Plan Overview

A Data Management Plan created using DMPTuuli

Title: Adaptation to Global Warming in Resurrected Phytoplankton - AWARE

Creator: Conny Sjöqvist

Principal Investigator: Conny Sjöqvist

Affiliation: Åbo Akademi University

Funder: The Research Council of Finland (former The Academy of Finland)

Template: Data Management Plan - Research Council of Finland

ORCID iD: 0000-0003-2562-0217

Project abstract:

There is an urgent need for a better understanding of the adaptive potential in marine phytoplankton if we are to understand and predict future ecosystem responses to global warming accurately. A key challenge is the limited awareness of evolution in a 'real-life' global warming scenario. The aim of AWARE is to reveal novel, yet undescribed biodiversity and assess the scope for realistic adaptation to global warming across European coastal waters. This is achieved by 1) resurrecting phytoplankton resting stages from the 1850s to current times and by producing thermal performance curves for key species across the European coastline. Further, we will 2) sequence the genomes of selected resurrected species to investigate the genetic basis for, and consequences of, temperature adaption in phytoplankton and 3) investigate potential ecological consequences of thermal adaption by specifically focusing on phytoplankton fatty acid profiles and the link to higher trophic levels. In this way, we will be able to study evolution that has already occurred in relation to documented global warming. Such knowledge is of pivotal value when trying to predict the responses of marine phytoplankton to future warming. My group's recent research using a resurrected diatom species from the Baltic Sea shows that adaptation to progressing global warming has been ongoing since at least the 1960s. This is the first discovery of its kind globally and warrants further research into the topic to clarify the exact mechanisms and the ecological consequences of thermal adaptation in phytoplankton. As a side product, this research will also establish the first paleo-algal strain collection in the world. This collection will be freely available to the research community, enabling studies of phytoplankton adaptation to environmental changes beyond temperature alone.

ID: 28406

Start date: 01-09-2025

End date: 31-08-2029

Last modified: 17-06-2025

Grant number / URL: 367876

Copyright information:

Adaptation to Global Warming in Resurrected Phytoplankton - AWARE

1. General description of data

1.1 What kinds of data is your research based on? What data will be collected, produced or reused? What file formats will the data be in? Additionally, give a rough estimate of the size of the data produced/collected.

The research is based on experimental data of microorganisms. This entails for example

- growth data (cell counts and fluorescence measurements)
- data on morphological features (cell size, width etc.)
- other molecular data such as fatty acid content of cells

We also reuse geological data from a "sister" project. This relates to data about chronological dating of marine sediment cores and is accessible via the open database created by TREC (Traversing European Coastlines, managed by the European Molecular Biology Laboratory) once all experimental data from the main project is published.

One part of the research is also focused on producing genomic data (DNA sequences)

The used file formats are

- .txt, xls in case of growth, morphological and geological data
- The raw genomic data is in .fastq format.
- We estimate to produce around 2TB of data, including all different formats.

1.2 How will the consistency and quality of data be controlled?

In all data conversions, maintaining the original information content will be ensured. We will organise training sessions and set data management guidelines for the project group.

2. Ethical and legal compliance

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

No ethical or GDPR aspects are related to our data management.

2.2 How will you manage the rights of the data you use, produce and share?

We make all of the research data and code created within the research project available for reuse under the Creative Commons license CC BY-NC. Regarding legal rights to the data among project partners, ÅAU will provide legal assistance to write the necessary agreement with the collaborators at the beginning of the project.

3. Documentation and metadata

3. How will you document your data in order to make the data findable, accessible, interoperable and re-usable for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

Once published in a scientific journal, all related data, code, and metadata are made available in a separate Zenodo data repository (https://zenodo.org/) (one per publication). The data and the creation of them are explained in detail in an associated README file to enable re-use by other users. Datasets published in Zenodo will have DOI identifiers to ensure long-term accessibility and citability.

4. Storage and backup during the research project

4.1 Where will your data be stored, and how will the data be backed up?

We store all experimental data on Åbo Akademi University OneDrive, with backups, and where previous versions of files are also accessible. The genomic data is stored partly on the CSC Puhti cluster (for data in active use) and we use the CSC Allas service for long-term storage. In addition, all data is backed up on an external personal hard drive monthly.

4.2 Who will be responsible for controlling access to your data, and how will secured access be controlled?

The PI will control the access to all data produced within this research project. We do not produce data with security issues.

5. Opening, publishing and archiving the data after the research project

5.1 What part of the data can be made openly available or published? Where and when will the data, or its metadata, be made available?

All produced data with accessory metadata will be made publicly available through Zenodo once the manuscripts are published.

5.2 Where will data with long-term value be archived, and for how long?

For this purpose, we use CSC's Allas service, where it will be stored as long as the project is active in CSC. After this, the data is stored on an external hard drive and is also accessible through the published articles. All genomic data will, in addition, be deposited in NCBI GenBank, where it is stored as open access indefinitely.

6. Data management responsibilities and resources

6.1 Who (for example role, position, and institution) will be responsible for data management?

The PI of the research project (researcher at Åbo Akademi University) is responsible for data management. The PhD student (NN) within the project will describe the data and metadata and also transfer it to Zenodo upon publication of the research articles.

6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?

Currently the data storage on CSC's Allas and Zenodo services is free of cost for academic users.