Gap-analysis of DNA barcoding reference libraries for two taxon inventories

Guy Paz (✉ guy@ocean.org.il)  
Israel Oceanographic and Limnological Research Institute  
https://orcid.org/0000-0003-2263-9019

Baruch Rinkevich  
Israel Oceanographic and Limnological Research Institute

Research

Keywords: Biological monitoring, Ascidians, Cnidarians, BoLD, Quality control, ERMS

DOI: https://doi.org/10.21203/rs.3.rs-50653/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background

All-inclusive DNA barcoding libraries in the storage and analysis platform of the BoLD (Barcode of Life Data) system are essential for the study of the marine biodiversity and are pertinent for regulatory purposes, including ecosystem monitoring and assessment, such as in the context of the EU Water Framework Directive (WFD) and the Marine Strategy Framework Directive (MSFD). Here we investigate knowledge gaps in the lists of DNA barcoded organisms within the Cnidaria (Anthozoa and Hydrozoa) and Asciidiae reference libraries of the European Register of Marine Species (ERMS) dataset (402 ascidians and 1200 cnidarian species). ERMS records were checked species by species, against publicly available sequence information and the other data stored in BoLD pages.

Results

Results revealed that just 22.9% and 29.2% of the listed ascidians and cnidarians species, respectively, are BoLD’s barcoded species, of which, 58.4% and 52.3% of the seemingly barcoded species, respectively, were noted to have complete BoLD pages. Thus, only 11.44% of the tunicate and 17.07% of the cnidarian data in the ERMS lists are of high quality. Deep analyses revealed seven common types of gaps in the list of the barcoded species in addition to a wide range of discrepancies and misidentifications, discordances and errors primarily in the GenBank mined data as with the BINs assignments, and more.

Conclusions

Gap knowledge in barcoding of important taxonomic marine groups exist and in addition, quality management elements (quality assurance and quality control) were not employed when using the list for national monitoring projects, for regulatory compliance purposes and other purposes. Even though Bold is the most trustable DNA barcoding reference library, worldwide projects of DNA barcoding are needed in order to close these gaps of mistakes, verifications, missing data and unreliable sequences labs. Tight quality control and quality assurance is important to close the knowledge gaps of Barcoding of the European recommended ERMS reference library.

Background

The accurate evaluation of the biodiversity for any given ecosystem is a keystone element, even imperative, in numerous biological and applied disciplines, including ecology, conservation biology, food regulatory compliance, forensics, and ecosystem monitoring and assessment [1–2]. In response to the needs, DNA-based taxon identification constituted on the Cytochrome Oxidase I (COI) gene are habitually used to assess biodiversity, including species identification, species boundaries and species diversity analyses. The inventory of DNA barcodes is deposited in BoLD [3], a cloud-based data storage and analysis platform, that further employed as a curation tool. BoLD currently contains (updated to April, 2020) about eight million barcodes, encompassing > 310,000 animal, plant and fungi species. The rationale for using the COI gene in species barcoding lies on the understanding that the intraspecific diversity for the COI gene is lower than the interspecific diversity backed by the difficulties associated with the traditional taxonomy, the morphologically based species identification [4]. The major benefit of using BoLD is immediately emerged when an unknown sequence is compared against a database to determine its closest species match, an evaluation strictly depends on the correctness and reliability of the data stored in BoLD and on the quality of the barcode libraries [5]. Yet, it should be noted that while barcodes may be excellent tools to identify species that are already in BoLD, they may have poor predictive power in identification of unknown species.

As a curator tool, it is inferred that all barcode sequences stored in the BoLD database are backed by vouchered specimens and thoroughly identified by taxonomy experts. Yet, being a public database, it is inevitable that BoLD, as any other similar curation tool, will not accrue erroneous data, sometimes significantly [2, 6]. Taxonomic misidentifications and/or taxonomic conundrums, composites of cryptic species complexes, technical faults such as deficient DNA extraction, primary bias and/or PCR-based errors and contaminations by foreign DNA, including bacterial COI sequences, are just part of the causes that may unavoidably generate erroneous data and inaccurate sequences [2, 6–10]. The above difficulties may affect dramatically the accuracy of barcoding. For example, it was claimed that they lead to the lack of an unambiguous species level identification by the BIN (Barcode Index Number) tool in the BoLD system, and to taxonomic conflicts by the assignment of more than a single species name per bin [11].

The European Register of Marine Species (ERMS [12]) is an authoritative taxonomic check-list of species that are found in all European marine environments (the all-taxon marine species inventory from the Canaries and Azores to Greenland and north west Russia, towards the Mediterranean sea and the Baltic Sea), from the deep sea, all continental shelf areas and up to the splash zone above the high tide mark, and in estuaries, down to 0.5 psu salinity. During 1997–1999, ERMS was published on the internet and subsequently as a book, containing a list of about 30,000 marine species of the kingdoms Animalia, Plantae, Fungi and Prototista occurring in the European marine environment [13]. It is projected that this marine species inventory will be used as the standard reference and technological tool for marine research and for management of the marine environment in Europe.

Until recently, the standardized methodologies available for biological monitoring and management in the marine environments, primarily for practitioners, were restricted to traditional morphological taxonomy, tedious and time-consuming methodologies that require the involvement of expert taxonomists with skills that can only be attained via years of practice. This line of analyses is currently being replaced by molecular approaches such as DNA barcoding, environmental DNA (eDNA) and metabarcoding [5, 14–17]. For these approaches, it is of special interest to identify gaps in already existing or developing DNA barcode reference libraries, primarily those that are pertinent in the context of the EU Water Framework Directive (WFD) and the Marine Strategy Framework Directive (MSFD). A recent global study on this perspective [5] has revealed that the barcoding coverage varies strongly among taxonomic groups, and among geographic regions, pointing to many missing species and imperfect data (e.g., errors in species identification, discordance among taxonomists) that are relevant to monitoring and highlighted the needs for improving quality assurance of the barcode reference libraries.
Following Weigand et al. [5] global analysis, we aim here to investigate potential gaps in already DNA barcoded organisms (based on publicly available data in BoLD database) listed in two reference libraries of the ERMS inventory. We discuss the necessity of quality control (QC) when building and curating a barcode reference library, and provide recommendations for filling the gaps in the barcode library of European aquatic taxa.

Methods

Each BoLD page consists of six sections: 1. A short taxon description with a link to a species-specific website; 2. Statistics data, including: the number of records, specimens with sequences, specimens with barcodes, subspecies, subspecies with barcodes, public records, public available subspecies and public BIN clustering; 3. Worldwide specimen depositories. 5. Collection sites including countries; 4. Origin of sequences (GenBank ID numbers or sequencing laboratories); 6. Species images gallery.

Species checklist of two distant taxa (the Anthozoa and Hydrozoa of the phylum Cnidaria and the class Asciidiacea of the phylum Chordata) were downloaded from the European Register of Marine Species dataset (ERMS [12]). These taxa were not analysed in Weigand et al. [5] and are used here for deep analyses on quality control (QC) of the barcoded species from these two lists. The conformity of taxonomy and assurance of correct spelling were performed manually, species by species, against the World Register of Marine Species database (WoRMS [18]) and assessed following the recommendations by Costello et al. [19]. Finalized species-level checklists were re-ordered and compared with the BoLD list. For the analyses on the barcoded species (the COI marker) we used the checklist on BoLD created by Dirk Steinke, titled ‘Marine Animals Europe’ (BoLD checklist code: CL-MARAE; last updated on March 20th, 2017). The full list contains 27,634 records of marine species belonging to 10 phyla, Annelida, Arthropoda (class Decapoda and superorder Peracarida), Brachiopoda, Chordata (class Asciidiacea- subphylum Tunicata and class Pisces), Cnidaria (classes Anthozoa and Hydrozoa), Echinodermata, Mollusca (classes Bivalvia and Gastropoda), Nemertea, Priapulida, and Sipuncula. Datasets were generated on two checklists (the cnidarians and the ascidians; updated 18 July 2019) that were compared, species by species, to all publicly available sequence information in BoLD system and to the other data stored in BoLD pages. Working species by species allows the discrimination and the analyses of records in BoLD, the number of sequences/species, bin numbers/species, specimen in public depositories, the number of barcodes publicly stored in BoLD including those mined from the GenBank database at the National Library of Medicine (NCBI [20]) U.S. and the number of privately stored barcodes/species in BoLD. No geographical data were considered for the taxa analysed.

Results

The analysis was performed on the BoLD [3] database, for the 1603 species extracted from the ERMS taxonomic checklists, including 402 ascidians species and 1200 hydrozoans and anthozoans species (Supporting Information Table 3). Checking against the BoLD database (July 18th, 2019 inventory) we found only 88 (22.9%) of the ascidians species and just 351 (29.2%) cnidarians species in the list of BoLD’s barcoded species (Table 1). Then, analyses were performed on all the available BoLD records at three levels: the statistics level, the repository level and at the sequence level. At the statistic level, data for each species was individually inspected for the number of BoLD specimen (records), records that hold full sequences, records with just barcodes, and the numbers of public records and public BINs. At the repository level, we searched for number of records mined from GenBank and the number of repository facilities where voucher specimen were deposited. At the sequence level we recorded the open to the public sequences and scored (good, medium or low) the sequence quality control (QC) of the barcoded species from these two lists. The conformity of taxonomy and assurance of correct spelling were performed manually, species by species, against the World Register of Marine Species database (WoRMS [18]) and assessed following the recommendations by Costello et al. [19].

| Taxonomic group | ERMS species (n) | No. of species in BoLD (% are calculated to the total number of species in ERMS) | Sequences origin | No. of Public Bins* | BoLD species with more the 1 Bin |
|-----------------|-----------------|-------------------------------------------------|-----------------|-------------------|-------------------------------|
|                 | with ≥ 1 barcode | with ≥ 5 barcodes | with sequences | without sequences | with full description | Mined from Genbank | Sequencing labs |
| Ascidians       | 402             | 88 (21.9%)      | 50 (12.4%)    | 81 (20.1%)      | 88 (21.9%)      | 78 (19.4%)      | 74 (18.4%)      | 77 (19.2%)      | 68 (16.9%)      | 32 (8.0%)       |
| Cnidarians      | 1200            | 351 (29.2%)     | 153 (12.7%)   | 310 (25.8%)     | 351 (25.8%)     | 297 (24.7%)     | 205 (17.1%)     | 278 (23.1%)     | 231 (19.2%)     | 65 (5.4%)       |

* = Bins total number of all 88 species appears in BoLD.
** = Species total number BoLD with more the one Bin in BoLD.

Ascidians (Tables 1, Supporting Information Table 1)

For only 50 species (12.4% of the whole ERMS list of ascidians) of the 88 ERMS species referenced in the BoLD database there are more than five specimens barcoded. The COI gene sequence was assigned to 81 species of this list and just 78 of the species (19.4% of total species) have full descriptions in the BoLD pages, including number of specimen records, sequences, specimens with barcodes, species names, public records, public species and public BINs. As for the COI sequences, we assigned three types of records: records with no sequences, records containing sequences downloaded from GenBank (hence with no trace files, without trusted curation), and records containing BoLD related new sequences (with trace files). Many species contained all three record types. Thus, in
77 species, the COI gene was sequenced and contains trace files, while for 74 species, the COI sequences were mined from the GenBank. A total of 68 public bins were assembled and 32 species contained more than a single Bin (Table 1).

Cnidarians (Tables 1, Supporting Information Table 2)

For only 153 species (12.7% of the whole 1200 ERMS list) of the 351 species found in the BoLD database contained more than five specimens barcoded. The COI gene sequence was assigned to 310 species of this list and just 297 species (6.5% of total species) had full descriptions in the BoLD pages, including number of specimen records, sequences, specimens with barcodes, species names, public records, public species and public BINS. As for the COI sequences, we assigned three types of records- records with no sequences, records containing sequences downloaded form GenBank (hence with no trace files, without trusted curation), and records containing BoLD related new sequences (with trace files). In 278 species COI was sequenced and contains trace files, however in 205 species, sequences were mined form the GenBank. A total of 68 public bins were assembled and 65 species contained more than a single Bin (Table 1).

Then, we assigned seven common types of gaps in the list of barcoded species, as: (a) records with no data available (empty pages on BoLD website)- 10 tunicate species and 41 cnidarians, (b) records with partial public data and no COI sequences- 10 tunicate species and 41 cnidarians, (c) no public available records- 17 tunicate species and 73 cnidarians, (d) records with sequences mined from the GenBank, many with gaps in sequences and all without trace files- 74 tunicate species and 205 cnidarians, (e) records dispensed between more than a single bin- 29 tunicate species and 65 cnidarians, (f) records with no bins- 25 tunicates and 120 cnidarians, (g) records with sequences dispersed within several bins- 29 tunicate species and 65 cnidarians (examples for the seven knowledge gap types are detailed in Table 2). In summary, only 52.3% and 58.4% for cnidarians and tunicates, respectively, of the seemingly ‘barcoded species’ on the BoLD website have complete BoLD pages, just 11.44% tunicate species and 17.07% cnidarian species appearing in the ERMS lists.

| Taxon          | Ascidians | Cnidarians | Ascidians | Cnidarians | Ascidians | Cnidarians | Ascidians | Cnidarians | Ascidians | Cnidarians |
|---------------|-----------|------------|-----------|------------|-----------|------------|-----------|------------|-----------|------------|
| Species       | Molguloides vitrea | Edwardsia claparell | Molgula oculata | Epizoanthus eardmanni | Molgula occidentalis | Epizoanthus incrustatus | Dendrodoa grossularia | Eunicella singularis | Botryllus schlosseri | Porpita porpita |
| Records in BoLD | 4         | 1          | 5         | 7          | 7         | 9          | 4         | 7          | 577       | 14         |
| No. of Sequences in BoLD | 0         | 0          | 1         | 0          | 2         | 5          | 3         | 7          | 540       | 12         |
| No. of Barcodes | 0         | 0          | 1         | 0          | 0         | 5          | 2         | 7          | 531       | 12         |
| No. of Public Barcodes | 0         | 0          | 0         | 0          | 0         | 0          | 3         | 7          | 531       | 8          |
| Mined from GenBank | 0         | 0          | 1         | 0          | 0         | 0          | 3         | 7          | 528       | 2          |
| No. of Specimen Depositories | 0         | 0          | 0         | 0          | 0         | 2          | 1         | 0          | 8         | 2          |
| Public Bins | 0         | 0          | 0         | 0          | 0         | 0          | 1         | 1          | 18        | 2          |

Discussion

The transformation of marine ecosystems and global biodiversity loss [21–22] pose challenges to the developing marine and water strategy directives (such as WFD and MSFD), as to the development of marine policy and management approaches [23–25]. Clearly, the current know-how of marine biodiversity is conclusive to the rigour of the science that underpins policy and management assessments and for the future of all marine ecosystems.

Reliable barcode reference libraries are of particular importance [5]. It is evident that the level of taxonomic detection and degree of accuracy is directly contingent on the newly developed molecular-biology depositories, the reference libraries completion and reliability of the DNA records [26–29]. Given our increased reliance on molecular taxonomy as a robust tool [4], strengthening the existing reference libraries is a necessitate for a wide range of scientific and applied purposes, such as monitoring, eDNA and metabarcoding approaches, all targeting matched identifications and the assessments of biodiversity and abundance [5, 14–17, 29].

To test if the available barcoding data is applicable, we analysed two major reference libraries, the ascidians and the cnidarians form the ERMS recommended list. Working on members of two taxa, the Anthozoa and Hydrozoa of the phylum Cnidaria and the class Asciidiacea of the phylum Chordata, our analyses showed that important reference libraries lack reliable barcodes for these dominant marine macroinvertebrate species. Results further revealed a wide range of difficulties and inconsistencies, including taxonomic congruency of the COI barcode records on the one hand and possible cryptic diversity (sensu Leite et al. [28]), that should be further studied, on the other hand. The above clearly affect the wholeness of the ERMS list, as only 52.3% and 58.4% of the cnidarians.
and tunicates, respectively, of the short list of seemingly well 'barcoded species' were noted to have confirmed complete BoLD pages, further lighting that only 11.44% of the tunicate data and 17.07% of the cnidarian data in the ERMS lists are reliable and fully supported (July 2019 state).

Further, two major results have been emerged. The first indicated that the BINs assignments revealed a sizeable amount of discordances, many are probably related to species misidentifications or synonyms [28] or the deficiency of the BIN clustering algorithm to correctly discriminate species [30]. The second outcome pointed towards the low power of GenBank results as compared to the BoLD, discrepancies that are already noted in the literature [2, 31], characterized by contaminations of the query sequences discordances and misidentifications. The BoLD and the GenBank data storage systems are highly intermingled. About 11% of COI barcode records on BoLD are mined from the GenBank, while 75% of the COI barcodes on the GenBank system originate from the BoLD system [31], yet our results point to the many weakness features associated with the GenBank data that are less informative and do not present extended data elements such as trace electropherograms, specimen images, voucher numbers or BIN assignments, and are usually poorly curated (see also Bridge et al. [26]) as compared to the BoLD.

Human-made artefacts during the barcoding processes, primarily for the GenBank storage data, affect the reliability of DNA barcoding to correctly assign a given specimen to species. The overall recorded knowledge gaps in the DNA barcodes found in the present study are considerably high. Clearly, the two types of gaps, the first of records not found in BoLD, and the second- errors and missing data in already DNA barcoded organisms that are found in the BoLD database, may impact dramatically the accuracy of any DNA-based assessment or biomonitoring approach that counts on BoLD's data. Many additional commonly accumulated errors target inaccurate taxonomic identifications of specimens by nonexpert taxonomists (in addition to the lack of voucher specimen), sequence contamination, incomplete reference data and insufficient quality of the uploaded molecular data [5, 30]. Working on reference libraries of DNA barcodes of marine organisms (invertebrates and fish taxa), Weigand et al. [5] recorded numerous identification errors, sequence contamination, incomplete reference (missing trace files or primer information) as inadequate data management. The results of the present study, as supported by other recent studies [2, 5, 17, 28, 30, 32], reveal that we are still away from possessing decent representative reference libraries for important taxonomic groups. In addition, new DNA barcode data are continuously made available for the already barcoded and additional species from the reference libraries, including additional auditing and annotation processes, altogether helping in closing gap knowledge and purging accumulated erroneous data.

Our global assessment on the completeness of the ERMS library elucidates that the available COI barcode data adequately covers just a small fraction of this reference library, raising an alarm for similar statuses in other reference libraries, such as the AZTI's Marine Biotic Index (AMBI, [33]).

**Conclusions**

DNA barcoding-based approaches are superior in issues like diminished ambiguity and improved accuracy of species identification with ultimate verification of results against repository documentations [14, 32]. Quality management elements (such as quality assurance and quality control) should be employed when using the list for monitoring and other purposes and for closing the knowledge gaps. Purging of errors from BoLD, the most trustable DNA barcoding reference library, will significantly contribute to future attempts in biodiversity monitoring efforts, in eDNA and metabarcoding approaches and their assessments, for various regulatory compliance purposes, forensics, and more [1–2, 27]. Given the increasing use of high throughput sequencing approaches and of automated pipelines, data quality aspects of DNA barcodes should be cogitated with higher priority.

**Abbreviations**

BoLD: Barcode of Life Data, COI: Cytochrome Oxidase I, CQ/QA: Quality control/Quality assurance, eDNA: environmental DNA, ERMS: European Register of Marine Species, MSFD: Marine Strategy Framework Directive, NCBI: National Library of Medicine, WFD: Water Framework Directive, WoRMS: World Register of Marine Species database.

**Declarations**

**Availability of data and materials:**

All data generated or analyzed during this study are included in this published article.

**Ethics approval and consent to participate**

The authors declare they have no competing financial interests.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare they have no competing interests.

**Funding**

The authors did not receive any funding for this research.
Authors contributions:
The authors conceived, wrote and approved the manuscript.

Acknowledgements
This study associates with the National Israeli Marine Barcoding project (BIM) at IOLR and is the outcome of activities and discussion in the COST Action CA15219 “Developing new genetic tools for bioassessment of aquatic ecosystems in Europe” (DNAqua-Net).

References
1. Baird DJ, Sweeney BW (2011) Applying DNA barcoding in benthology: the state of the science. J N Am Benthol Soc 30:122–124. 10.1899/10-154.1.
2. Mekie john KA, Damaso N, Robertson JM (2019) Assessment of BOLD and GenBank- Their accuracy and reliability for the identification of biological materials. PLoS ONE 14(6): e0217084. doi.org/10.1371/journal.pone.0217084.
3. BoLD- The Barcode of Life Data System http://v4.boldsystems.org. Accessed 19 July 2020
4. Blaxter ML (2004) The promise of a DNA taxonomy. Phil Trans R Soc Lond B Biological Sciences 359:669–679. 10.1098/rstb.2003.1447
5. Weigand H et al. (2019) DNA barcode reference libraries for marine biota in Europe: Gap-analysis and recommendations for future work. Sci Total Environ 678: 499–524 doi.org/10.1016/j.scitotenv.2019.04.247.
6. Shen YY, Chen X, Murphy RW (2013) Assessing DNA barcoding as a tool for species identification and data quality control. PLoS ONE 8(2):e57125. 10.1371/journal.pone.0057125
7. Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. App Environ Microbiol 71: 7724–7736. 10.1128/AEM.71.12.7724-7736.2005
8. Wells JD, Stevens JR (2008) Application of DNA-based methods in forensic entomology. Annu Rev Entomol 53:103–120. doi.org/10.1146/annurev.ento.52.110405.091423.
9. Sonet G, Jordens K, Braet Y, Bourguignon L, Dupont E, Backeljau T et al. (2013) Utility of GenBank and the Barcode of Life Data Systems (BOLD) for the identification of forensically important Diptera from Belgium and France. ZooKeys 365: 307–328 10.3897/zookeys.365.6027.
10. Mioduchowska M, Czyż MJ, Goldyn B, Kur J, Sell J (2018) Instances of erroneous DNA barcoding of metazoan invertebrates: Are universal cox1 gene primers too “universal”? PLoS ONE 13(6): e0199609 https://doi.org/10.1371/journal.pone.0199609.
11. Collins RA and Cruickshank RH (2013) The seven deadly sins of DNA barcoding. Mol Ecol Resour 13: 969–975 10.1111/1755-0998.12046.
12. ERMS- European Register of Marine Species. http://www.marbef.org/data/erms.php. Accessed 19 July 2020
13. Costello MJ, Emblow C, White RJ (Ed) (2001) European register of marine species: a check-list of the marine species in Europe and a bibliography of guides to their identification. Collection Patrimoines Naturels, 50. Museum national d’Histoire naturelle: Paris ISBN 2-85653-538-0, 463 pp.
14. Leese F, Altermann F, Bouchet A et al. (2016) DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. Res Ideas Outcomes 2: e11321 https://doi.org/10.1371/journal.pone.011321.
15. Karahan A, Douek J, Paz G, Stern N, Kideys AK, Shaish L, Goren M, Rinkevich B (2017) Employing DNA barcoding as taxonomy and conservation tools for fish species censuses at the southeastern Mediterranean, a hot-spot area for biological invasion. J Nat Conserv 36: 1–9 doi.org/10.1016/j.jnc.2017.01.004.
16. Pawlowski, J Kahlert M, Kelly-Quinn M et al. The future of biotic indices in the ecogenomic era: integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. (2018) Sci Total Environ 637–638: 1295–1310 doi.org/10.1016/j.scitotenv.2018.05.002.
17. Paz G, Yudkovsky Y, Shaish L, Stern N, Lubinevski H, Mienis HK, Douek J, Galil B, Goren M, Rinkevich B. Initiating DNA barcoding of Eastern Mediterranean deep-sea biota. (2018) Mediterr Mar Sci 19/3: 416–429 https://doi.org/10.1080/1229134X.2018.1489194.
18. WoRMS- World Register of Marine Species. www.marinespecies.org. Accessed 19 July 2020
19. Costello MJ, Bouchet P, Boxshall G, Fauchald K, Gordon D, Hoeksema BW, Poore GCB, Fauchald K, Gordon D, Hoeksema BW, Poore GCB, Van Soest RWM, Stöhr S, Walter TC, Vanhoomae B, Decock W, Appelans W (2013) Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. PLoS ONE 8(1):e51629. 10.1371/journal.pone.0051629. 
20. NCBI- National Center for Biotechnology Information U.S. National Library of Medicine. . Last accessed 19 July 2020
21. Johnson CN, Balmford A, Brook BW, Buttell JC, Galetti M, Guangchun L, Wilmshurst JM (2017) Biodiversity losses and conservation responses in the Anthropocene Science. 356(6335): 270–275 10.1126/science.aam9317
22. Rinkevich B (2019) Coral chimerism as an evolutionary rescue mechanism to mitigate global climate change impacts. Glob Chang Biol 25:1198–1206. https://doi.org/10.1111/gcb.14576
23. Creighton C, Hobday AJ, Lockwood M, Pecl GT (2016) Adapting management of marine environments to a changing climate: a checklist to guide reform and assess progress. Ecosystems 19:187–219. 10.1007/s10021-015-9925-2
24. Cormier R, Elliott M, Rice J (2019) Putting on a bow-tie to sort out who does what and why in the complex arena of marine policy and management. Sci Total Environ 648:293–305. doi.org/10.1016/j.scitotenv.2018.08.168
25. Fine M, Cinar M, Voolstra CR, Safa A, Rinkevich B, Laffoley D, Hilmi N, Allemand D (2019) Coral reefs of the Red Sea: Challenges and potential solutions. Reg Stud Mar Sci 25:100498. doi.org/10.1016/j.rsma.2018.100498

26. Bridge PD, Roberts PJ, Spooner BM, Panchal G (2003) On the unreliability of published DNA sequences. New Phytol 160:43–48. doi.org/10.1046/j.1469-8137.2003.00861.x

27. Trebitz A, Sykes M, Barge J (2019) A reference inventory for aquatic fauna of the Laurentian Great Lakes. J Great Lakes Res 45:1036–1046. doi.org/10.1016/j.jglr.2019.10.004

28. Leite BR, Vieira PE, Teixeira MAL, Lobo-Arteaga J, Hollatz C, Borges LMS, Duarte S, Troncoso JS, Costa FO (2020) Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Reg Stud Mar Sci 36:101307. doi.org/10.1016/j.rsma.2020.101307

29. Stoeckle MY, Das Mishu M, Charlop-Powers Z (2020) Improved Environmental DNA Reference Library Detects Overlooked Marine Fishes in New Jersey United States. Front Mar Sci 7:226. 10.3389/fmars.2020.00226

30. Oliveira LM, Knebelsberger T, Landi M, Soares P, Raupach MJ, Costa FO (2016) Assembling and auditing a comprehensive DNA barcode reference library for European marine fishes. J Fish Biol 89:2741–2754. https://doi.org/10.1111/jfb.13169

31. Pentinsaari M, Ratnasingham S, Miller SE, Hebert PDN (2020) BOLD and GenBank revisited – Do identification errors arise in the lab or in the sequence libraries? PLoS ONE 15(4):e0231814. doi.org/10.1371/journal.pone.0231814

32. Kelly RP, Closek CJ, O’Donnell JL, Kralj JE, Shelton AO, Samhouri JF (2017) Genetic and manual survey methods yield different and complementary views of an ecosystem. Front Mar Sci 3:283. doi.org/10.3389/fmars.2016.00283

33. Borja A, Franco J Pérez V (2000) A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Mar Pollut Bull 40:1100–1114. http://doi.org/10.1016/S0025-326X(00)00618

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS3.xlsx
- TableS2.docx
- TableS1.docx