Marine

Microbial Biodiversity,

Bioinformatics & Biotechnology

Grant agreement n°287589

Acronym: Micro B3

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

Deliverable 2.10

Final registry completed and made public through the Micro B3 Catalogue

Version: 3.0

Circulated to: OSD Core Team, 16/07/2015

Approved by: Prof. Frank Oliver Glöckner, 17.07.2015

Expected Submission Date: 30/06/2015

Actual Submission Date: 24/07/2015

Lead Party for Deliverable: UOXF
Mail: mesude.bicak@oerc.ox.ac.uk, osd-contact@microb3.eu   Tel.: +441865610600

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

Following on from the success of OSD 2014, OSD Coordination Team received significant interest from the OSD Consortium towards holding another event in 2015. This encouraged the OSD Coordination Team to organise another big OSD sampling campaign in 2015 on the June Solstice. This time efforts were heavily focused on growing OSD Citizen Science and consequently a full-fletched MyOSD 2015 campaign was organised in cooperation with the OSD sites as MyOSD hubs. This deliverable summarises the latest statistics from the publicly available OSD Sites Registry after the recent OSD and MyOSD 2015 Events held in June 2015 Solstice. As explained in D2.9, an OSD Analysis Core Group (OACG) was established in October 2014 to coordinate analysis of OSD 2014 dataset. This has now transformed into the OSD Analysis Consortium, an active group of 132 experts from the OSD Consortium, aiming to collectively analyse the OSD 2014 dataset. The establishment, recent progress and future plans of OSD Analysis Consortium are detailed in this report.

Table of Contents

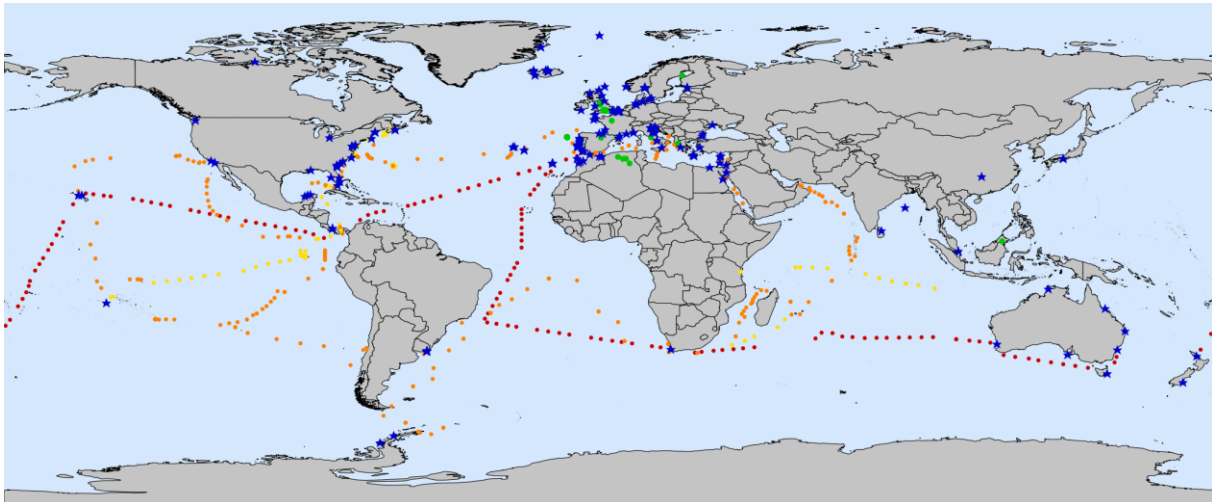## 1. OSD Main Events & OSD Sites Registry Statistics

The OSD campaigns have achieved a glorious success curve over the years, starting with only 20 sampling sites in the first OSD pilot event back in June 2012, now resulting in worldwide recognition with numerous news articles in press, featuring in TV programs as well as taking over social media.

As explained in D2.9, the OSD Sites Registry is an online database of all OSD sampling sites that have registered to take part in OSD, which is a mandatory process for participation as well as recognition as part of the OSD Consortium (introduced in D5.7). Through this resource, it is possible to view a list of all participating sites together with their basic metadata (Site name, Site coordinator(s), Institution affiliated, Country, Sampling site latitude and longitude, Institution latitude and longitude) in a tabular format, as well as via a dynamic interactive map of all sites, referred to as the OSD Map. OSD 2014 Event was held on June Solstice of 2014, with 191 registered OSD sites reported in the OSD Sites Registry. The list of OSD sites along with their basic metadata has been made publicly available via the OSD Sites Registry soon after the OSD 2014 event.

The global success of OSD 2014, strong interest from all existing participants resulted in OSD Coordination Team deciding to hold another event in June 2015. Funds were identified to be sufficient to cover an exact repeat of the previous event with sequencing of same number of samples. However soon after OSD Coordination Team announced an OSD Event 2015 among the OSD Consortium, tremendous interest for participation swarmed in from brand new institutions and sampling sites. Funds permitting, OSD Coordination Team decided to accept 9 more sampling sites prioritising their geographical coordinates, with an aim to expand the global coverage of OSD Map. As a result, there are currently 200 OSD sampling sites registered in the OSD Sites Registry, publicly available along with their basic metadata. We are currently unable to report exact numbers for participation since July 10th has been set as the deadline for all OSD 2015 participants to ship their samples to MPIMM in Bremen.

The OSD Sites Registry along with the OSD Map are continuously maintained by the OSD Coordination Team. Both can be viewed via the following link:

http://mb3is.megx.net/osd-registry/list

Figure 1: World map illustrating the current status of global marine sequencing projects. OSD: blue stars, RSD: green dots, Tara Oceans: orange dots, Malaspina cruise: red dots, Global Ocean Sampling (GOS): yellow dots.
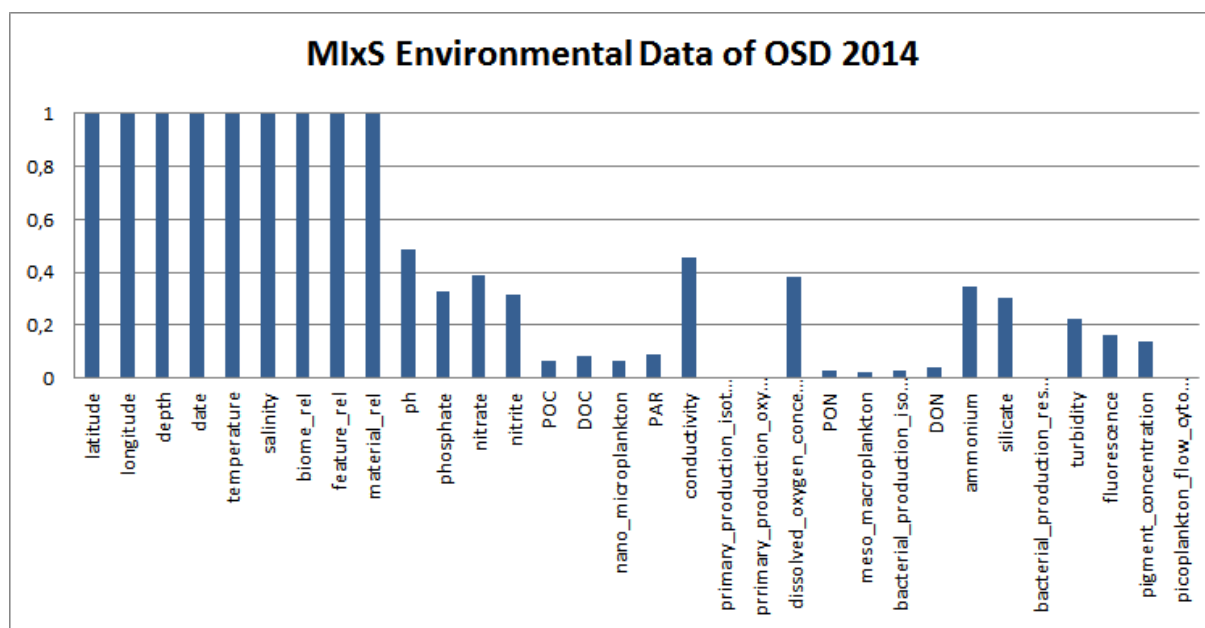
2. OSD 2015 Event & Metadata Capture

Similar to the process in the OSD 2014 event, within the OSD 2015 event too, all OSD sampling sites were asked to complete OSD Logsheets (Annex I of OSD Handbook 3.0, released June 2015) for immediate capturing of their sample metadata during the sampling process. They then sent the hard copy of the OSD Logsheets together with their OSD samples to the members of the OSD Core Team at the MPIMM and JacobsUni in Bremen, Germany. OSD sampling sites were then asked to transfer ALL information from the hardcopy of the OSD Logsheets into the Micro B3 Information System (Micro B3-IS) via the OSD Sample and Registration Page, the online submission portal, via the following link:

http://mb3is.megx.net/osd- registry/sample-registration

For OSD 2015, the existing protocol was updated (NPL022_rnalater) by introducing the use of RNALater by OSD participants, to allow for improved conservation of samples, reducing the risks of degradation of samples arriving especially from long distances, and to reduce the shipping costs for participants.

Furthermore, the submission portal was updated to accommodate internal submission of MyOSD metadata by MyOSD Core Team, further to OSD metadata, allowing MyOSD Core Team to add missing data via the online form as some people only send the data via hard copy of the MyOSD Logsheet.
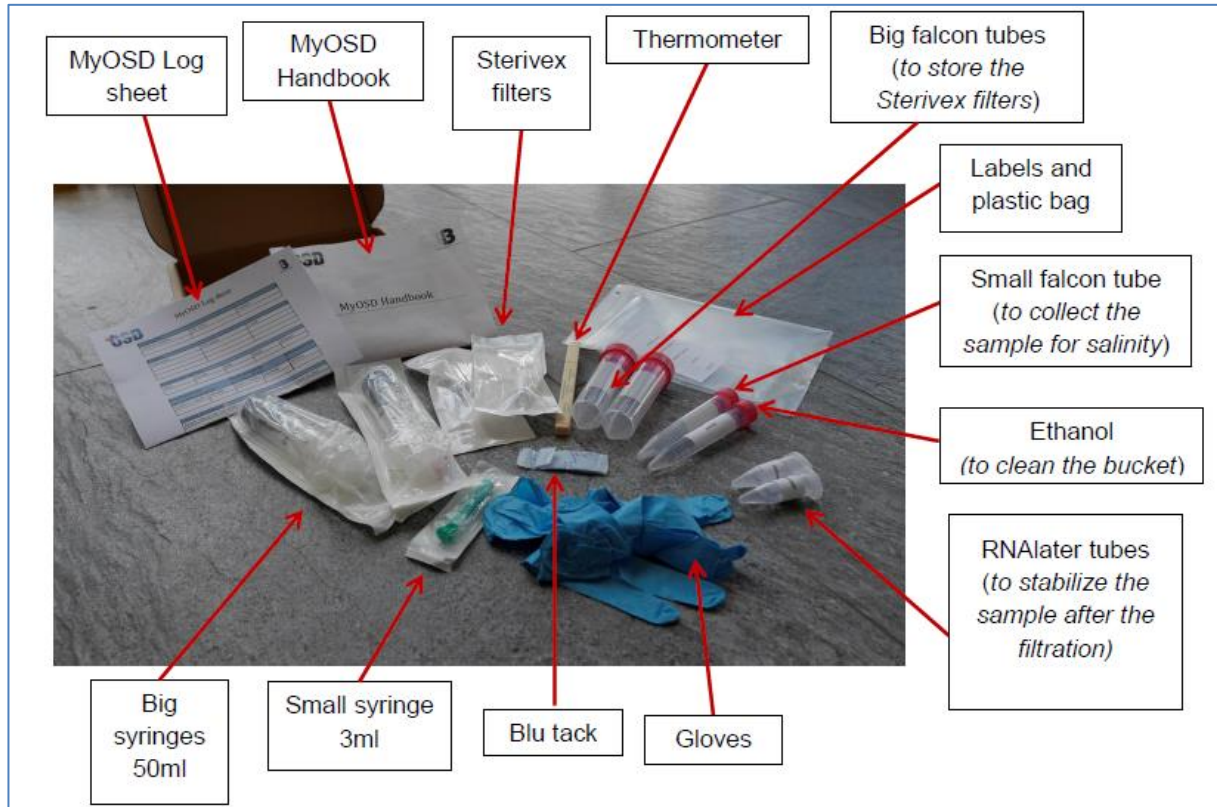
Figure 2: Graphical representation of OSD 2014 Metadata conforming to MIxS and M2B3 standard (Ref 4), meeting all mandatory metadata criteria.

The OSD Sample and Registration online submission portal once again successfully allowed all OSD participants to enter all associated measured metadata directly into an online form (as reported in D5.7 "Complete Micro B3 Catalogue documentation"). All mandatory metadata (at a minimum, meeting the criteria for the M2B3 standards and numerous relevant domain-specific data archives) were again successfully received from all participating OSD sites, proving yet another success by itself, as it will provide the marine community with a follow-up collection of standardized marine data. Figure 2 above further demonstrates this with a graphical representation of OSD 2014 metadata.

3. MyOSD 2015 & Metadata Capture

This year we put a considerable amount of effort into further promoting OSD Citizen Science project called MyOSD (http://www.my-osd.org). In MyOSD 2014, citizen scientists measured environmental parameters, such as water temperature, salinity and uploaded such metadata either via the OSD Citizen Smartphone App or the online submission portal (http://mb3is.megx.net/osd-app/myosd-form) to the Micro B3-IS database. After the event we received a lot of requests from many enthusiastic citizen scientists stating that they would also be very keen to collect actual microbial samples in addition to measuring and providing environmental parameters. Appreciating their enthusiasm, the MyOSD Core Team instantly took this request on board and designed a MyOSD Sampling Kit (shown in Figure 3) that would contain everything a citizen scientist would need to take a microbial sample. This would entail filtration of marine water through a Sterivex filter. We are pleased to state that

MyOSD Kits were sponsored by Millipore, which provided in-kind Sterivex filters to be included as part of the kits.

Figure 3. MyOSD Kit

To follow the guidelines of the Nagoya Protocol, MyOSD Core Team asked existing OSD Sites if they would like to function as a MyOSD Hub. MyOSD Sampling Kits were then distributed to all registered Hubs, which then distributed them to their citizen scientists and taught them about the sampling procedure. After sampling, the citizen scientists returned their Sterivex filters together with marine water samples back to the Hubs, which then sent them all back to the MPIMM. To reach out to further keen citizen scientists far away from the coast and provide them with the opportunity to take part in OSD 2015, as well as to support OSD's sister initiative project River Sampling Day (RSD 2015), we also accepted samples from rivers.

MyOSD's citizen scientists were asked to conduct sampling for MyOSD as well as measure environmental parameters and upload their metadata to the OSD server. Detailed information on participation details were made publicly available via the following link:

http://www.microb3.eu/myosd/how-join-myosd

A MyOSD Handbook (available via the MyOSD homepage, www.my-osd.org) was also prepared by the OSD Coordination Team, which was included within the MyOSD Kits and shipped to all MyOSD Hubs. Moreover, a video tutorial was prepared to guide the citizen scientists in a more informative and efficient manner, which can be found via the following link:

https://www.youtube.com/watch?v=Mtmp4Ltr4xo

Just like OSD Logsheets, MyOSD Logsheets (available via the MyOSD homepage) were prepared and distributed along with the MyOSD Kits, to allow their efficient capture of environmental parameters manually. They are to be filled manually during sampling, hard copies to be shipped to OSD Coordination Team together with samples, and its contents to be digitally submitted to Micro B3-IS either via the OSD App or via a separate online submission portal especially put together for MyOSD 2015 (http://mb3is.megx.net/osd-app/myosd-form).

With an aim to reach citizen scientists globally in a more effective way, the MyOSD webpage was updated in multiple languages, including Spanish, Portuguese, German and Italian. To promote the effort, various dissemination material, such as posters, flyers and drawings for children were prepared. To allow existing OSD Site Coordinators to spread the word locally among their MyOSD Hubs, all such material have been made publicly available and can still be found via the following link:

https://zarafa.mpi-bremen.de/owncloud/public.php?service=files&t=cc452ab3301c7e6061ebb934dc288fed

Social media being one of the key instruments to promote citizen science initiatives, Twitter, Instagram and Facebook accounts for Micro B3 and OSD have been heavily used via the following links with official hashtags of #osd2015, #oceansamplingday, #myosd2015.

Overall, 31 MyOSD Hubs were formed in 17 different countries, and 270 MyOSD Kits (shown in Figure 4) have been distributed to MyOSD Hubs' citizen scientists. As of July 13, 2015 144 MyOSD samples have been returned and this number is expected to increase until the end of the month. A full list of MyOSD Hubs can be found via the following link:

http://www.microb3.eu/myosd/osd-hubs

Figure 4: MyOSD Kits ready to be distributed to the MyOSD Hubs' citizen scientists

Figure 5: MyOSD Kit after successful sampling by a citizen scientist from Germany. Samples were taken at the North Sea**.**

4. OSD Citizen App

The OSD Citizen App was updated to match the MyOSD Sampling Kits; additional metadata fields, such as MyOSD ID and filtration volume were added. Furthermore, smooth functionality of the App on iOS 8 versions was also developed. Finally, the App is now available in six different languages; Catalan, German, English, Italian, Turkish and Spanish.

5. First OSD Publication Published

It was reported in the previous deliverable that the first OSD manuscript titled "Ocean Sampling Day Consortium" was prepared and submitted to the GigaScience Journal. This manuscript aims to formally announce and establish the OSD Consortium, outlining aims, objectives and future plans of OSD serving Micro B3 EU FP7 Project.

We are very pleased to report that this paper was published on June 19th, 2015, just in time before the OSD 2015 Event (Ref 3), which was a huge source of added motivation for the entire OSD Consortium, especially for those preparing for their upcoming sampling efforts. The paper is open-access and can be found via the following link:

http://bit.ly/1CgPmWq

6. OSD Analysis Consortium

6.1 OACG Becomes OSD Analysis Consortium

As explained in D2.9, an OSD Analysis Core Group (OACG) was established in October 2014, to coordinate the analysis of all OSD data in line with the analysis pipeline devised by the Micro B3-IS as well as the submission of all OSD raw sequences and metadata to relevant databases. OACG performed quality checking for the entire OSD 2014 raw sequences and ensured its timely submission to EBI's European Nucleotide Archive for public use by the marine community, including the OSD Consortium.

This was followed by the organisation of an OSD Analysis Jamboree week at the EBI, explained in D9.8 in detail, which brought 30 participants (the majority of whom represented OSD 2014 sampling sites) together with OSD Coordination Team and OACG members. The Jamboree proved extremely successful, resulting in analysis strategies for the OSD 2014 dataset as well as a start to their collaborative implementations. The major outcome of the jamboree was the transformation of OACG into **the establishment of the OSD Analysis Consortium**, consisting of OACG, OSD Jamboree participants, as well as any other experts among the OSD Consortium with a strong interest and passion for the analysis

of OSD 2014 dataset. It was also agreed that given the timely organization and completion of OSD Analysis Consortium efforts and analysis of the OSD 2014 dataset, a collective manuscript will be prepared around the analysis of the OSD 2014 dataset.

6.2 OSD Paper Taskforce

During discussions at the EBI, an OSD Paper Taskforce was also established, which are tasked with organising distinct categories of the paper among OSD Analysis Consortium participants, collating their work and writing the paper. The OSD Paper Taskforce consists of the following experts:

Dr Linda Amaral-Zettler, Dr Mesude Bicak, Professor Chris Bowler, Dr Pier Luigi Buttigieg, Professor Jed Fuhrman, Professor Frank Oliver Glöckner, Dr Antonio Fernandez-Guerra, Dr Pascal Hingamp, Dr Renzo Kottman, Dr Anna Kopf, Dr Daniele Lucidone and Dr Francesca Malfatti.

A mailing list (osd-paper@microb3.eu) was set up to moderate communication among the Taskforce. It was further agreed for both OSD Analysis Consortium and OSD Paper Taskforce to be led by Dr Francesca Malfatti (Scientific Lead) and Dr Mesude Bicak (Scientific & Coordination Lead).

6.3 Open Invitation to Join OSD Analysis Consortium

Shortly after the OSD Jamboree Week, an open invitation was circulated among the OSD Consortium polling for their interest to join and take part in collective analysis of the OSD 2014 dataset. All interested parties were asked to send a page long proposal to osd-paper@microb3.eu by March 31st containing below details. Joint proposals were also highly encouraged to foster collaboration among the Consortium.

- Title

- Feasibility

- Preliminary analysis results

- Needs (i.e. compute power, collaborator)

- Reference paper from literature

- What input data required

- What software tool is to be used

39 proposals (1 individual, rest joint proposals) were received and reviewed by the OSD Paper Taskforce. Consequently all proposal senders were invited to join the OSD Analysis Consortium, and a mailing list (osd-analysis@microb3.eu) was set up to moderate

communication across the OSD Analysis Consortium. The Consortium currently has 132 members.

6.4 OSD Analysis Paper Outline and Organisation

A draft paper outline was prepared by the OSD Paper Taskforce and all proposals were grouped under three main categories with key questions per category. Further, every category was assigned Task Leaders. These were circulated among the OSD Analysis Consortium and received positive feedback.

List of categories along with Task Leaders are as follows:

1. Diversity (Using OTU-based metrics and alternatives such as MED, UniPept etc.)

Task Leaders: Linda Amaral-Zettler, Jed Fuhrman & Pier Luigi Buttigieg

2. Insights metabolic functions (with focus on human impact) and their role in the ecosystems from metaG

Task Leaders: Francesca Malfatti & Chris Bowler

3. Towards an understanding of broad-scale ecological patterns

Task Leaders: Francesca Malfatti & Gian Marco Luna

The draft OSD Analysis Paper Outline can be found here: https://goo.gl/pYBE71

All OSD Analysis Proposals can be found here: https://goo.gl/u43fMf

6.5 OSD GitHub Account

An OSD GitHub account was created and is continuously maintained by the OSD Bremen Team, who are in charge of OSD 2014 data release and its relevant documentation, all announced via this account. OSD GitHub Page also has an "Issues" section which serves as a forum/helpdesk to inform OSD Analysis Consortium of any updates/fixes as well as address their question and concerns.

OSD GitHub Page can be accessed via here: https://github.com/MicroB3-IS/osd-analysis

6.6 OSD Analysis Consortium Recent Progress & Next Steps

Based on the needs specified in the proposals, or where possible, proposal groups were either merged or matched together with each other or with the experts on the Consortium. Since almost all OSD 2014 has been publicly released by the OSD Bremen Team, OSD Analysis Consortium PIs were recently asked to report progress of their analyses briefly, giving them an opportunity to notify us of any potential issues or needs.

Next steps are as follows:

i. Investigation of Primers

In June 2014, a new study by Apprill et al. (Ref 1) was published (http://www.int-res.com/articles/ame_oa/a075p129.pdf) showing that a minor adjustment to the 806R primer would greatly increase the detection of the globally abundant SAR11 clade in marine and lake environments. These new findings are in line with an in-house study by Professor Jed Fuhrman and he pointed out that the primers to be used for the analysis of OSD 2014 samples (see Ref 2 for more details) would miss out on the detection of SAR11. We had looked into addressing this issue back then, together with Professor Fuhrman and Earth Microbiome Project Leaders Drs Jack Gilbert and Rob Knight, whom which we share the same primers, however there was no better primer available to be used in time for the OSD 2014 Event.

This time around, in order to address this issue, the company LGC was contacted, who are currently performing re-sequencing of five OSD and one RSD 2014 samples with different primer pairs. Professor Fuhrman also kindly shipped DNA samples from two different mock community to LGC to be included in the test runs. Details are provided below. Once the results are available, they will be cross-compared and measures will be taken accordingly.

Re-sequencing test-run being performed using the following combinations:

- B515F (OSD, Ref 2) and B806R (Ref 1)

- B515F (OSD, Ref 2) and 926R (Parada et al, submitted for publication)

- 515F-Y (Parada et al, submitted for publication) and B806R (Ref 1)

- 515F-Y (Parada et al, submitted for publication) and 926R (Parada et al, submitted for publication)

Primer sequences:

- B515F (OSD, Ref 2) - GTG CCA GCM GCC GCG GTA A

- 515F-Y (Parada et al, submitted for publication) - GTG YCA GCM GCC GCG GTA A

- B806R (Ref 1) - GGA CTA CNV GGG TWT CTA AT

- 926R (Parada et al, submitted for publication) - CCG YCA ATT YMT TTR AGT TT

Test samples:

- OSD3 Helgoland

- OSD14 Banyuls

- OSD69 Venice - Marghera

- OSD95 Singapore Indigo_V

- OSD133 Robben Island

- RSD5

Mock communities provided by Professor Fuhrman:

Tube 1 (Jed): staggered community, 0.001 ng/µL, 100 µL (in TE buffer)

Tube 2 (Jed): even community, 0.001 ng/µL, 100 µL (in TE buffer)

ii. Submission of OSD 2014 Environmental Data to PANGAEA

During final thorough curation of OSD 2014 environmental data, it was noticed that some of the OSD sampling stations do not have accurate coordinates for latitude and longitude. Their relevant OSD Site Coordinators were all contacted individually. These will soon be updated accordingly, submitted to PANGAEA and made publicly available.

iii. Gathering Ancillary Data from Relevant Public Initiatives

An online Ancillary Data Request Form (http://goo.gl/forms/pvZAz2dDrO) was prepared and circulated among the OSD Analysis Consortium to allow request of further ancillary data for the sampling sites under investigation, to aid with their analyses. We are currently in the progress of pulling the requested ancillary data from relevant public initiatives. This is being performed based on the latitude and longitude of all sampling sites and will soon be made available for the use of OSD Analysis Consortium.

iv. Announce Submission Deadline for Final Reports on OSD 2014 Data Analysis

OSD Analysis Consortium PIs have already been informed of the structure for their analysis reports. This is provided below. Once all above remaining tasks are finalised, OSD Paper Taskforce will get together and decide on a submission deadline to be announced.

- Proposal Number (Subject of email as well as within report)

- Title

- All contributors: Name, email & affiliation

- Summary: Aim & main findings

- Input data: Which OSD 2014 data did you use?

(i.e. 16S raw, 16S workable, MG raw, MG workable by EBI, MG workable by Bremen,..)

- Methodology: Software tools/Algorithms/Computing infrastructure

- Results including figures/graphs

- Conclusion

- Next steps along with % Completion (if any)

Proposals are to be emailed to osd-paper@microb3.eu in pdf format and not exceeding 5 pages.

v. Collate Reports & Write Paper

Upon receiving of all analysis reports, a telecon will be organised by the OSD Paper Taskforce to review them all in detail. This will potentially be followed by an in-person meeting where Task Leaders will get together to write the paper. Details to be arranged in due time.

References

Apprill A, McNally S, Parsons R, Weber L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology. 75: 129-137.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J.

Kopf A, Bicak M, Kottmann R, Schnetzer J, Kostadinov I, Lehmann K, Fernandez-Guerra A, Jeanthon C, Rahav E, Ullrich M, Wichels A, Gerdts G, Polymenakou P, Kotoulas G, Siam R, Abdallah RZ, Sonnenschein EC, Cariou T, O'Gara F, Jackson S, Orlic S, Steinke M, Busch J, Duarte B, Caçador, I, Canning-Clode J, Bobrova O, Marteinsson V, Reynisson E, Loureiro CM, Luna GM, Quero GM, Löscher CR, Kremp A, DeLorenzo ME, Øvreås L, Tolman J, LaRoche J, Penna A, Frischer M, Davis T, Barker K, Meyer CP, Ramos S, Magalhães C, Jude-Lemeilleur F, Aguirre-Macedo ML, Wang S, Poulton N, Jones S, Collin R, Fuhrman JA, Conan P, Alonso C, Stambler N, Goodwin K, Yakimov MM, Baltar F, Bodrossy L, Kamp JV, Frampton DMF, Ostrowski M, Ruth RV, Malthouse P, Claus S, Deneudt K, Mortelmans J, Pitois S, Wallom D, Salter I, Costa R, Schroeder DC, Kandil MM, Amaral V, Biancalana F, Santana R, Pedrotti ML, Yoshida T, Ogata H, Ingleton T, Munnik K, Rodriguez-Ezpeleta N, Berteaux-Lecellier V, Wecker P, Cancio I, Vaulot D, Bienhold C, Ghazal H, Chaouni B, Essayeh S, Ettamimi S, Zaid EH, Boukhatem N, Bouali A, Chahboune R, Barrijal S, Timinouni M, Otmani FE, Bennani M, Mea M, Todorova N, Karamfilov V, Hoopen P, Cochrane G, L'Haridon S, Bizsel KC, Vezzi A, Lauro FM, Martin P, Jensen RM, Hinks J, Gebbels S, Rosselli R, Pascale F, Schiavon R, Santos A, Villar E, Pesant S, Cataletto B, Malfatti F, Edirisinghe R, Silveira JAHS, Barbier M, Turk V, Tinta T, Fuller WJ, Salihoglu I, Serakinci N, Ergoren MC, Bresnan E, Iriberri J, Nyhus PAF, Bente E, Karlsen HE, Golyshin PN, Gasol JM, Moncheva S, Dzhembekova N, Johnson Z, Sinigalliano CD, Gidley ML, Zingone A, Danovaro R, Tsiamis G, Clark MS, Costa AC, Bour ME, Martins AM, Collins ER, Ducluzeau A, Martinez J, Costello MJ, Amaral-Zettler LA, Gilbert JA, Davies N, Field D and Glöckner FO. 2015. The ocean sampling day consortium. GigaScience. 4:27.

Hoopen P, Pesant S, Kottmann R, Kopf A, Bicak M, Claus S, Deneudt K, Borremans C, Thijsse P, Dekeyzer S, Schaap DMA, Bowler C, Glöckner FO and Cochrane G. 2015. Marine microbial biodiversity, bioinformatics and biotechnology (M2B3) data reporting and service standards. Standards in Genomic Sciences. 10:20.