Marine

Microbial Biodiversity,

Bioinformatics & Biotechnology

Grant agreement n°287589

Acronym: Micro B3

Start date of project: 01/01/2012, funded for 48 month

Deliverable 2.7

Report at mid-term meeting

Version: 2.0

Circulated to: OSD Core Team (02.05.2014)

Approved by: Prof. Frank Oliver Glöckner, (06.05.2014)

Expected Submission Date: 31/12/2013

Actual Submission Date: 08/05/2014

Lead Party for Deliverable: UOXF
Mail: osd-contact@microb3.eu

Public (PU)	X
Restricted to other programme participants (including the Commission Services) (PP)	
Restricted to a group specified by the consortium (including the Commission Services) (RE)	
Confidential, only for members of the consortium (including the Commission Services) (CO)	

Summary

We reported on the progress of OSD at the General Assembly (GA) in Bremen, April 2014. Here we summarize the highlights of the formal presentation and decisions taken during the meeting with respect to OSD pilot events so far, plans for the main event on June 21st 2014 and forward planning for 2015 and beyond. Key highlights include: 1) progress towards the formation of an OSD Consortium and a first international meeting at GSC 16 in Oxford April 2014, 2) completion of 4 pilot events to date, 3) registration of 130+ sites for OSD 2014. In addition, at the GA we voted to spend Micro B3 budget on data generation for 2014 and completed Deliverable D2.6 with regards to this process and outcome. We also ratified the OSD Data Policy. Overall support for OSD continues to grow with Micro B3 and the external community and we look forward to focusing on data analysis in the coming two years. We will also focus on exploring options for the continuation of OSD into 2015 and beyond the duration of Micro B3.

1. Formation of an OSD Consortium and 1st International OSD Meeting at GSC 16 in Oxford, April 2014

In order to sustain the efforts of OSD and give credit to all members we continue to work towards formalizing an OSD Consortium. We have agreed that our first joint publication will be on the sites registry as this is a way to unify the community. We had the first meeting of this Consortium within the GSC 16 meeting in Oxford in April 2014. OSD Coordinators as well as many OSD Site Coordinators were present and gave talks. The whole workflow was reviewed, lessons learned from sites were presented as well as discussion panels were held on future and strategic developments of OSD. Strong interest in taking OSD forward into the future was the main highlight of the meeting.

2. Completion of 4 Pilot Events to Date

Figure 1 shows the summary of progress and participation at the 4 OSD pilot events held over the past two years. As OSD progresses we are adding more features. The next features will be submission to ENA and analysis of data. As can be seen by reading the numbers in the figure from top left to bottom right the numbers of participating sites and the outputs generated continues to grow.
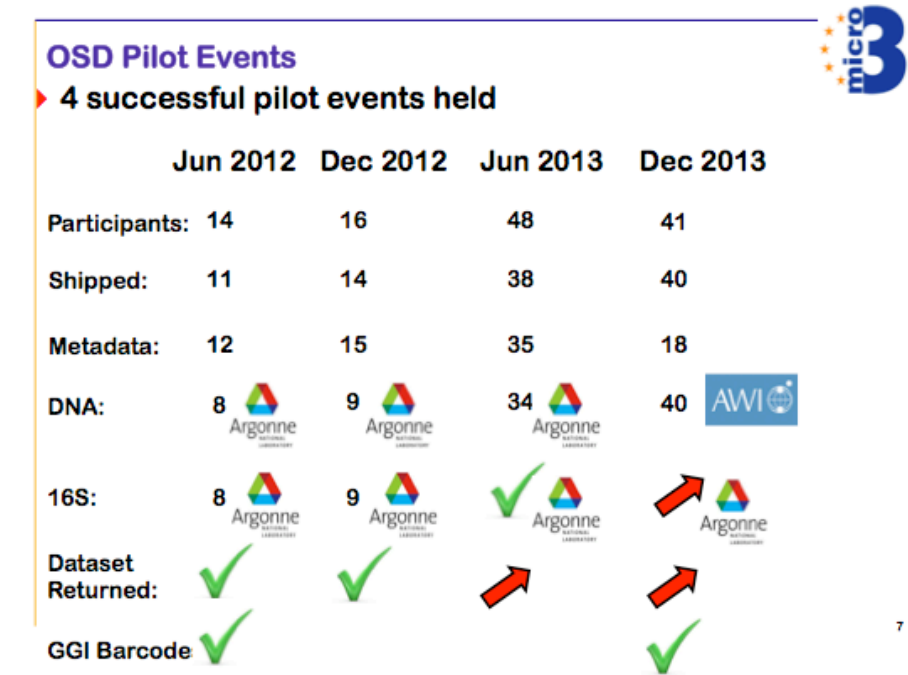
Figure 1. OSD Pilot Events

3. Registration of 130+ Sites and Plans for OSD 2014

Active recruitment of sites has led to the formal registration of 130+ sites. The formal registration deadline was May 1. The OSD workflow is continuing to be put in place and is described in Figure 2. Through communications and consultations with other WP Leaders, OSD Core Team has defined the timeline for 2014. Key milestones and targets are shown in Figure 3.
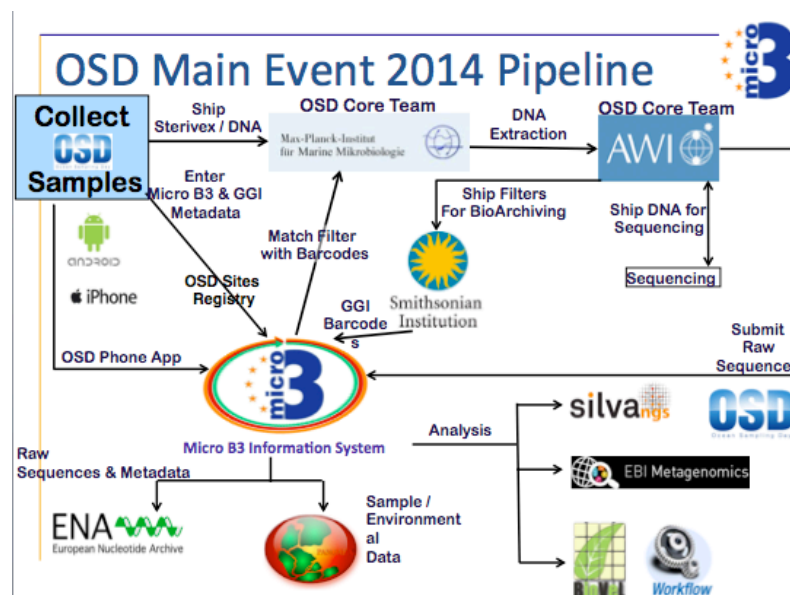

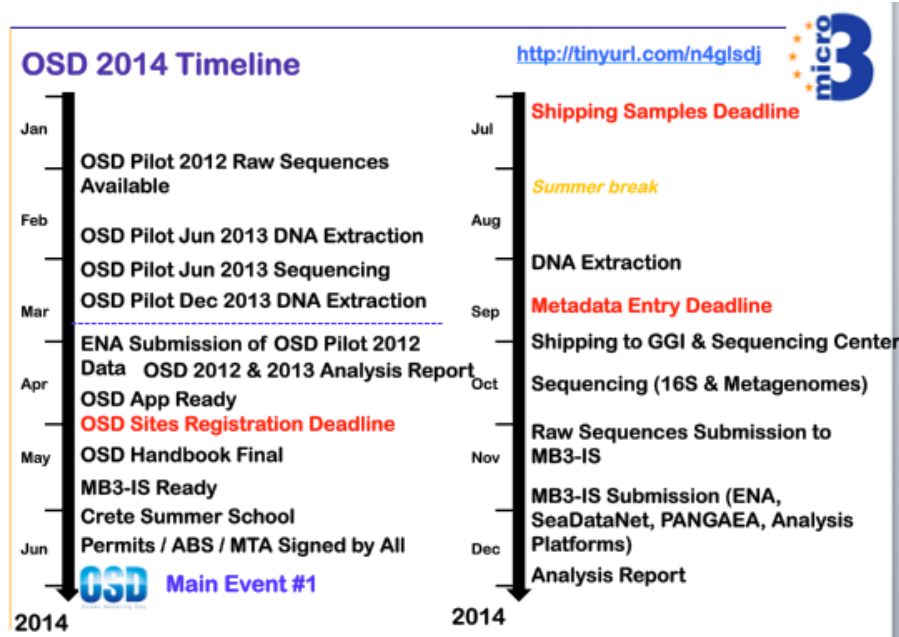Figure 2. OSD Main Event 2014 Pipeline

Figure 3. OSD 2014 Timeline

4. Micro B3 Budget on Data Generation for 2014

Figure 4 shows a summary of the data to be generated via the Main OSD Event on June 21st, 2014. A part of the GA we completed D2.6 describing how we formed an OSD Committee for deciding how to spend Micro B3 budget on OSD. We agreed to allocate 80K Euro for sequencing and will spend 43K Euro on rRNA and shotgun metagenomic sequencing for the June 2014 event.

Figure 4. Data Generation

5. OSD Data Policy

At the GA we also reviewed the OSD Data Policy and decided to add a clause on the fact that minimum required metadata must be submitted to Micro B3/OSD prior to release of the sequencing data.

Micro B3 / OSD Data Policy is provided in Annex 1, and can also be found on the OSD Blog via the following link:

http://oceansamplingday.blogspot.co.uk/2014/05/final-micro-b3-osd-data-policy.html

We agreed that all sites would electronically sign the data policy and include the original paperwork together with shipping of samples after the main event.

6. Formalized plans to continue OSD into 2015.

Due to ongoing support for OSD we formally decided to continue and host a 2015 event. Our focus going forward for the next two years is on continuing to ramp up sites / involvement and on data analysis. We presented preliminary data analysis of pilot data (Figures 5-7) showing that early data suggests high diversity in winter and expected geographical clustering by water body. We will look for future funding to sustain OSD beyond 2015 and welcome input from the community on ways to do so.
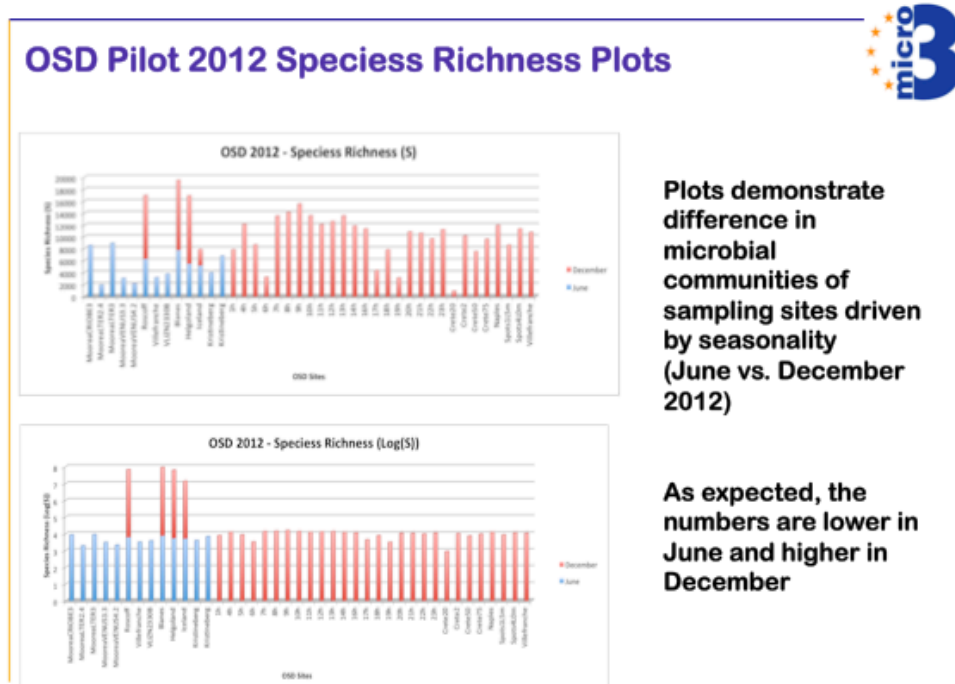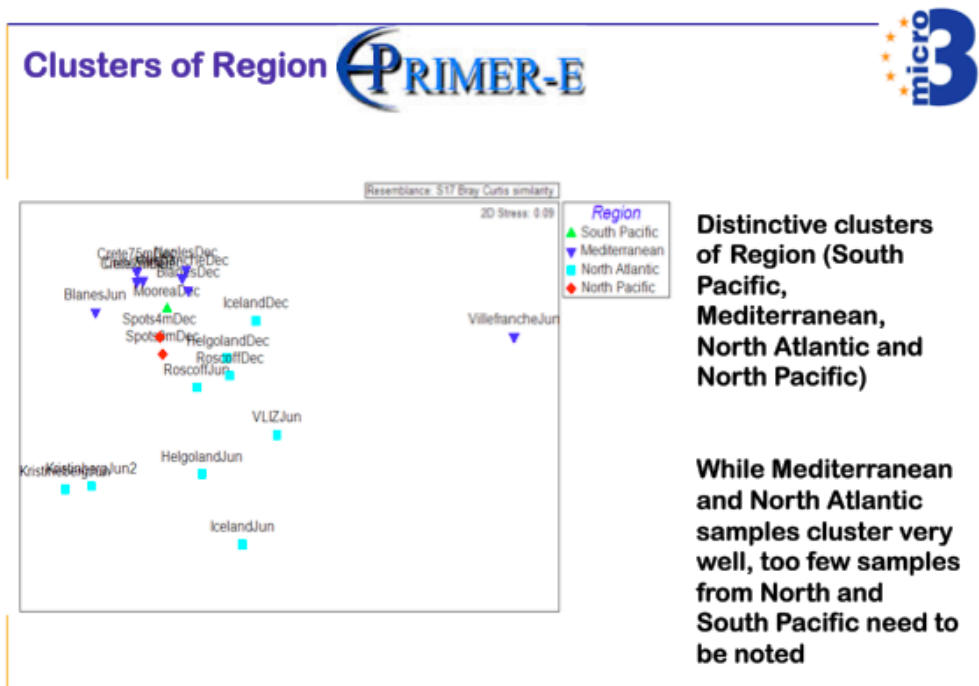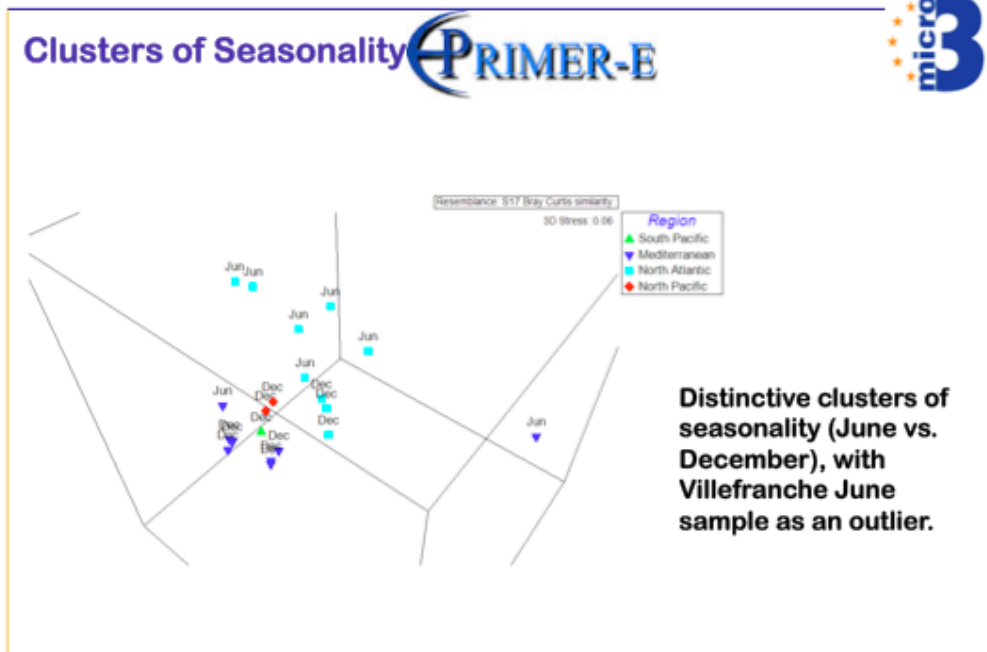
Figures 5-7. Preliminary data analysis of pilot data showing diversity may be higher in winter (5), seasonal variation (6) and geographic clustering (7).

ANNEX I

Final Micro B3 / OSD Data Policy

Authored by the OSD Data Policy Working Group with the OSD Consortium

OSD Data Policy Working Group

Dawn Field, Oxford University, OSD, WP2 Leader

Mesude Bicak, Oxford University, OSD Co-ordinator/Bioinformatician

Anna Klindworth, MPI Bremen, OSD Co-ordinator

Frank Oliver Glöckner, Jacobs University Bremen & MPI Bremen, Micro B3 Coordinator

Arianna Broggiato, BIOGOV, ABS (WP8)

Arul Scaria, BIOGOV, Data Policy (WP8)

Tom Dedeurwaerdere Univ Louvain WP8 Leader

Alban Ramette, AWI, Biostatistics (WP5 & WP6)

Linda-Amaral Zettler, Woods Hole, MicroB3 Scientific Advisory Role (ICOMM MIRADA LTER)

Petra ten Hoopen, EBI, Metadata & Standards (WP2 & WP4)

Michelle Barbier, CIESM, Med Co-ordinator (WP2, 4, 8 & 9)

Neil Davies, UC Berkeley, GOs Network, WAML, Pacific Region

Jack Gilbert, Argonne National Labs, EMP

Katie Barker, Smithsonian Institution, GGI Bio-Archiving

Status to date

This data policy covers the collection, dissemination, analysis and publication of OSD data. It was developed over the past years of OSD pilot events. Within Micro B3, WP8 is responsible for the Micro B3 data policy. The OSD policy is based on this policy and consultation with the OSD Consortium. It was finalized at the Micro B3 General Assembly (GA) April 2014. It remains the responsibility of the OSD Data Policy working group to make sure the policy is fit for purpose and all are aware of it when submitting and using data of the OSD Consortium. If you would like to join the group please contact osd-contact@microb3.eu.We welcome input from all OSD participants on shaping of the OSD data policy. Please send comments to osd-contact@microb3.eu.

Background

The OSD Data Policy was created by drawing on existing policies (See Biosharing: http://biosharing.org/policies) and current best practice in genomics and the needs of Micro B3 and the OSD community. It was developed and will be maintained according to the "12 steps to a data policy" outlined in *Field et al (2009) Science:*

12 Step Path to a Data Policy:

http://blog.biosharing.org/2010/02/12-steps-to-creating-data-sharing.html

Tenets of Ocean Sampling Day (The policy applies to all in the OSD Consortium. It is a mandatory part of taking part in OSD)

The OSD data set is generated by the members of the OSD Consortium - to be part of the Consortium you must be formally registered and substantively contribute to the collaboration.

The OSD dataset will be a reference data set and should be as widely accessible and used to support downstream research as possible.

As OSD is one event (a one-off event) it should be easier to get all contributors (their official sites) to release the data publicly.

All OSD data must be standards compliant - must include metadata - this includes compliance with the Micro B3 standard, which includes the GSC's MIxS standard.

All OSD genomic data must be tightly linked to environmental data - contextual information is a top priority.

Sequence data without deposition of accompanying metadata (minimum of lat, long, time, depth, temp, salinity according to OSD Handbook) will not be released after the June solstice 2014 event. The OSD Consortium will work towards release of all data.

The OSD data policy covers the pilot events, the main OSD event and any subsequent events.

Data must be submitted to the public repositories (e.g. INSDC). Micro B3 will make sure 2012-2015 data sets are submitted to the ENA/MG-Portal.

All data will be processed/stored in the Micro B3 infrastructure as it is built and the full data set will eventually be available through this infrastructure. Data can flow in many analysis pipelines/databases - we welcome its wider use

All samples must be collected legally (with appropriate permits) and individual collectors are responsible for making sure of this. Specifically ABS/MTA/DTA agreements will be in place as required (bioarchiving at GGI requires OSD MTA, proof of relevant ABS and metadata)

The OSD Consortium

This data policy applies to everyone in the OSD Consortium (to date this means everyone who has constructively contributed to OSD sampling events).

Release of the data and metadata

At the 2013 EEB meeting of the Micro B3 project we agree on the importance of the OSD data set as a whole (the sum of participating sites) and the desire to support open sharing of the data. The EEB agreed to work towards: "immediate release of all OSD data as soon as sequenced and quality checked to the public". In this, sense "public" means beyond Micro B3 and OSD participants to the public at large. To protect the rights of the consortium to publish the first global analysis of the data, the data will be released under Ft Lauderdale principles. This is routinely used in large sequencing projects and promotes the use of the data by the wider community while safeguarding the scientific contributions of the data generators. For example, a complete Neanderthal genome was released to the public prior to publication in 2013 under these principle (see: http://www.eva.mpg.de/neandertal/index.html).

This statement will accompany all access to OSD sequence data.

Use of the sequence data

All data will be made freely available. However, we ask users to observe the Ft. Lauderdale principles, which entitles the data producers to make the first presentation and publish the first genome-wide analysis of the data. The data can be used freely for studies of individual genes or other individual features of these sequences.

Collaborative Data Analysis

We hope that the OSD reference data set will be helpful in a variety of ways and we encourage those interested in analysis of the data to declare their interests and formally

join the OSD Data Analysis working group. Together, we are working towards "OSD reports" that will be generated following each event. Ideally, each report will provide a high level overview of the event and, if multiple events are available, comparison and contextualization with past data. Predictions about future events may also be included. It is our vision that this type of transparent collaboration where interested parties declare how they would like to contribute to the data analysis, will help maximize the efforts of the community and build the strongest possible biological interpretation of the data.

Publication

We expect the authorship of the global analysis of the OSD data set to belong to the OSD Consortium. Authorship will be defined on a case-by-case basis for more specialist publications.

Evolution of this Policy

According to the "12 steps for policy development" process, this policy may evolve according to best practice. As agreed at the May 2013 EEB meeting of Micro B3, changes to the policy will be done in consultation with the community. Specifically, policy issues will be taken forward through the OSD Data Policy Working group who is responsible for building consensus on any issues or changes. Final authority for any changes and for upholding the policy will rest with the leadership of Micro B3

Signed by

Site ID: …………………………………………….............

Site Name: …………………………………………….............

Institution: …………………………………………….............

Country: …………………………………………….............

Site Coordinator: …………………………………………….............

Signature: ....

Date: ………………………………………………….