

Open camera or QR reader and  
scan code to access this article  
and other resources online.



## The Ocean Genome Legacy: A Genomic Resource Repository for Marine Life

Rosalia Falco,\* Hannah J. Appiah-Madson,\* and Daniel L. Distel

**A**DVANCES IN GENOMICS HAVE revolutionized many areas of research, providing new insights into the extraordinary biochemical diversity of life. The rapidly decreasing cost of DNA sequencing now enables unprecedented insights into biological diversity, species and ecosystem conservation, human and animal health, agriculture, food production, and biotechnology. However, environmental degradation and anthropogenic decimation of natural populations increasingly make biological collection more costly and less environmentally acceptable. Changing environmental regulations and complex permitting processes also complicate collecting practices and add to their cost. As a result, the rising cost and difficulty of sampling now frequently exceed sequencing costs as obstacles to biological research.

These considerations are particularly important for marine environments. The World Register of Marine Species<sup>1</sup> names >239,000 accepted marine species, however, estimates suggest that >90% of the ocean's species remain unnamed, undiscovered, undescribed, and poorly understood.<sup>2</sup> Although the ocean contains fewer named species than terrestrial environments, its deep phylogenetic biodiversity is evident at higher taxonomic levels where much of the fundamental biochemical diversity of life is found.

For example, two-thirds of all animal phyla are mainly or exclusively marine, whereas only about a quarter are commonly represented on land. Only one animal phylum has no known marine representatives. Thus, to fully understand the diversity of life, the vast biodiversity of marine life must be addressed. The marine realm is also substantially larger in area and volume than the terrestrial world and is often more difficult and dangerous to access. Marine sampling requires specialized knowledge, skills, and equipment, and often poses significant health risks to collectors.

The Ocean Genome Legacy Center (OGL) of Northeastern University was created to address these issues. OGL is a nonprofit open-access marine biorepository and research laboratory dedicated to collecting, preserving, and sharing genomic resources—primarily preserved DNA and tissue samples—from diverse marine taxa. OGL's mission is to acquire, authenticate, study, preserve, develop, and distribute marine genomic materials, biological specimens, information, technology, and standards needed to advance basic and applied research.

OGL operates under policies that ensure legal, equitable, and ethical use of genomic resources in compliance with existing laws and the principles established by the Nagoya Protocol (<https://www.cbd.int/abs/>). OGL operates as a collaborative research collection and welcomes sample contributions and requests from individuals and institutions. Its primary goal is to accelerate research that can protect marine ecosystems and improve the human condition by (a) facilitating open sample sharing, (b) enabling easy access to rare and hard-to-find materials, (c) allowing multiple uses of collected materials, (d) mitigating the costs and hazards of marine sampling, (e) reducing the need for destructive sampling in sensitive environments, and (f) facilitating dissemination of contextual, taxonomic, and biogeographic data.

OGL's collection (Fig. 1) is located at Northeastern University's Marine Science Center in Nahant, MA, in a secure cryostorage facility that includes mechanical (−80°C) and liquid nitrogen (−186°C) freezer units with backup power, controlled entry, and temperature surveillance provided around-the-clock, monitored both by security personnel and electronic security systems.

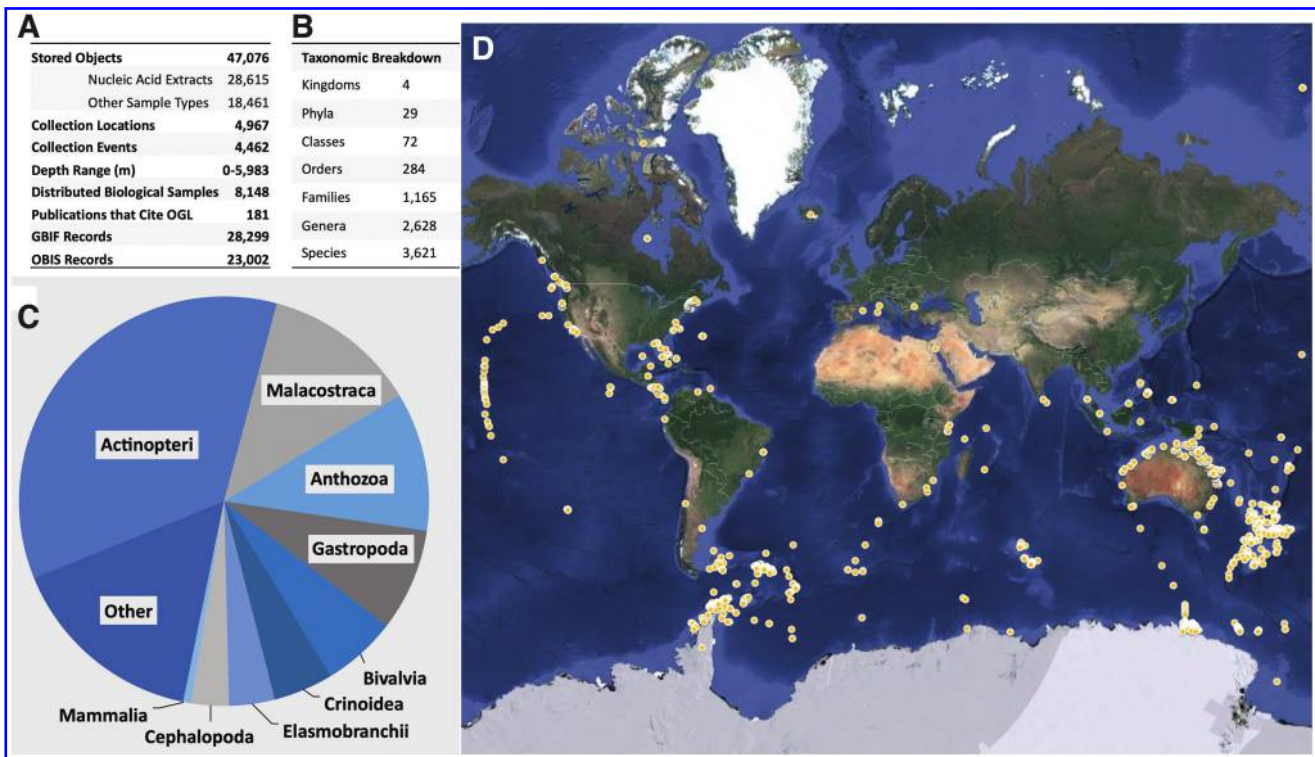
OGL's collection policy may be viewed at <https://bit.ly/OGLCollectionPolicy>. In brief, OGL accepts legally

Ocean Genome Legacy Center, Department of Marine and Environmental Science, Northeastern University, Nahant, Massachusetts, USA.

\*These authors contributed equally to this study.

© Rosalia Falco et al. 2022; Published by Mary Ann Liebert, Inc. This Open Access article is distributed under the terms of the Creative Commons License [CC-BY] (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Correction added** on February 11, 2022 after first online publication of February 2, 2022: The article reflects Open Access, with copyright transferring to the author(s), and a Creative Commons License (CC-BY) added (<http://creativecommons.org/licenses/by/4.0>).



**FIG. 1.** The OGL Genomic Resource Collection. **(A)** Description of OGL holdings and products, including distributed biological materials and data. **(B)** Taxonomic ranks represented by stored objects held in the OGL collection. **(C)** Proportion of stored objects in the OGL collection by class. **(D)** Map displaying collection locations associated with stored objects held in the OGL collection. All data reflect the collection as of October 22, 2021. GBIF, Global Biodiversity Information Facility; OBIS, Ocean Biodiversity Information System; OGL, Ocean Genome Legacy. Color images are available online.

collected biological materials from all marine taxa. Samples are vetted for provenance and permits, and taxonomic names are verified against the World Register of Marine Species ([www.marinespecies.org](http://www.marinespecies.org)).<sup>3</sup> Physical voucher materials are held at OGL or other public collections, with appropriate accession numbers provided upon submission. Digital vouchers may include photo, video, audio, and/or diagnostic molecular sequence files. OGL operates in compliance with best practices established by the International Society for Biological and Environmental Repositories ([www.isber.org](http://www.isber.org)).<sup>4</sup>

Taxonomic, contextual, collection, storage, and biogeographic data, as well as associated documents and permits are maintained in a database using Darwin Core controlled vocabulary.<sup>5</sup> OGL is a core member of the Global Genome Biodiversity Network ([www.ggbn.org](http://www.ggbn.org))<sup>6</sup> and shares biogeographic occurrence data through the Global Biodiversity Information Facility ([www.gbif.org](http://www.gbif.org)) and the Ocean Biodiversity Information System ([www.obis.org](http://www.obis.org)). There is no charge for samples deposited to or received from the OGL collection, however, modest cost recovery fees are typically charged for sample requests to cover sample processing, shipping, handling, and any additional services requested.

Materials are distributed under Material Transfer Agreements (<https://bit.ly/OGLDistributionMTA>), which govern their noncommercial use. Commercial use requires appropriate licensing agreements formed directly with the nations, depositors, and other stakeholders owning rights to those materials. OGL is not typically a party to such licensing agreements and claims no commercial rights to materials deposited to its collections.

Instructions for sample deposits and requests may be found at <https://bit.ly/OGLDepositSamples> and <https://bit.ly/OGLRequestSamples>. OGL facilitates sample deposit by providing depositor kits that contain labeled tubes, appropriate preservatives, and an electronic spreadsheet for data deposition. Upon receipt of biological materials, DNA is extracted and purified according to taxon- and tissue-specific protocols. Extracted DNA and residual tissues are preserved at ultralow temperatures ( $-80^{\circ}\text{C}$  to  $-186^{\circ}\text{C}$ ) to ensure their long-term viability. Quality control checks are performed as needed by gel electrophoresis, spectrophotometry, and other appropriate methods.

In addition to preservation and distribution of biological materials, OGL offers consultation, assistance in locating and procuring materials not held in the OGL collection, and DNA barcoding for species identification. Examples of OGL use cases include the following:

1. Storage, documentation, and sharing of biological materials for researchers and institutions.
2. Collection planning for retirement: OGL works with researchers to ensure proper documentation, preservation, and secure maintenance of their collections beyond retirement. Past retirement rescues have included valuable collections of Antarctic fishes, deep-sea hydrothermal vent and cold seep invertebrates, and polyclad flatworms.
3. Specimen and permit management for common, rare, and endangered species: OGL facilitates sample storage, import, export, and compliance with CITES, Marine Mammal Protection Act, and Endangered Species Act.

4. Specimens for medical research: OGL samples have been used for discovery of antibiotic<sup>7,8</sup> and antiparasitic<sup>9</sup> drug lead candidates.
5. Research consultation: OGL provided taxonomic guidance and >500 marine fish samples for a survey of DNA methylation across the animal kingdom.
6. Species identification: OGL performed DNA barcode analysis on 286 supermarket fish samples for the New York Attorney General's office<sup>10</sup> and routinely identifies marine samples for government and commercial entities, researchers, and private citizens.

For more information, please see <https://ogl.northeastern.edu/> or contact OGL at [oglinfo@northeastern.edu](mailto:oglinfo@northeastern.edu).

### Authors' Contributions

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the study to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the article.

### Acknowledgments

We dedicate this article to the memory of OGL founder Dr. Donald G. Comb. We also express our deep appreciation to the individual and institutional collaborators, depositors, donors, staff, students, and users of the OGL collection and to New England Biolabs Inc. and Northeastern University for their support in realizing the mission of OGL.

### Data Sharing Statement

Data describing the OGL collection can be accessed on GBIF at (<https://www.gbif.org/dataset/968dfa3e-3614-4bab-a3b5-e9209bdd6078>), OBIS at (<https://obis.org/dataset/9d629feb-ba03-4a83-908c-794d94d30531>), GGBN at [https://www.ggbn.org/ggbn\\_portal/stats/details?registry=OGL%2C+Nahant](https://www.ggbn.org/ggbn_portal/stats/details?registry=OGL%2C+Nahant), and at (<https://ogl.northeastern.edu/catalog/>).

### Author Disclosure Statement

The authors declare no conflicts of interest.

### Funding Information

The study described in this article was supported by numerous grants, gifts, and donations including those from Cell Signaling Technology Inc., the Comb Family, the Egan Family Foundation, Elysium Health Inc., the Francis Goelet Charitable Lead Trust, the Global Genome Biodiversity Network, the Good Works Foundation, the National Fish and Wildlife Foundation, National Oceanographic and At-

mospheric Administration Grant No. NA19OAR0110303, National Science Foundation Grant No. DBI 1722553 to Northeastern University, New England Biolabs Inc., the Paul M. Angell Family Foundation, the Richard Lounsbery Foundation, the Seafood Industry Research Fund, and numerous generous private donors.

### References

1. WoRMS Editorial Board (2022). World Register of Marine Species. <https://www.marinespecies.org> at VLIZ (accessed January 21, 2022).
2. Mora C, Tittensor DP, Adl S, Simpson AG, Worm B. How many species are there on Earth and in the ocean? *PLoS Biol* 2011;9:e1001127.
3. Costello MJ, Bouchet P, Boxshall G, et al. Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. *PLoS One* 2013;8:e51629.
4. Campbell LD, Astrin JJ, DeSouza Y, et al. The 2018 revision of the ISBER best practices: Summary of changes and the editorial team's development process. *Biopreserv Biobank* 2018;16:3–6.
5. Wieczorek J, Bloom D, Guralnick R, et al. Darwin Core: An evolving community-developed biodiversity data standard. *PLoS One* 2012;7:e29715.
6. Droege G, Barker K, Astrin JJ, et al. The Global Genome Biodiversity Network (GGBN) data portal. *Nucleic Acids Res* 2014;42:D607–D612.
7. Elshahawi SI, Trindade-Silva AE, Hanora A, et al. Boronated tetracycline antibiotic produced by symbiotic cellulose-degrading bacteria in shipworm gills. *Proc Natl Acad Sci USA* 2013;110:E295–E304.
8. Miller BW, Lim AL, Lin Z, et al. Shipworm symbiosis ecology-guided discovery of an antibiotic that kills colistin-resistant *Acinetobacter*. *Cell Chem Biol* 2021;28:1628.e4–1637.e4.
9. O'Connor RM, Nepveux VF, Abenoja J, et al. A symbiotic bacterium of shipworms produces a compound with broad spectrum anti-apicomplexan activity. *PLoS Pathog* 2020;16:e1008600.
10. Underwood BD. *Fishy Business: Seafood Fraud and Mislabeling in New York State Supermarkets*. New York: Office of the New York State Attorney General; 2018: 1–13.

Address correspondence to:

Daniel L. Distel, PhD

Ocean Genome Legacy Center

Department of Marine and Environmental Science

Northeastern University

Nahant, MA 01908

USA

E-mail: [d.distel@northeastern.edu](mailto:d.distel@northeastern.edu)