

mARS Workshop Report

Marine Biology Lab, ULB, Brussels, 9-13 December 2013

The mARS (Microbial Antarctic Resource System) workshop was organised in Brussels, BELGIUM, and was hosted by the Marine Biology lab at the Université Libre de Bruxelles (ULB).

Participants

- Dr Alison Murray (USA, Desert Research Institute)
- Dr Anton Van de Putte (BE, Royal Belgian Institute for Natural Sciences)
- Dr Bruno Danis (BE, Université Libre de Bruxelles)
- Ir Nabil Youdjou (BE, Université Libre de Bruxelles)
- Dr Isabelle George (BE, Université Libre de Bruxelles)
- Dr Thomas Pommier (FR, Université Lyon1) (remote participation)

Background

The concept for mARS originated as a scientific community need - recognition of the increasing levels of Antarctic microbial diversity being revealed through DNA sequencing efforts, which to date, has not yet been captured in a single database. The Scientific Committee on Antarctic Research (SCAR) has long been interested in biodiversity, and has had a long-standing recognition in the importance of data management representing the extent of Antarctic biodiversity, starting initially in the marine environment with the SCAR MARBIN database. This effort subsequently expanded to include terrestrial biodiversity reflected in the ANTABIF database, which has evolved now to Biodiversity.aq. The mARS project is currently an element of the biodiversity.aq platform being supported primarily by the Belgian Science Policy Office (BELSPO).

The mARS workshop gathered participants from different backgrounds (e.g. Microbial Ecology, Programme Management, IT, Biodiversity Informatics). The main objective of the workshop was to implement a (meta)data publication process for molecular microbial diversity information in order to allow the integration of microbial resources and diversity-oriented research projects using the biodiversity.aq information system.

Progress

The workshop focused on the examination of data standards, the development of new documentation describing the protocols for using mARS, and the implementation of dataflows, including their testing by selected users.

The mARS workshop covered the following items:

1. Synthesized comments on the White Paper documenting the elements envisioned to achieve a usable data system (available to the Antarctic microbial community July 2013). This paper is now available on the mARS website as a **Vision Paper** which sets the ground for further development of the mARS initiative.
2. Creation of draft version of a user-information interface for submission of sequence set information to communicate to mARS database administrators, including definitions of terms.
3. Preparation of example data sets for both sequence set and environmental data information that will be made available to users.
4. Development of the first version of the mARS Standard Operating Procedures (**SOPs**)
An initial iteration of mARS Standard Operating Procedures (SOPs) was prepared evaluated and published online (mars.biodiversity.aq/howto). The SOP describes, step-by-step how users can participate in the mARS project.
5. Preparation of a set of **FAQ**
Based on user feedback a set of Frequently Asked Questions (FAQ's) was prepared (mars.biodiversity.aq/howto)
6. Deployment of a new version of the **mARS website**, offering access to all new documents and templates prepared during the workshop.
7. First round of **Data upload testing**
SOP's were tested both by workshop participants and by an external microbial ecology expert, in order to adjust the content of the SOP and FAQ sections.
8. Development of the **2014 roadmap** (see below)
The MARS workshop produced a workplan to reach a determined set of milestones prior to the upcoming SCAR Open Science Conference (August 2014).

Perspectives

By harboring this information directly at mARS.biodiversity.aq, Antarctic scientists will have

the ability to archive project information that will be accessible through common language queries. There is currently NO exhaustive resource that provides the level of information and integration which is envisaged for the mARS initiative.

For the first time, mARS will make sequence-based Antarctic microbial diversity information discoverable through a single, integrated and open access point on the Internet. An additional aim of this initiative is to provide information access for the broad scientific community, including conservation and management interests, in the spirit of the Antarctic Treaty.

Roadmap for 2014

Outreach/communication

1. Organise mARS event at SCAR OSC
2. Disseminate the results of the mothur analysis
3. Communicate to a large set of potential users once steps 1-3 of the mARS vision paper have been implemented

Beta-testing

1. Communicate to a targeted set of beta testers with sequence set(s) in public repositories (GenBank) and agree to upload their metadata.
2. Beta-test the mother pipeline that will be used to analyze data in mARS; Write up a literal protocol; and teach this pipeline to whoever will actually do the data analysis.

Technical Developments to achieve steps 1-3 of the mARS vision

1. Develop a graphical tool to represent the 16S rRNA gene sequence data (in mARS) based on gene coordinates (e.g. variable regions 6-8 in Bacteria, Variable region 13 in Archaea, and 1-3 in Eukarya) and sequence effort (1-100,000's per run).
2. Related to (1), generate sequence alignment graph that has links to the sequence sets represented in mARS to make this information discoverable.
3. Initiate ability of users to register metagenome/ metatranscriptome/ metaproteome/ metabolomics data with links to archived data.
4. Represent the data types in #'s 1 and 2 using map server to enable visualization of the geographical scope of mARS data sets.
5. Explore the possibility to mine PDFs for GenBank sequences
6. Make files uploaded through the IPT searchable and retrievable (for the user & the managers), through new application (Note check if this can be achieved with an IPT update)
7. Explore the possibility to automate metadata flows to GCMD/AMD (liaise with SCADM/OBIS)
8. Scoping and requirements of data retrieval application (Advanced search)

Future Ideas (towards achieving step 4 of the mARS Vision)

1. Explore the inclusion of a dedicated Antarctic BLAST server to serve selected collections of Antarctic-specific sequence sets (full mARS or subsets).
2. Provide mothur output files as downloadable elements.
3. Construct workflow for web-based representation of microbial diversity for the three domains of life represented in mARS.

Publication Goals

1. mARS Database description in Nucleic Acids Research Database Issue [notify by July 1]
http://www.oxfordjournals.org/our_journals/nar/for_authors/msprep_database.html
2. Develop a publication based on mARS data exploration.

This workshop was funded by the Belgian Science Policy (BELSPO) with additional support from Desert Research Institute (DRI), and Université Libre de Bruxelles (ULB).

