

# Contents

|     |  |   |
|-----|--|---|
| 1   | Minimum Information about an Omics Protocol (MIOP) . . . . . | 1 |
| 1.1 | Methodology Categories . . . . .                             | 3 |
| 1.2 | Maturity Levels . . . . .                                    | 5 |
| 1.3 | Helpful Ontologies . . . . .                                 | 5 |

## 1 Minimum Information about an Omics Protocol (MIOP)

| MIOP Term            | Example Value                                   | MIOP Definition  |
|----------------------|---|--|
| methodology category | sample extraction and purification              | Methodology category which the uploaded protocol belongs to. This links to the associated methodology categories which precede and succeed it in the workflow, to facilitate the linking of protocols into entire workflows, while keeping granularity and flexibility. This will enable the mixing and matching of protocol modules from various uploaded workflows |
| project              | Marine Biodiversity Observation Network (MBON)  | Details about the project that generated this version of the protocol (e.g., full name of project, acronym(s), principle investigator, lead institute(s), persistent website, funding agency). If possible please provide the most consistent URI associated with your project (e.g. Cordis Project ID for EU projects).   |
| purpose              | biodiversity assessment objective [OBI:0001969] | Terms to describe the purpose of the omics research. [e.g., time series design (OBI:0500020) or taxonomic diversity assessment by targeted gene survey (OBI:0001960)].   |
| analyses             | DNA extraction [OBI:0000257]                    | Terms to describe the types of analyses used in the protocol [e.g., amplicon sequencing assay [OBI:0002767] or polymerase chain reaction [OBI:0002692]]  |
| geographic location  | Monterey Bay [GAZ:00002509]                     | Geographic location/s in which the protocol has been used [e.g., Hawaii Ocean Time-series Site [GAZ:00187530], Western English Channel Sampling Stations [GAZ:00187525]]   |

| MIOP Term                         | Example Value  | MIOP Definition   |
|-----------------------------------|--|---|
| broad-scale environmental context | marine biome [ENVO:00000447]   | Biomes in which the protocol was successfully used [e.g., oceanic epipelagic zone biome [ENVO:01000033]]  |
| local environmental context       | oceanic epipelagic zone biome [ENVO:01000033]  | Environmental features targeted using the protocol [e.g., seasonal marine thermocline [ENVO:01000107]]  |
| environmental medium              | sea water [ENVO:00002149]  | Identify the environmental or organismal material from which the biological molecule (e.g., DNA/RNA/Protein) was extracted [e.g., sea water [ENVO:00002149]]  |
| target                            | deoxyribonucleic acid [CHEBI:16991]  | Identify the target taxa, gene and/or molecule for the protocol [e.g., <i>Polaribacter</i> [NCBITaxon:1642819], 16S Mitochondrial Ribosomal RNA [NCIT:C131261]].                                    |
| creator                           | Jacoby Baker,<br><a href="https://orcid.org/0000-0002-0673-7535">https://orcid.org/0000-0002-0673-7535</a> | An entity responsible for making the resource.  |
| materials required                | centrifuge [OBI:0400106] ;<br>incubator [OBI:0000136]  | Terms to identify the key resources needed to complete the protocol [e.g., Illumina MiSeq (EFO:0004205), centrifuge (OBI:0400106)]. When possible use controlled vocabularies using (namespace:ID). |
| skills required                   | sterile technique ; pipetting skills   | Description of expertise required to execute protocol.  |
| time required                     | 1320   | Number of minutes required to complete protocol (round up to nearest whole number).   |
| personnel required                | 1  | Recommended number of people required to execute protocol. For example, when deploying an instrument or when scuba diving, or doing ice work you may need a minimum number of individuals.          |
| language                          | en   | A language of the resource.   |
| issued                            | 2023-11-07   | Date of formal issuance of the resource.  |
| audience                          | scientists   | A class of agents for whom the resource is intended or useful.  |
| publisher                         | Monterey Bay Aquarium<br>Research Institute, Chavez Lab  | An entity responsible for making the resource available.  |

| MIOP Term      | Example Value | MIOP Definition   |
|----------------|---------------|---|
| hasVersion     | 3             | A related resource that is a version, edition, or adaptation of the described resource. |
| license        | CC BY 4.0     | A legal document giving official permission to do something with the resource.          |
| maturity level | Mature        | Enter the maturity level of the methodology in the document.                            |

## 1.1 Methodology Categories

Our 2021 paper describes the overall vision and organizational principle of defining protocols by methodology categories:

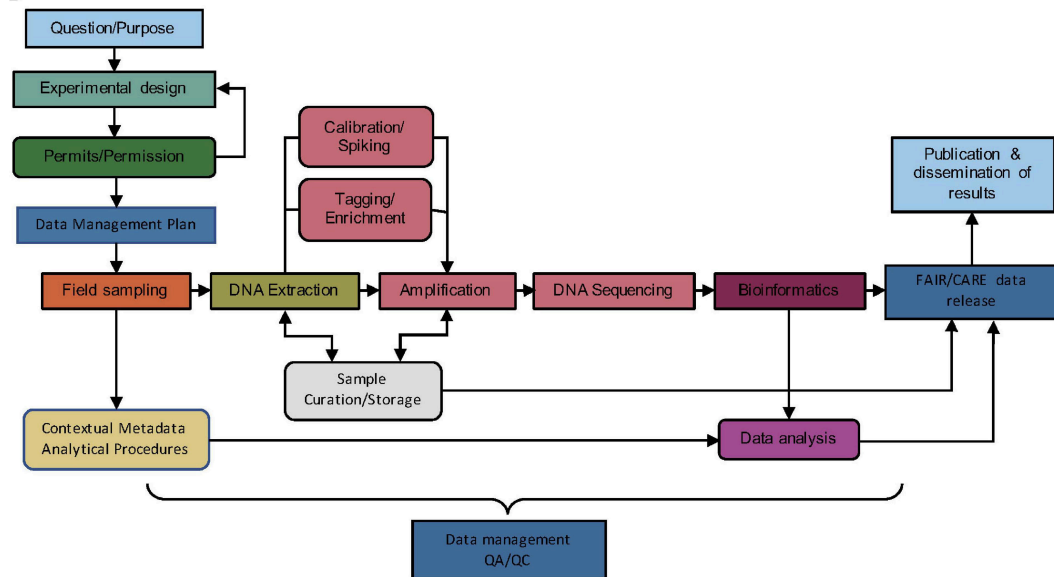
Samuel, Robyn M., et al. "Toward a global public repository of community protocols to encourage best practices in biomolecular ocean observing and research." *Frontiers in Marine Science* 8 (2021): 758694. <https://doi.org/10.3389/fmars.2021.758694>

**Figure 1 (A)** Proposed methodology categories to enhance exchange of ocean omics analysis knowhow. Protocols, guidelines, and other methodologies in some of these categories (such as Sample archiving/biobanking, Data Management, and Society) are cross-cutting and may apply at multiple points in the workflow. **(B)** Example workflow for a DNA metabarcoding project. Colors correspond to the methodology categories outlined in panel (A) and arrows indicate the order of the workflow. Square boxes show essential steps in a metabarcoding workflow, whereas rounded boxes indicate non-essential steps. Data management and QA/QC are required throughout the entire workflow.

**A**

|  |  |
|--|--|
| 1. Society                                   | Documenting societal inputs and outputs (e.g. mission statements (linked to broadly accepted societal priorities e.g. SDGs EOVs, EBVs. etc), stakeholder consultation outputs, associated publications/media, links to relevant data repositories, educational materials associated with the workflow).  |
| 2. Design and logistics                      | Documenting plans for the experimental/observational design (e.g. cruise schedules/reports, deployment regimes, statistical replication, logistics, costing/resources/material lists)  |
| 3. Ethics and law                            | Documenting permits, permissions and guidelines for obtaining samples and release data. (e.g., Internationally Recognized Certificate of Compliance (IRCC) code for sampling, or relevant Due Diligence documentation if an IRCC code is not required, for countries that have signed and ratified the Nagoya protocol (Nagoya Protocol, 2010)) checklists showing compliance to ethical principle (e.g. OCAP) for ethical sampling on indigenous lands, etc |
| 4. Data management                           | Protocols for data and quality management. This will include data management plans with quality assurance and quality check measures to be undertaken throughout the workflow, as well as checklists and guidance documents for the compliance with FAIR/CARE data standards; and links to all data repositories to which the omics study (meta)data should be uploaded.   |
| 5. Sampling Collection Procedure             | Protocols for the physical collection of samples to be used in omics research. Including both environmental samples (e.g. water, sediments, biofilms) and organismal samples (e.g. gut contents, bulk plankton, sponge).   |
| 6. Contextual Metadata Analytical Procedures | Protocols used for the measurement of additional parameters (e.g. temperature, salinity, pH, chlorophyll etc.)   |
| 7. Sample Extraction & Purification          | Protocols for the extraction and purification of biological molecules (e.g. DNA, RNA, Proteins) from environmental and/or organismal samples.  |
| 8. Omics Sequencing Procedures               | Protocols for sequencing nucleic acids. This stage may contain sub-stages that occur during the sample extraction and purification stage, if so, links between the two stages should be provided.  |
| 9. Bioinformatics                            | Bioinformatics pipeline for processing sequence data. This can include links to github, docker etc.  |
| 10. Quantitative amplification               | Protocols for the quantitative amplification of DNA and RNA. Including quantitative PCR (qPCR), reverse transcriptase quantitative PCR (rt-qPCR), nucleic acid sequence-based amplification (NASBA), Recombinase Polymerase Amplification (RPA) etc.   |
| 11. Data Analysis                            | Protocols for analysis of data. This will include scripts for statistical tests, modelling and data visualisations.  |
| 12. Sample archiving/biobanking              | Protocols for archiving physical samples at all stages of laboratory processing (e.g. whole environmental/organismal samples, DNA/RNA/protein extracts and/or amplicons)   |

**B**



Since publication we have further refined these methodology categories to include:

- **society:** All workflows should begin and end with society; societal needs inform the question or purpose behind the research, and societal impacts show the value in the research once it has been completed.
- **design and logistics:** This category covers the practical logistics for implementing ocean omics research and operations, including the experimental/observational design formulated to address the societal priorities
- **ethics and law:** A survey of workshop participants highlighted a need for guidance on sharing data and complying with important ethical and legal requirements (Simpson et al., 2021). This category will include information on permits and permission required to obtain samples and release data. Collating and publishing this information will firstly provide examples for how previous projects have adhered to legal requirements/ethical principles and secondly stimulate discussion on how to facilitate adherence

to these requirements and principles, perhaps through checklists, templates, or training materials.

- **data management:** The data management plan (DMP) is designed to support all the downstream steps according to the ethics, legalities and societal needs identified in society, design and logistics, and ethics and law, while making sure that the (meta)data flows to the right stakeholders in society that we need to interface with. DMPs should be drafted prior to data collection and referred to throughout the workflow to ensure that quality assurance and quality checks take place, and that detailed information on (meta)data requirements for both short and long-term (meta)data storage is given. There is a growing body of tools and best practices surrounding DMPs, including principles for making them more machine-actionable, that should be leveraged in omic protocols and associated infrastructure (see Miksa et al., 2019). Publishing documentation on omics specific DMPs will increase transparency for funders by providing direct links to the protocols they refer to. Furthermore, collating examples of omics specific DMPs will provide insight into what the community needs from omics specific data management tools.
- **sampling collection procedure:** Protocols for the physical collection of the sample.
- **contextual field data:** (*Adapted from Contextual Metadata Analytical Procedures*) Protocols for the measurement of additional parameters.
- **sample extraction and purification:** Protocols for the extraction and purification of biological molecules.
- **omics analysis:** (*Adapted from Omics Sequencing Procedures and Quantitative amplification*) Protocols for wet lab procedures to examine biological molecules such as amplification, sequencing, or calibration.
- **bioinformatics:** Bioinformatics pipeline for processing sequence data.
- **data analysis:** Protocols for analysis of data including statistical tests, modelling, and data visualizations.
- **sample archiving or biobanking:** Protocols for archiving physical samples at all stages of processing.

## 1.2 Maturity Levels

- **Mature:** Methodologies are well demonstrated for a given objective, documented and peer reviewed; methods are commonly used by more than one organization.
- **Pilot or Demonstrated:** Methodologies are being demonstrated and validated; limited consensus exists on widespread use or in any given situation.
- **Concept:** A methodology is being developed at one institution(s) but has not been agreed to by the community; requirements and form for a methodology are understood.

## 1.3 Helpful Ontologies

Although not required, it's helpful to use terms that are machine readable so in the future we can more easily organize protocols by MIOP terms.

- Ontology for Biomedical Investigations (**OBI**) (<https://obi-ontology.org/>)
- The Environment Ontology (**ENVO**) (<https://sites.google.com/site/environmentontology/>)

A site to search for ontologies or terms: <https://www.ebi.ac.uk/ols4/ontologies>

---

To suggest changes to MIOP terms or definitions please create a github issue or suggest changes to the definitions defined here: <https://github.com/BeBOP-OBON/miop/blob/main/model/schema/terms.yaml>