**MoBE Metadata & Standards**

**What is metadata?**

Metadata is ‘data’ about data. In practical terms, metadata is the information describing a sampling event and subsequent sequencing efforts. For studies of the Microbiome of the Built Environment, metadata describes the study site (type of building) and sampling data (building and room descriptors where the sampling occurred).

**Why implement metadata standards?**

Utilizing metadata standards to annotate the data describing the sample, sampling environment and sequencing methodology will vastly improve our ability to mine and integrate the built environment sequence data collection for knowledge and application driven research. Collection and reporting of a common, minimal set of metadata across the Microbiome of the Built Environment projects will foster data comparisons and analysis. Combining studies in a standard way will allow for more powerful analyses of Built Environment data.

**MoBE Metadata Standard Implementation Guide**

Study and sample metadata may be collected and organized into tab-delimited metadata files that are MIxS/MIMARKS compliant for submission to QIIME, EBI or NCBI’s BioSample repository.

**Metadata Standard**

The MIxS metadata standard is a checklist of metadata terms utilized to annotate sample records. MIxS is a unified standard developed by the Genomic Standards Consortium (GSC) for reporting of minimum information about any (x) nucleotide sequence. It consists of MIGS, MIMS and MIMARKS standards and describes fourteen environments.

Fifteen environmental checklists, developed by environment-specific community experts, provide extensive range of environmental and epidemiological contextual data fields for accurate and consistent description of the sequenced sample, its environment and sequencing experiments performed with the sample.

GSC MIxS Project: <http://gensc.org/projects/mixs-gsc-project/>

MIxS checklists (EBI): <http://www.ebi.ac.uk/ena/about/mixs-checklist>

**Built Environment Metadata Standard (MIxS-BE)**

The MIxS-BE checklist represents a minimal metadata description of the built environment to be collected and reported for each sequenced sample.

MIxS-BE includes MIxS core fields (e.g. sample name, geographic location, collection date, sample material), BE core fields (e.g. surface material, surface humidity, air temperature), BE building properties (e.g. indoor surface, filter type) and BE sample properties (e.g. occupancy at sampling).

**Prepare and load metadata files**

**Templates:** [available under the microbe.net Resources menu]

Template files can be utilized to collect and organize study and sample metadata. The templates (.xls, .xlxs, or .txt) include specifications on allowable terms for each metadata field and term definitions for MIxS (MIMARKS) compliant metadata.

**Available Templates:**

1. QIIME MoBE prep\_template
2. QIIME MoBE sample\_template
3. MIxS template
4. MIxS-BE template

**QIIME**

QIIME tutorials: <http://qiime.org/documentation/index.html>

QIIME Metadata tutorial:

[http://www.microbio.me/qiime/docs/tutorials/tutorial.html#](http://www.microbio.me/qiime/docs/tutorials/tutorial.html) provides step-by-step instructions.

QIIME documentation: <http://qiime.org/documentation/index.html>

QIIME Metadata Mapping Files: <http://qiime.org/documentation/file_formats.html>

Steps:

1. Log into QIIME: <http://www.microbio.me/qiime/>
2. Create a New study: from the QIIME Home Page, select ‘Create a New Study’. Fill in Study information and click on “Create”. Study metadata fields include: Study Name, Study Title, PubMed ID, Investigation Type and Environmental Package. Select package type “miscellaneous natural or artificial environment”.
3. Create MIMARKS-Compliant Templates: select metadata fields, select “Generate Template”, download the generated templates, fill in sample metadata to the sample template, and sequencing information to the prep\_template and save (.txt file) and compress the metadata file (.zip).
4. Upload the metadata and study (.zip) file. Send the .zip file to QIIME (Gail Ackermann: <mailto:gail.ackermann@colorado.edu>)

**QIIME Metadata Loader:**

The QIIME loader can be accessed through the database websites ([www.microbio.me/emp](http://www.microbio.me/emp) and [www.microbio.me/qiime](http://www.microbio.me/qiime)).

The QIIME loader currently supports Firefox. If you experience any difficulties contact Gail Ackermann (<mailto:gail.ackermann@colorado.edu>).

Studies uploaded to QIIME can be submitted to EBI if requested.

All MoBE studies have a link to VAMPS.

**NCBI BioSample Metadata Submissions**

Following registration of a NCBI BioProject, metadata can be submitted through the NCBI BioSample portal (<https://submit.ncbi.nlm.nih.gov/subs/biosample/>) with data formatted in the provided template format (**as .txt files**). The BioSample loader will not accept excel files.

**EBI European Nucleotide Archive SRA Submissions**

<http://www.ebi.ac.uk/ena/about/environmental_sequencing_submissions>

View and select the MIxS checklists and their environmental packages.