

eDNA Metabarcoding Data in the Global Biodiversity Information Facility (GBIF)

Tobias Guldberg Frøslev[‡], Thomas S. Jeppesen[‡], Tim Robertson[‡], Dmitry Schigel[‡]

[‡] Global Biodiversity Information Facility, Copenhagen, Denmark

Corresponding author: Tobias Guldberg Frøslev (tfroeslev@gbif.org)

Abstract

The Global Biodiversity Information Facility ([GBIF](https://www.gbif.org)) is an international network and data infrastructure that provides free and open access to biodiversity data from around the world, enabling scientists, policymakers, and the public to explore and analyze information about the Earth's living organisms. Originally developing at a distance from GBIF, metabarcoding of DNA has become a standard tool for detecting species in bulk samples or environmental samples such as soil, water, and air. Raw sequence data (fastq files) are often shared and deposited in dedicated repositories. Seen from a biodiversity documenting perspective, raw sequences have limited value, as several steps of bioinformatic processing and filtering are needed to arrive at a credible set of sequences that can be interpreted by comparing to a sequence reference database. Most often, such interpreted DNA metabarcoding data come in the shape a table with abundances of so-called Amplicon Sequence Variants (ASV) or Operational Taxonomic Units (OTU) across samples—a so-called ASV/OTU table—and some associated files, e.g., spatiotemporal and other sample metadata, and taxonomic inferences of sequences. In this session we present GBIF state-of-work and plans for developing and improving publishing and standardisation of OTU table-like biodiversity data for easier and broader reuse, including a prototype tool using the OTU table as a publishing model, mapping this research-familiar data format to the Darwin Core standard (Darwin Core Task Group 2009).

Keywords

environmental DNA, OTU tables, FAIR data, biodiversity

Presenting author

Tobias Guldberg Frøslev

Presented at

TDWG 2023

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Darwin Core Task Group (2009) Darwin Core. Biodiversity Information Standards (TDWG). URL: <http://www.tdwg.org/standards/450>