Marine
robial Biodiversity,
ics & Biotechnology

Grant agreement n°287589

Acronym : Micro B3
Start date of project: 01/01/2012, funded for 48 month

Deliverable 4.6

Approved concepts and software transferred for long-term maintenance

Version: 1.0
Circulated to: Peter Thijsse, MARIS, Renzo Kottmann, MPI-BREMEN, Catherine Borremans, IFREMER, Stefanie Dekeyzer, VLIZ (20.12.2014)
Approved by: WP3-5

Expected Submission Date: 31.12.2014
Actual submission Date: 22.12.2014

Lead Party for Deliverable: EMBL-EBI (Petra ten Hoopen and Guy Cochrane)

Mail: petra@ebi.ac.uk Tel.: +44 1223 492565

Dissemination level:	
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)	

The Micro B3 project is funded from the European Union's Seventh Framework Programme (Joint Call OCEAN.2011-2: Marine microbial diversity – new insights into marine ecosystems functioning and its biotechnological potential) under the grant agreement no 287589. The Micro B3 project is solely responsible for this publication. It does not represent the opinion of the EU. The EU is not responsible for any use that might be made of data appearing herein.

Summary

The Micro B3 project as a platform for multidisciplinary community discussions led, among others, to development of best practice in collection and archiving of contextual data associated with marine microbial samples with focus on the Ocean Sampling Day campaign (OSD). The Micro B3 standards and interoperability structures were proven during the primary Micro B3 sampling campaign in June 2014, the OSD, and summarised in a manuscript undergoing currently a peer review process.

We have identified two suitable communities that offer registration of data standardisation efforts, such as the Micro B3 consortium standards and interoperability structures, at their portal, which allows experts from the community and beyond to be aware of and benefit from the harmonisation efforts. A proposal has been formulated to (1) the IODE-OSDBP project designed for registration of standards and best practices related to ocean data management and exchange, and (2) the BioSharing project designed for registration of standards and guidelines in life sciences.

Furthermore, a proposal to the Genomic Standards Consortium (GSC) as the long-term guardian of the Micro B3 contextual data reporting standard has been formalised. A strong focus of the GSC-developed MIxS molecular data standards on microbial environments and Micro B3 emphases on nucleotide sequence analysis of the OSD samples entitles GSC to the long-term guardian role.

Table of contents

1. Objectives of the Deliverable 4.6

The legacy of the Micro B3 project reaches many different areas, including advances in bioinformatics infrastructures, a unique microbial sample collection, a global network of marine stations, increased interoperability across multidisciplinary data repositories and established best practice for contextual data collection of marine microbial samples. Exposure of the last two of these to relevant expert communities is the objective of this Deliverable 4.6, which builds on results described in the previous deliverables.

Deliverable 4.3 summarises two years of effort resulting in the best practice guide for the Ocean Sampling Day initiative (OSD, http://www.oceansamplingday.org). It specifies contextual data reporting requirements for OSD sampling stations, which are compatible with data reporting standards of three Micro B3-connected domains, genomic, biodiversity and oceanographic. The OSD Handbook is a concise guide advising OSD marine laboratories on steps they need to take before, during and after the sampling in order to ensure harmonised sample and data collection.

Intensive discussions among data management centres from the three above-mentioned domains resulted in an agreement on interoperability minimum and interoperability protocols, both summarised in the Deliverable 4.4. The former is essentially a minimal set of descriptors with specified formats that each interoperable resource will support in its discovery and delivery services. The latter comprises programmatic service protocols, which each interoperable resource will implement for serving their data.

In order to deliver harmonised and compliant datasets from Micro B3 data providers (OSD marine stations) to the Micro B3 data recipients (primary and secondary archives of molecular, oceanographic and biodiversity data), software tools were adapted and developed supporting the digitisation of collected data and use of the developed standards. Deliverable 4.5 describes these tools available from the Micro B3 Information System and European Nucleotide Archive.

Micro B3 harmonisation of data acquisition and archiving is the main responsibility of work package 4 but a result of an intensive cross-work package collaboration. Experts in Micro B3-related communities can benefit from the Micro B3 developed standards and interoperability structures only if they are aware of them. Deliverable 4.6 addresses this aspect by identifying community portals appropriate for registration of the Micro B3 work package 4 outputs. Chapter 2 contains a registration proposal to two community portals: (1) a proposal to the IODE-OSDBP project designed for registration of standards and best practices related to ocean data management and exchange, and (2) a proposal to the BioSharing project designed for registration of standards and guidelines in life sciences.

Micro B3 work package 4 outputs will also be handed over to a long-term guardian. Considering the microbial nature of the OSD samples and focus on molecular analysis of these samples, the most obvious long-term guardian is the Genomic Standards Consortium (GSC, http://gensc.org/). Chapter 3 of the Deliverable 4.6 describes the long-term guardian proposal to the GSC.

2. Community portals registration

Partners of the Micro B3 work package 3, 4 and 5 met in October 2014 for a follow up interoperability meeting reviewing their progress on implementation of interoperability structures. Experts from three scientific domains: (1) oceanographic (represented by MARIS and MARUM, (2) biodiversity (represented by VLIZ) and (3) molecular (represented by EMBL-EBI and MPIMM) discussed also potential guardians of the Micro B3 standardisation assets.
Most standardisation efforts are led by communities with a long-term approval process and a tracker system of a standard maturity level, such as the Biodiversity Information Standards (TDWG, http://www.tdwg.org/). The Micro B3 project served as a platform for multidisciplinary community discussions. The Micro B3 standards and interoperability structures were approved by the Micro B3 consortium in the course of the project and tested during the OSD in June 2014.
Therefore, we seek community portals offering solely a standard effort registration that enables extended awareness of the standard within the community and beyond.
From a number of potential candidates two portals described in the following chapters were selected for registration.

2.1. IODE-ODSBP proposal

This chapter describes in full a proposal to the Ocean Data Standards and Best Practices Project (ODSBP) of the International Oceanographic Data and Information Exchange (IODE) programme. The proposal aims to register the multidisciplinary Micro B3 data reporting and service standards under a new name 'Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards'. The IODE-ODSBP proposal follows the recommended proposal template available at the link: http://www.oceandatastandards.org/proposal-template-mainmenu-46.

---------IODE-ODSBP proposal ----------start ----------

Title

Titles should provide an indication of the subject, scope and purpose of the Proposal (for example: Adoption of ISO 3166-1 and 3166-3: Country Codes)

Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards

Scope

Provide a clear indication of the extent of the Proposal ™s application. Identify any specific data communities (based on data management functions, types of data, observations, etc.) to which this Proposal could apply. Indicate any known limitations or exclusions where the Proposal is not adequate.

Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) standards consist of data Reporting and Service standards.

The M2B3 Reporting Standard describes contextual information recommendation for marine microbial samples relevant to researchers in the oceanographic, biodiversity and molecular disciplines and adoptable by any marine laboratory with minimum sampling resources.

The M2B3 Service Standard defines bioinformatics means through which these data can be discovered and explored in data repositories.

Envisaged publication type

Briefly describe the target audience for this Proposal and any outreach plans to be considered.

The M2B3 is targeted to a community of marine laboratories with biological sampling program and data management centers dealing with marine data.

The M2B3 Reporting Standard forms guidelines to marine laboratories performing sampling for biological samples with focus on marine microbial biodiversity.

The M2B3 Service Standard specifies service protocols and identifies a minimum set of descriptors essential to interoperability across disciplines that a compliant data resource shall support.

A manuscript with detailed specification of the M2B3 standard has been submitted to the Standards in Genomic Sciences (http://www.standardsingenomics.com/) peer review journal.

Purpose and Justification:

1. *Describe the specific aims and reason for this Proposal, with particular emphasis on the aspects of standardization covered, the problems it is expected to solve or the difficulties it is intended to overcome.*

In recent years large-scale initiatives were formed to address the wealth of genetic, functional and morphological diversity in marine ecosystems. Such planetary-scale surveys originate from realization that only a holistic approach and powerful data infrastructure can extend our understanding of the composition and ecology of marine microbial ecosystems. Prerequisite

for the successful exploitation are standards that enable interoperability in the data infrastructure. Biological sampling is an expensive effort that requires expertise as well as material resources. Collection of contextual data at the time of sample collection is critical in the field of marine biodiversity. The M2B3 Reporting Standard addresses these issues by harmonising contextual data reporting associated with sampling for microbial organisms in a global network of marine stations. This harmonisation allows a multiple use of a single sample, meaningful comparisons between samples as well as reproducible molecular data analysis of collected samples. The M2B3 Service Standard specifies service protocols and a minimal set of core interoperability descriptors, which are essential to integration of data from independent resources and allow connection between diverse data across molecular, biodiversity and oceanographic domains.

2. *Describe how this proposed standard supports data management, exchange or interoperability. When applicable include mention of what data management functions (e.g. date transport, quality control, archive, ¦) the proposal supports.*

The M2B3 Reporting standard spans the biodiversity, molecular and oceanographic domains and adopts existing standards of each discipline. It represents a unique intersection of existing reporting requirements across all three domains. The Standard specifies descriptor usage, requirements for formats and appropriate units, where applicable.
Compliance of a dataset to the M2B2 Reporting Standards allows data mangers create a molecular data record compliant with the MIxS molecular data standard, Version 4 [3], an oceanographic data record compliant with the oceanographic Common Data Index (CDI), Version 3 [4] and a biodiversity data record compliant with the biodiversity OBIS Schema, Version 1.1 [5].

The M2B3 Service Standard specified descriptors from the M2B3 Reporting Standard, which are essential to data interoperability across the three mentioned disciplines. These descriptors allow connections from data points of one discipline to data points in another. The six interoperability descriptors are presented by a compliant data resource using a programmatic service interface that follows Open Geospatial Consortium (OGC) standards, the Web Map Service (WMS), Web Feature Service (WFS) and/or the OpenSearch protocol.

3. *Describe the main interests benefitting from or affected by the proposed standard, such as industry, consumers, governments, distributors. Identify any relationships and/or dependencies.*

The M2B3 Reporting Standard is primarily relevant to the marine microbial community. The M2B3 Service Standards is of primary importance to data distributors, i.e. data centers that collect, archive and present information of their respective scientific domain, wishing to make their data interoperable and available beyond their discipline.

4. *Describe the feasibility of implementing the proposed standard. Include any factors that could hinder the successful establishment or global application of the Proposed standard. Are there any associated issues? Identify resource implications resulting from the recommendations.*

The M2B3 Reporting Standard is directed to marine scientists who should be able to report this minimum contextual information about each marine microbial sample irrespective of their sampling resources.
The six interoperability descriptors supported by the M2B3 Service Standard form a subset of minimum reporting requirements applied in each respective domain and are therefore available in data resources of all three disciplines. The proposed support for discovery and delivery of data using the OpenSearch query and result service, OGS WMS and/or OGS WFS service is based on their use by search engines.

5. *Considering the needs of other fields or organizations, indicate the timeliness, target date(s), or if proposing a series of standards, suggest priorities. List any statutory requirement or other driving factors.*

The M2B3 Standards resulted from discussions of a multidisciplinary community of stakeholders (research and data management centers) where working practices established in each particular area of expertise were revised in order to find a common consensus upon which interoperability across disciplines could be built. The proposal is tailored to needs of the oceanographic/biodiversity/genomic communities with a focus on marine microbial biological sample collection, archiving and discovery. However, a relatively small revision would be needed to apply the M2B3 to the full range of pelagic and benthic organism. The M2B3 can also serve as a solid base for other standardisation efforts within each domain. For instance, the M2B3 serves a basis for standardisation of all marine molecular data available in the public domain.

6. Describe the possible benefits gained by the implementation of the proposed standard. Alternatively, describe the loss or disadvantage(s) if no standard is established within a reasonable time.

Compliance to the M2B3 Reporting Standard allows a marine laboratory to join a global network of marine stations, where collected samples can be compared irrespective of the sampling locality and meaningful analysis of all samples as a single dataset performed due to harmonization of their contextual data. The M2B3 Service Standard addresses the need to build an interoperable infrastructure across scientific disciplines allowing multidisciplinary questions to be asked and data across disciplines discovered and integrated.

7. Indicate whether the proposed standard is or may become the subject of regulations or may require the harmonization of existing regulations. Describe any impacts of this activity.

The M2B3 Standards development was based on the pragmatic approach of supporting data flow across disciplines using extensively established standardization methodologies in each respective domain. The M2B3 fully respects harmonization efforts in each scientific domain and does not impose alteration of existing established practice in either domain. It's value lies in the identification and definition of data reporting and interoperability minimum allowing harmonized contextual data collection and sharing across disciplines.

***Current Operational Implementations**: Provide information about organizations, programs or projects which currently use the Proposed standard as part of an operational environment. If there are none, please indicate organizations that are testing the standard.*

The M2B3 Data Standards were developed by the European project Micro B3 [6], funded under 7th Framework Programme "Ocean of Tomorrow".

The M2B3 Reporting Standard was first used with the Ocean Sampling Day initiative [7]. Over 150 marine stations around the world sampled simultaneously for marine microbes according to standardized sampling protocols and collected sample contextual data using the M2B3 Reporting Standard. This network of marine stations, one of the largest networks ever made, generated a unique collection of samples where one set of replicates is archived for future use as technologies advances and another set currently undergoes molecular data analysis.

To date, the European Nucleotide Archive [8], European Ocean Biogeographic Information System (EurOBIS) [9], Micro B3 Information System [10], PANGAEA [11] and SeaDataNet [12] have committed to supporting the M2B3 Service Standard for OSD data.

Relevant Documents:

1. Provide the reference(s) to all documents or materials associated with this Proposal (e.g. standards, specifications, regulations). Where Proposals comprise multiple documents or files, include a brief description of the relevancy of each as well as any dependencies among these materials.

The M2B3 Reporting Standard is described in the Micro B3 project deliverable D4.3 [1].

The M2B3 Service Standard is described in the Micro B3 project deliverable D4.4 [2].

All relevant documents and data resource mentioned in this proposal are listed below.

[1] **Micro B3 – D4.3**
[http://www.microb3.eu/sites/default/files/deliverables/MB3_D4_3_PU.pdf]

[2] **Micro B3 – D4.4**
[http://www.microb3.eu/sites/default/files/deliverables/MB3_D4_4_PU.pdf]

[3] **MIxS, v.4.0** [http://wiki.gensc.org/index.php?title=MIxS]

[4] **CDI, v.3.0** [http://www.seadatanet.org/Data-Access/Common-Data-Index-CDI]

[5] **OBIS, v.1.1** [http://www.iobis.org/node/304]

[6] **Micro B3** [http://www.microb3.eu/]

[7] **OSD** [http://www.oceansamplingday.org]

[8] **ENA** [http://www.ebi.ac.uk/ena/home]

[9] **EurOBIS** [http://www.eurobis.org/]

[10] **Micro B3 IS** [http://mb3is.megx.net/]

[11] **PANGAEA** [http://www.pangaea.de/]

[12] **SeaDataNet** [http://www.seadatanet.org/]

[13] Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G *et al.* **Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications.** Nature Biotechnology 2011; 29: 415–420.

[14]] **MEDIN** [http://www.oceannet.org/marine_data_standards/]

2. Attach copies of all relevant documents or materials to this proposal. Where copyright policies restrict the attachment of the documents, indicate these by providing a listing along with the resource through which these documents may be obtained.

Cooperation and liaison:

1. ***Existing Community***: *List relevant organizations, bodies, work groups or related projects which currently use the Proposed standard and through which cooperation and liaison could be extended to the broader community. Include organizations, programs, etc. supporting the submission of this proposal.*

SeaDataNet – Pan-European infrastructure for ocean and marine data management
EurOBIS – European Node of the international Ocean Biogeographic Information System
ENA – European Nucleotide Achive
PANGAEA – Data Publisher for Earth and Environmental Science
Micro B3 IS – Micro B3 Information System

2. ***Expanded Community:*** *List relevant organizations, bodies, work groups or related projects not currently employing the Proposed standard and with which cooperation and liaison should exist.*

GSC – Genomic Standards Consortium [13]
MEDIN – Marine Environmental Data and Information Network [14]

Contact information: *Provide the contact information of the Proponent. This individual acts as the key point of contact for interaction with the ETDMP Chair on this proposal. Include the Proponents name, organization, email address, and telephone number.*

Name: Dr. Petra ten Hoopen
Organization: European Nucleotide Archive, The EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom
Email Address: petra@ebi.ac.uk
Telephone Number: *+44(0)1223 492565*

Proposal version: v.1, September 2014

List of Acronyms:
CDI – Common Data Index
ENA – European Nucleotide Archive
EurOBIS – European Ocean Biogeographic Information System
M2B3 – Marine Microbial Biodiversity, Bioinformatics and Biotechnology
Micro B3 – Marine Microbial Biodiversity, Bioinformatics, Biotechnology
Micro B3 IS – Micro B3 Information System
OGC – Open Geospatial Consortium
OSD – Ocean Sampling Day
PANGAEA – Data Publisher for Earth and Environmental Science
WFS – Web Feature Service
WMS – Web Map Service

---------IODE-ODSBP proposal ----------end ----------

2.2. BioSharing proposal

This chapter describes in full a registration record to the BioSharing portal for community-developed standards in the life sciences, broadly covering biological, biomedical and natural sciences. The record aims to register the multidisciplinary Micro B3 data reporting and service standards under a new name 'Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards'. The BioSharing record follows the recommended record template available at the link: http://www.biosharing.org/new/biosharing.

---------BioSharing record ----------start ----------

Record Type:
Reporting guidelines

Umbrella Foundry:
None

Short Name:
M2B3

Full Name:
Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards

Description:
The M2B3 consists of a Data Reporting Standard and Data Service Standard. The M2B3 Data Reporting Standard harmonises contextual data reporting associated with sampling for microbial organisms in a global network of marine stations. The M2B3 Reporting standard spans the biodiversity, molecular and oceanographic domains and adopts existing standards of each discipline. It represents a unique intersection of existing reporting requirements across all three domains. The Standard specifies descriptor usage, formats and indication of appropriate units, where applicable.
The M2B3 Service Standard specifies service protocols and a minimal set of core interoperability descriptors, which are central to integration of data from independent resources and allow connection between diverse data across molecular, biodiversity and oceanographic domains.

Homepage:
http://www.ebi.ac.uk/ena/submit/microb3-checklist

Contact Name:
Petra ten Hoopen

Contact ORCID:
http://orcid.org/0000-0003-4242-4015

Contact Email:
petra@ebi.ac.uk

Countries:
United Kingdom, Germany, Netherlands, Belgium,

Domains:
Biodiversity, Oceanographic, Genomic

Taxonomy:
NCBI Taxonomy (http://www.ncbi.nlm.nih.gov/taxonomy/)
WoRMS (http://www.marinespecies.org/about.php)

Related BioDBCore Entries:
-

Related Standards:
Minimum information about any (x) sequence, Version 4.0
(http://gensc.org/projects/mixs-gsc-project/)

Ocean Biogeographic Information System schema, Version 1.1,

(http://www.iobis.org/data/schema-and-metadata)

Common Data Index, Version 3.0
(http://www.seadatanet.org/Data-Access/Common-Data-Index-CDI)

Associated Organisations:
EMBL-EBI (http://www.ebi.ac.uk/)
IBENS (http://www.biologie.ens.fr/depbio/?lang=eng)
ICES (http://www.ices.dk/Pages/default.aspx)
IFREMER (http://wwz.ifremer.fr/institut_eng/)
MARIS [(http://www.maris.nl/)
MARUM (http://www.marum.de/en/)
MPIMM (http://www.mpi-bremen.de/Institut.html)
University of Oxford (http://www.ox.ac.uk/)
VLIZ (http://www.vliz.be/en/organisation)

Publications:
A manuscript with detailed specification of the M2B3 standard has been submitted to the Standards in Genomic Sciences (http://www.standardsingenomics.com/) peer review journal.

Support:
Online documentation,
URL: http://www.microb3.eu/sites/default/files/osd/OSD_Handbook_v2.0.pdf

---------BioSharing record ----------end ----------

3. GSC as the long-term guardian

This chapter describes in full a proposal to the Genomic Standards Consortium (GSC) as the long-term guardian of the Micro B3 standards and interoperability results. The proposal aims to introduce the multidisciplinary Micro B3 data reporting and service standards under a new name 'Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards'. The GSC proposal follows the recommended project description template available at the link: http://gensc.org/gsc-project-description-template/.

---------GSC proposal ----------start ----------

Project Title

The Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data reporting and service standards

Project Lead (s) (name, email, institute(s))

Guy Cochrane, cochrane@ebi.ac.uk, EMBL-EBI
Petra ten Hoopen, petra@ebi.ac.uk, EMBL-EBI

Team members (name, email, institute(s)) (those introducing and leading the project in the initial phases)

Guy Cochrane, cochrane@ebi.ac.uk, EMBL-EBI
Petra ten Hoopen, petra@ebi.ac.uk, EMBL-EBI
Frank Oliver Glöckner, fog@mpi-bremen.de, MPI
Renzo Kottmann, rkottman@mpi-bremen.de, MPI
Anna Klindworth, aklindwo@mpi-bremen.de, MPI
Dawn Field, fiedawn@googlemail.com, Univ. Oxford
Mesude Bicak, mesude.bicak@oerc.ox.ac.uk, Univ. Oxford
Stéphane Pesant, spesant@marum.de, MARUM
Chris Bowler, cbowler@biologie.ens.fr, IBENS
Simon Claus, simon.claus@vliz.be, VLIZ
Klaas Deneudt, klaas.deneudt@vliz.be, VLIZ
Stefanie Dekeyzer, stefanie.dekeyzer@vliz.be, VLIZ
Dick M.A. Schaap, dick@maris.nl, MARIS
Peter Thijsse, peter@maris.nl, MARIS
Catherine Borremans, Catherine.Borremans@ifremer.fr, IFREMER
Neil Holdsworth, NeilH@ices.dk, ICES

Elevator pitch

The Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) data Reporting Standard describes contextual information recommendation for marine microbial samples relevant to researchers in the oceanographic, biodiversity and molecular disciplines and adoptable by any marine laboratory with minimum sampling resources.

Project Summary (two or three paragraphs of background, purpose and plans)

The *M2B3 Reporting and Service Standard* was developed by the European project Micro B3 [8], funded under 7th Framework Programme "Ocean of Tomorrow" as a collaboration of experts from biodiversity, genomic and oceanographic domain, representing microbial biodiversity and function in the marine environment. This proposal concerns the *M2B3 data Reporting Standard*.

Development of *the M2B3 Reporting Standard* was based on a combination of a hypothesis-driven approach and current best-practice driven approach. The theoretical approach involved analysis of use case studies in the field of marine microbiology. The practical approach involved detailed analysis of current sampling practice by consulting experts in marine microbial sampling and research from leading marine Institutes in Europe and USA.

Further insight into the current best practice in data capture was provided by major European data management and research centers in each domain, for biodiversity data the VLIZ, [15], IBENS [3] and University of Oxford [14], for oceanographic data the MARUM [7], MARIS [6], IFREMER [5] and ICES [4], and for molecular data the ENA at EMBL-EBI [2] and MPI [11].

The *M2B3 Reporting Standard* sums minimum reporting information about microbial sampling in the epipelagic zone for marine stations and cruises participating in the Ocean Sampling Day campaign, the principal sampling campaign of the Micro B3 project.

The *M2B3 Reporting Standard* represents minimum information about marine microbial sample and is compliant to the MIxS core standards, Version 4.0, [12]. Moreover, data reporting information of this standard allows create records compliant with the biodiversity OBIS schema, Version 1.1, [13] and oceanographic Common Data Index metadata standard, Version 3 [1].

The M2B2 serves as a basis for standardisation of all marine molecular data available in the public domain. For example, contextual data of molecular samples generated during the Tara Oceans Expeditions are compliant to *the M2B3 Reporting Standard* and form one of the biggest, most consistent and richly described dataset in the public domain. Other marine communities already expressed interest to the European Nucleotide Archive to use this standard for description of their marine molecular samples.

Since assets of the Micro B3 standardization can serve as an inspiration for other joint marine enterprises and other minimum information efforts, the Micro B3 consortium seeks a handover of *the M2B3 Reporting Standard* to a long-term guardian.

This proposal indicates how the M2B3 fits into the mission of the Genomic Standards Consortium (GSC) and what contribution can the M2B3 project bring.

Project initiation date (can be a date or a narrative about how and when the project started and how it is proceeding)

01.01.2012

What will this project aim to contribute to the GSC?

The M2B3 is an asset and legacy of the GSC-registered Micro B3 project. It represents a unique intersection of existing standards in the oceanographic, molecular and biodiversity domains.

Have you spoken about the project already within GSC? (on a call, at a formal GSC meeting, would like to request time to present at a future meeting).

Handover of Micro B3-developed standards to the Genomic Standards Consortium as the long-term guardian has been proposed in the Micro B3 project, which is coordinated by the GSC member Frank Oliver Gloeckner and its work package 2, 4 and 5 led by GSC members Dawn Field (former), Guy Cochrane and Renzo Kottmann, respectively.

Which existing projects, if any, does this one

replace/complement/subsume/expand? Explain briefly why an extra project is needed/justified (what gap does it fill?)

The M2B3 specifies contextual information of a microbial sample in the marine environment and was developed in collaboration of oceanographic, molecular and marine biodiversity experts.
Its multidisciplinary nature allows interoperability of data and compliance of generated records in all three scientific domains.

How does this project fit into GSC's mission statement (might also expand it)?

The M2B3 describes in a consistent manner a marine research sampling site and contextual information around collected marine microbial samples.

Will you start a GSC working group (how far along are you?)? If not, why not (i.e. subgroup within developers group, existing external community, etc)

There is no need for a GSC Working Group since the Micro B3 consortium served as a platform for multidisciplinary community discussions.

How do you wish to further engage the GSC (recruit members to project, get consultation, link to other GSC projects, etc)?

The M2B3 is a sub-project of the GSC-registered Micro B3 project.

Do you already have a website or do you wish to create a home page for the project in the GSC website (GSC maintains an open wiki at present, all working groups have a page)?

Currently, the M2B3 is available from the Micro B3 project website in the OSD Handbook, Version 2 [11] and on external pages of the ENA project [2].

What other resources might you like from what the GSC can offer (mailing lists, etc)?

The M2B3 project leads would appreciate a summary of resources that GSC can offer to minimum information efforts, such as the M2B3 project, in order to evaluate and maximise mutual benefit arising from GSC as the long-term guardian of the M2B3.

What kind of timeline are you working to for building consensus, releasing a first version etc.

In June 2014 the second version of the Ocean Sampling Day Handbook has been released describing the final version the Micro B3 Checklist, which forms the content of the M2B3.

How is this work currently funded (list grants, funders, in kind contributions, etc)?

Funding of the project has been granted and finishes in December 2015.

What are your current plans for publishing/promoting the project?

A manuscript with detailed specification of the M2B3 standard has been submitted to the Standards in Genomic Sciences (http://www.standardsingenomics.com/) peer review journal.

References or relevant websites (for further reading)

[1] **CDI, v.3.0** [http://www.seadatanet.org/Data-Access/Common-Data-Index-CDI]
[2] **ENA at EMBL-EBI** [http://www.ebi.ac.uk/ena/home]
[3] **IBENS** [http://www.biologie.ens.fr/depbio/?lang=eng]
[4] **ICES** [http://www.ices.dk/Pages/default.aspx]
[5] **IFREMER** [http://wwz.ifremer.fr/institut_eng/]
[6] **MARIS** [http://www.maris.nl/]
[7] **MARUM** [http://www.marum.de/en/]

[8] **Micro B3** [http://www.microb3.eu/]
[9] **Micro B3 – WP4** [http://www.microb3.eu/work-packages/wp4]
[10] **Micro B3 – OSD Handbook**
[http://www.microb3.eu/sites/default/files/osd/OSD_Handbook_v2.0.pdf]
[11] **MPI** [http://www.mpi-bremen.de/Institut.html]
[12] **MIxS, v.4.0** [http://wiki.gensc.org/index.php?title=MIxS]
[13] **OBIS, v.1.1** [http://www.iobis.org/node/304]
[14] **University of Oxford** [http://www.ox.ac.uk/]
[15] **VLIZ** [http://www.vliz.be/en/organisation]

---------GSC proposal ----------end ----------