxa from the subspecies level to the species level. This was required because online databases occasionally contained duplicate records, with the same entry classified at the species level and others at the subspecies level. Additionally, subspecies recognition is a contentious topic within Hexactinellida [11], so their usage in both online databases and literature is inconsistent, and dependent on the author's recognition of the subspecies rank. To ensure data consistency, standardization to species level-classification was necessary. Additionally, the raw dataset also contained numerous duplicate records across databases and many entries lacked basic information, such as depth and coordinates. These issues can hinder accurate analysis to fully understand distribution patterns.

Ethics Statement

The authors confirm their adherence to the ethical requirements for publication in Data in Brief. This study does not involve human subjects, animal experiments, or data obtained from social media platforms.

Credit Author Statement

Celso Domingos: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Ana Sofia Soares**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Andreu Santín**: Conceptualization, Methodology, Writing - Review & Editing, Project administration. **Joana R. Xavier**: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Data Availability

An expert-curated dataset of glass sponges (Porifera, Hexactinellida) distribution across the Atlantic and Arctic Oceans. (Reference data) (Zenodo).

Acknowledgements

Celso Domingos would like to thank the Blue Young Talent program and the "Fundação para a Ciência e Tecnologia" (FCT) for their financial support through the doctoral grant reference UI/BD/150940/2021. **Ana Sofia Soares** would like to thank the "Fundação para a Ciência e Tecnologia" (FCT) for their financial support through the doctoral grant reference 2022.12800.BD. The authors are further supported by Portuguese national funds through FCT Foundation for Science and Technology within the scope of UID/04423/2025, (https://doi.org/10.54499/UID/04423/2025), UID/PRR/04423/2025, (https://doi.org/10.54499/UID/PRR/04423/2025), LA/P/0101/2020 (https://doi.org/10.54499/LA/P/0101/2020), as well as 2023.08132.CEECIND/CP2848/CT0008 (https://doi.org/10.54499/2023.08132.CEECIND/CP2848/CT0008) (to Joana Xavier) and 2023.11714.PEX and 2023.06804.CEECIND (https://doi.org/10.54499/2023.06804.CEECIND/CP2848/CT0009) (to Andreu Santín).

This research was performed in scope of the SponBIODIV project, which is funded by Biodiversa+, the European Biodiversity Partnership under the 2021–2022 BiodivProtect joint call for research proposals, co-funded by the European Commission (GA No.101052342) and with the funding organization FCT - Foundation for Science and Technology (DivProtect/0011/2021) and the Research Council of Norway (ref. 343194). It is also a contribution to the IUCN SSC Sponge Specialist Group.

Declaration of Competing Interest

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

References

[1] M.M Hogg, O.S. Tendal, K.W. Conway, S.A. Pomponi, R.W.M. van Soest, M. Krautter, J.M. Roberts, Deep-sea sponge grounds: Reservoirs of biodiversity, UNEP-WCMC Biodivers. (2010) Series No. 32.

[2] H. Meyer, E. Roberts, H. Rapp, A. Davies, Spatial patterns of arctic sponge ground fauna and demersal fish are detectable in autonomous underwater vehicle (AUV) imagery, Deep Sea Res. I 153 (2019) 103137, doi:10.1016/j.dsr.2019.103137.

[3] N.J. de Voogd, et al., World porifera database. https://www.marinespecies.org/porifera, 2025 (Accessed: Feb. 04, 2025).

[4] H.M. Reiswig, Class hexactinellida schmidt, 1870, in: J.N.A. Hooper, R.W.M. van Soest (Eds.), Systema Porifera. A Guide To The Classification Of Sponges, Systema Porifera. A Guide To The Classification Of Sponges, Vol. 2, Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow, 2002, pp. 1201–1202.

[5] J.J. Bell, E. Mcgrath, A. Biggerstaff, T. Bates, C.A. Cárdenas, H. Bennett, Global conservation status of sponges, Conserv. Biol. 29 (2014) 42–53, doi:10.1111/cobi.12447.
[6] C. Domingos, A.S. Soares, A. Santín, J.R. Xavier, An expert-curated dataset of glass sponges (porifera, hexactinellida) distribution across the Atlantic and Arctic Oceans, Zenodo Dataset (2025), doi:10.5281/zenodo.17575712.
[7] OBIS, Ocean biodiversity information system. https://obis.org (Accessed October 2024).
[8] GBIF, The Global Biodiversity Information FacilityWhat is GBIF?, 2024. Available from Accessed October 2024 https://www.gbif.org/what-is-gbif .
[9] National Oceanic and Atmospheric Administration (NOAA), Deep Sea Coral Research and Technology Program (DSCRTP) [2024].Observations of Deep-Sea Coral and Sponge Occurrences from NOAA's National Database for Deep-sea Corals and Sponges, 1842-Present, version [20250714-0] (NCEI Accession 0145037). [Hexactinellida]. NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/archive/accession/0145037. (Accessed January 2024).
[10] Vulnerable Marine Ecosystems Database, International council for the exploration of the sea (ICES). https://vme.ices.dk (Accessed October2024).
[11] D.A. Lopes, E. Hajdu, H.M. Reiswig, Taxonomy of farrea (Porifera, Hexactinellida, Hexactinosida) from the southwestern Atlantic, with description of a new species and a discussion on the recognition of subspecies in Porifera, Can. J. Zool. 89 (2011) 169–189, doi:10.1139/Tz10-105.