

# 1

## Organisms associated with anthropogenic marine debris reported in the scientific literature globally from 1990 to 2019

### 1.1 Abstract

- Humans have created a novel habitat in the ocean through the discharge of artificial materials like plastics, rubber, and manufactured wood products. The organisms that live and circulate through the marine environment in this new form of habitat have been identified, described, and otherwise published in the scientific record in growing numbers over the past 30 years. Comprehensively collating these records with their standardised taxonomy and corresponding spatial coordinates may aid in efforts to understand this ecosystem and its potential impacts on the rest of the marine environment.
- This database includes 3,294 of these recorded instances of organisms on anthropogenic marine debris (AMD) from 131 papers in every ocean basin, as well as the Mediterranean and several other seas published between 1992 and 2019.

- Records were pulled from peer-reviewed scientific publications found using tailored keyword searches in Scopus, Web of Science, and ProQuest.
- Papers returned in the initial queries were then manually assessed for their relevance according to their title and abstract. Those identified as potentially relevant were then read by the researchers to locate and incorporate relevant data included in the text, figures, tables, and appendices into the database. The results of this effort were then joined with validated taxonomic nomenclature from the World Register of Marine Species, Global Biodiversity Information Facility, and Integrated Taxonomic Information System. The broad marine region specified in each publication was manually noted along with the available spatial coordinates for the study as a whole from in-text description and georeferenced maps.
- The resulting database can be accessed from the stable link provided along with this publication.
- Combining the taxonomic information in this database with trait or behavioural databases, such as XXXX, could yield insights into how these ecosystems assemble and function. Further detail could be added and spatial analysis performed by merging this database with oceanographic data available from XXXX or XXXX.

### 1.1.1 Corresponding author

Correspondence and requests for materials should be directed to **C. M. (email: c.c.mcmains@gmail.com)**.

**Table 1.1:** Summary description of synthetic database

Design type(s)	Synthetic database - Marine ecology
Measurement type(s)	Taxonomic record
Technology type(s)	Specialist identification - Environmental DNA - DNA Sequencing
Factor type(s)	Scientific publication DOI
Sample characteristics	Spatial coordinates - Lowest taxonomic unit - Substrate material type - Publication year

## 1.2 Background & Summary

- Popularisation of plastic as a material for diverse uses in our daily lives has lead to a large and increasing amount of artificial materials in the world’s oceans. Collectively referred to as anthropogenic marine debris (AMD), these substrates comprised of plastics, rubber, and manufactured wood products, among other solid wastes from human activities in oceans and seas, also represent a novel type of habitat that is growing increasingly common. [cite]
- Scientific literature documented organisms associated with AMD as early as the 1970’s [cite], with more frequent and detailed records growing in number since the 1990’s [cite]. Researchers investigating the phenomenon of plastic pollution across its varied dimensions have often included secondary remarks of life attached to their study materials. A smaller subset of researchers has specifically studied the organisms themselves. Due to the diverse nature of the disciplines and modalities, records vary significantly across all meaningful dimensions, from collection methods through to how they are reported.
- Previous efforts towards compiling these records have resulted in several review papers [cites]. In one case, collaboration across disciplines resulted in a comprehensive database of all organisms attached to catalogued debris suspected to originate from the 2015 tsunami that occurred along the coast

of Japan's Tohoku prefecture [cite]. This database however, represents the first resource tailored for analysis of organisms associated with AMD globally from a comprehensive search of scientific publications and includes all of the records from these cited previous efforts.

- The database also not only contains records of identified organisms, but further includes information on methodology, substrate materials, and spatial co-ordinates. However, these variables were often only reported at the study level and not associated with individual organisms. Associations between them and the organisms are therefore not suitable for investigation using most types of statistical analysis.
- Taxonomic identifications of organisms in their source material have been joined to their currently accepted nomenclature as of March 2020 available from specialist-maintained open taxonomic databases and include stable identifiers. Substrate materials as reported in the literature have also been associated with standardised contemporary nomenclature from IUPAC [cite] and the DOI or ISBN is included for each publication within the database. Where relevant, each variable in the database has been named according to the Darwin Core Standard [cite]. This has all been done to facilitate the use of this database alongside others and as disciplines inevitably update their nomenclature.
- It may be useful to combine this database with others containing more detailed data on the organisms' evolutionary biology, physical traits, behaviour, or ecological function. Spatial analyses may georeference the coordinates of this database into marine regions of varying scope, such as through almanacs

available through XXXX [cite].

- Used alone, this database shows an increasing scientific interest in this novel ecosystem and changes in the way that it is studied. Molecular techniques have yielded a boom in the known diversity of organisms on AMD, particularly for bacteria and fungi. It also highlights the need for more research in the Southern Hemisphere and research in general. Almost all organisms were recorded only once or twice.
- While this database has the potential to yield important findings, its sparse nature relative to the known diversity of organisms and environments in the world's oceans is also likely to support the need for more research across disciplines.
- In its current state, it may be particularly useful for those researching the impacts of AMD on coastal environments. This database can be easily combined with new findings to provide background and context. It may also be useful for tracking the locations or spread of invasive species for biosecurity researchers and marine resource managers.
- Trends found in this database should be presented with the caveat that there is an immense number of potential external sources that could not be captured or reported. Findings related to materials should be carefully considered because reporting of which organisms was attached to which substrate was so rarely done that it could not be reliably recorded in this database. Spatial trends should also be considered judiciously due to the lack of tie between specific points and organisms. Studies with large spatial scopes should be

removed prior to any spatial analysis. It is also worth noting that data from the Japanese Tsunami Marine Debris (JTMD) database [cite], paper ID 999, has far more detailed and validated records than any other, which may skew results from the North Pacific unless accounted for.

- The database is separated into five primary tables linked by two primary identifiers (Table XX). Identifiers are unique keys at the highest record levels: study and accepted taxonomic nomenclature. Study level variables include data that was reported most often in the methods section of the paper, like methodology and sampling locations. While material types were reported in the results section, these too were often reported at the level of the study itself and are therefore linked to the unique study identifier. Both accepted and utilised taxonomic nomenclature for the most detailed identification available, or lowest taxonomic unit (LTU), of every identified organism is the second primary identifying key. This links each organism to its correct nomenclature and enables systematic validation and updates as necessary.
- All of the data originates in peer-reviewed scientific literature published and returned in searches for the specified queries in three major research repositories (Table XX) between 1992 and 2019. Organisms were recorded on debris items collected from, or deliberately immersed and naturally colonised in, the marine environment during the course of formal scientific research. The database contains records from all marine environments globally, from beaches to salt marshes and deep sea trenches. It only contains what the researchers reported in their publications, not necessarily all of the organisms

**Table 1.2:** Keywords used in the systematic database search that returned potentially relevant papers.

Primary	Secondary	Combined
marine OR ocean	debris OR litter	'marine OR ocean' AND 'debris OR litter'
marine OR ocean	wrack OR flotsam OR jetsam	'marine OR ocean' AND 'wrack OR flotsam OR jetsam'
marine OR ocean	macroplastic* OR microplastic* OR mesoplastic* OR 'plastic pollution'	'marine OR ocean' AND 'macroplastic* OR microplastic* OR mesoplastic* OR 'plastic pollution''

found or identified.

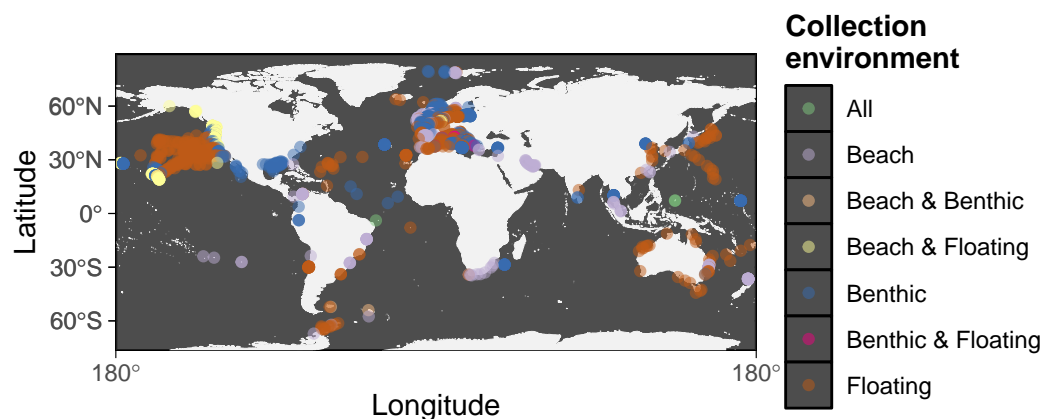
- It is recommended that users of this database update the taxonomic nomenclature using the noted databases and database IDs provided. Where material types are relevant, it is also recommended to reference current IUPC nomenclature and ensure that materials are described using contemporary naming conventions. Spatial coordinates are provided with descriptive quality ratings to assist in analyses where accuracy is important. However, no particular level of accuracy is guaranteed. For a guide on how to join the different tables, see the diagram included in Appendix XX.
- This database is provided in static format at the stable location associated with this publication. Any dynamic or updated versions of this database should reference this publication and be discoverable using that reference.
- At the time of this publication there is no update schedule and no plans exist to provide this resource in any other format.

```
df_keywords %>%
  kable(., "latex", booktabs = T,
        caption = "Keywords used in the systematic database search that returned po
  kable_styling(., latex_options = "scale_down")
```

## 1.3 Methods

- Overview of methods used to collect the field data.
- Explain where to find examples of each general method or a review paper describing various methods in detail.
- Overview of most common methods to process the field data.
- Brief comparison of these methods.
- Mention of other methods possible to use but not included in the dataset and description of reason why they were excluded.
- Describe how methods may be affected by different common scenarios that present in the field and lab.
- Delineate sources of data compiled into the database.
- Describe how data were accessed.
- Describe how data was filtered for relevancy.
- Describe how data was made to fit into a standardised data resource.
- Describe how ongoing maintenance or updating would be performed.





**Figure 1.1:** Global distribution of sampling sites within publications describing organisms associated with AMD.

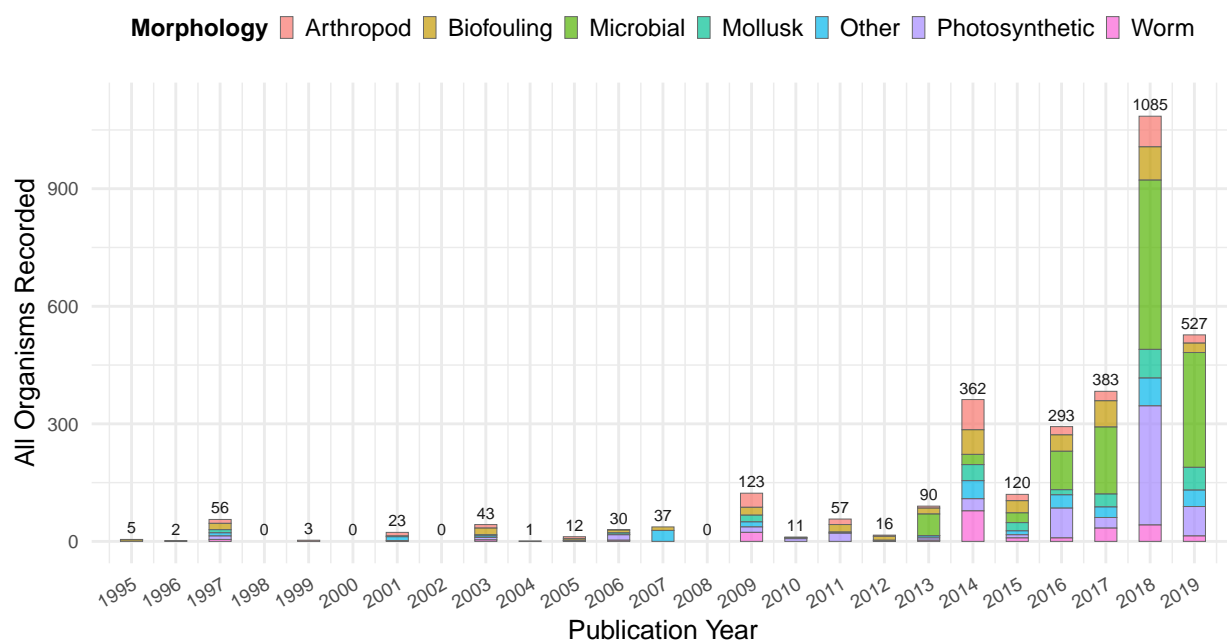
**Table 1.3:** Grouping of identified organisms into morphological categories by taxonomy.

Group	Level	Taxa
Arthropod	Phylum	Arthropoda
Biofouling	Class	Ascidiacea
Biofouling	Phylum	Bryozoa, Cnidaria, Echinodermata, Entoprocta, Porifera
Microbial	Kingdom	Bacteria, Archaea, Eukaryota, Protozoa, Metazoa, Fungi
Microbial	Phylum	Gastrotricha, Loricifera
Mollusk	Phylum	Mollusca
Other	Class	Actinopterygii, Elasmobranchii, Teleostei
Other	Phylum	Ctenophora, Nematoda, Priapulida
Photosynthetic	Kingdom	Plantae, Chromista
Worm	Phylum	Annelida, Nemertea, Platyhelminthes, Sipuncula, Chaetognatha

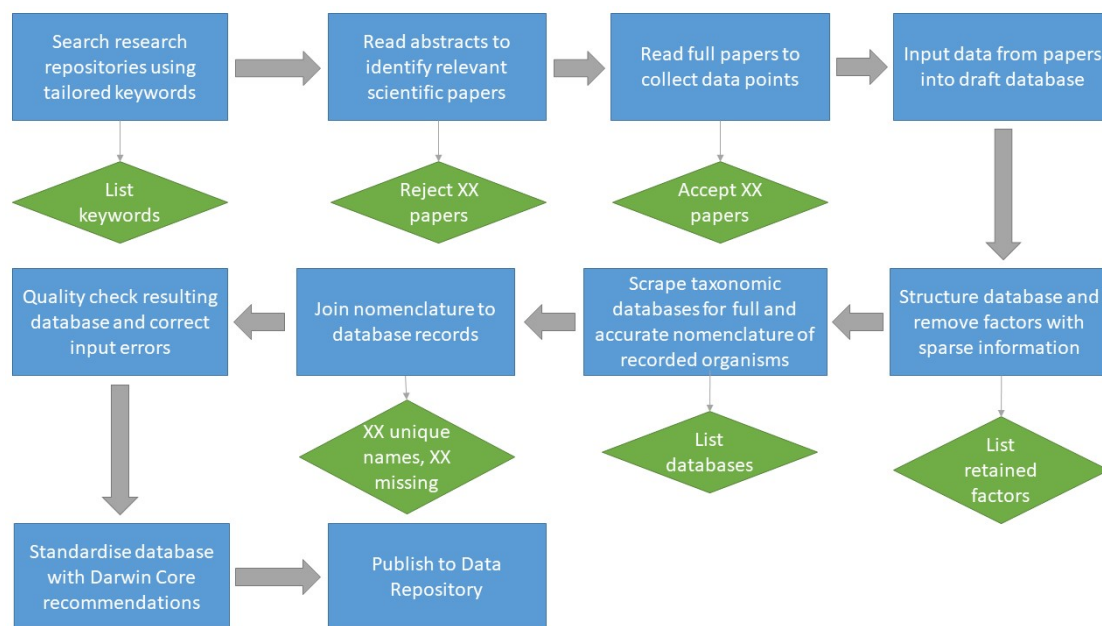
### 1.3.1 Figures

## 1.4 Data Records

- Define the lowest level of record and what it describes (include def of identifier)



**Figure 1.2:** Yearly counts of publications and records of identified organisms on AMD in the scientific literature.



**Figure 1.3:** Diagram of the process used to gather and integrate published records of organisms associated with AMD into the database.

- Describe how lowest records are linked through to other levels.
- Describe highest level of record and why it collects the lower levels.
- How metadata is linked to the dataset.
- Any caveats to the above, like why multiple names of a species might be attached to a record.
- Quality assurances for each record and how they were validated.

## 1.5 Technical Validation

- How data was cleaned and prepared in a way that needs to be acknowledged in any analysis.
- How data quality issues were addressed in the records that were lower quality but still met the minimum standards to be included.
- Any conversions of preparations that may be confusion to a user at first.
- Null or missing value signifiers.

## 1.6 Usage Notes

- Where the data has been made available for use.

- Suggest complementary datasets.
- Suggest how to link the data through to those complementary sets.
- Examples of how this type of data has been used in projects or publications in combination with other sources.
- Common reasons that the dataset admin would be contacted.
- How to contact the dataset admin.
- Where to go for more information about the dataset.

### **1.6.1 Data Citation**

*citation of dataset in stable repository*

### **1.6.2 Acknowledgements**

The records synthesised in this database were collected and described through the efforts of academic scientists and researchers, community scientists, and scientific professionals from around the world. We acknowledge their contributions to the body of knowledge compiled in this database and encourage further engagement with their published literature in our citation list. Where records from only one publication are used in continuing research, we recommend citing the publication directly or per the custodian's information. In the case where records from multiple sources were used from this database, this publication should be cited directly. This provenance can be determined using the publication identifier linked to every record in the

database. For data originating from the Japanese Tsunami Marine Debris database, the work of Carlton et. al should be cited directly or in addition to this database.

### **1.6.3 Author Contributions**

C.M. designed the methodology, performed all steps of the compilation and preparation of this database, and wrote the manuscript. C.L.H. and M.L.C. advised on the project concept and database synthesis. M.B. and C.L.H. provided significant assistance with and review of the manuscript and advised on data cleaning and validation. J.M. and M.L.C. also assisted with the manuscript review. Further details on author contributions can be found in Table 1 provided with the online version of this manuscript.

### **1.6.4 Additional Information**

The table of detailed authorship contributions is only available in the online version of this paper.

#### **Competing Interests**

The authors declare no competing financial interests.

#### **Appendices**

Create a table of detailed description of each authors contribution per the publisher's requirements.