Title

MIxS-BE: a MIxS extension defining a minimum information standard for sequence data from the built environment.

Permalink
https://escholarship.org/uc/item/6f7348rr

Journal

The ISME journal, 8(1)

ISSN
1751-7362

Authors

Glass, Elizabeth M
Dribinsky, Yekaterina
Yilmaz, Pelin
et al.

Publication Date

2014

DOI
10.1038/ismej.2013.176

The ISME Journal (2014) 8, 1–3
© 2014 International Society for Microbial Ecology All rights reserved 1751-7362/14
www.nature.com/ismej

COMMENTARY

MIxS-BE: a MIxS extension defining a minimum information standard for sequence data from the built environment

Elizabeth M Glass[1], Yekaterina Dribinsky, Pelin Yilmaz, Hal Levin, Robert Van Pelt, Doug Wendel, Andreas Wilke, Jonathan A Eisen, Sue Huse, Anna Shipanova, Mitch Sogin, Jason Stajich, Rob Knight, Folker Meyer and Lynn M Schriml[1]

The ISME Journal (2014) **8,** 1–3; doi:10.1038/ismej.2013.176; published online 24 October 2013

The need for metadata standards for microbe sampling in the built environment

The composition of indoor microbial communities has the potential to profoundly affect human health. A number of factors within a building or room can alter the microbial abundance and diversity, such as occupancy, temperature and humidity, which in turn impacts indoor air quality. Researchers Hospodsky *et al.* (2012); Kembel *et al.* (2012) and Dunn *et al.* (2013) are exploring the intersection of microbial ecology, building materials and architectural design to understand microbial diversity and abundance within a building. Metadata (data describing data) provides an essential complement to experimental data, helping to answer questions about its source, mode of collection and reliability. As the impact and prevalence of large-scale metagenomic surveys grow, so does the need for more complete and standards compliant metadata. While marine (Gilbert *et al.*, 2012), soil and the human microbiome environments have representation in the standards being developed the built environment (BE) represents a unique context in need of standards development for the use in the study of microbial sequences. Metadata collection and interpretation have become vital to the genomics and metagenomics community to share information and integrate data across resources and within data repositories. The Genomic Standards Consortium (GSC, http://gensc.org; Field *et al.* (2011)) has developed widely accepted metadata MIxS (MIGS, MIMS, MIMARKS; Yilmaz *et al.* (2011)) standards for genomic, metagenomic and amplicon (for example, 16S rRNA) sequence data sets. These standards have been developed within a framework that is both modular and extensible. The MIxS-BE, as a minimal metadata standard, represents a unique extension to the GSC's MIxS standard as a rigorous and structured tool for the analysis of microbial sequences and ecosystems of the indoor environment. The MIxS-BE package provides the BE community with a suggested list of parameters to record and report for each sequenced sample and to compare data across studies. The MIxS-BE standard incorporates the core set of required MIxS fields for a bacterial sequence along with a Built Environment (BE) package (http://gensc.org/index.php?title=MIxS_extensions) comprised of a BE core, MIxS air environmental package terms, BE building properties and BE sample properties. The MIxS-BE package has been incorporated into the MIxS checklist and integrated into the QIIME (Caporaso *et al.*, 2010), MoBeDAC and MG-RAST (Meyer *et al.*, 2008) databases to foster metadata submission compliance across BE projects. Updates to the MIxS specification will be included in BioSample (Barrett *et al.*, 2012) as part of the annual MIxS release. Requests for additions or changes to the MIxS checklists can be directed to the GSC Developers mailing list: developers@gensc.org or by contacting Lynn Schriml or Elizabeth Glass.

Development of the MIxS-BE package has been an open and iterative process engaging the GSC community, the GSC's MIxS developers, stakeholders across the BE community including microbial ecologists, microbiologists, architects and engineers. The Microbiology of the Built Environment Alfred P. Sloan funded initiative established a metadata working group as part of the MoBeDAC, bringing together Sloan-funded researchers, architects, civil engineers, bioinformaticists and computational biologists to discuss the need and context for standards to describe the most relevant metadata to be collected for BE samples. Led by Elizabeth Glass and Lynn Schriml, the working group proposed the development of the BE-MIxS package to the GSC and gained the GSC's board approval to develop the standard. The working group initially identified a comprehensive list of metadata terms reported in BE literature. Based on feedback solicited from industry experts and microbial ecologists, the working group established a minimal set of metadata terms for the BE package environment samples with a subset of the MIxS-BE terms classified as M (Mandatory),

Table 1 MIxS-BE metadata package terms

MIxS-BE term
Carbon dioxide[a]
Ventilation type[a]
Organism count[a]
BE core
Surface material
Surface-air contaminant
Relative air humidity
Absolute air humidity
Surface humidity
Air temperature
Surface temperature
Surface moisture pH
Surface moisture
Dew point
Building occupancy type
BE Building properties
Indoor space (room type)
Indoor surface
Filter type
Heating and cooling system type
Substructure type
Building setting
Light type
BE Sample properties
Sample size sorting method
Space typical state
Typical occupant density
Occupancy at sampling
Occupant density at sampling

Terms, definitions, required or optional status (for reporting compliance) and syntax are presented at: MIxS project, (http://gensc.org/gc_wiki/images/7/70/Built_environment-metadata-terms-v51.xls).
[a]The MIxS-BE package includes MIxS-air environmental package terms and the built environment core, building and sample property terms.

which indicates that the term has to be reported for the metadata to be considered compliant to the MIxS checklist. The MIxS-BE minimal set will be complemented by a MIxS-BE-Building package (under development) describing the larger set of building and room metadata pertinent to describing a BE sample (Table 1).

Acknowledgements

We would like to thank the Alfred P Sloan Foundation grant FP047325-01-PR for support for this project. We also thank the members of the BE community and the GSC for their comments on this minimal standard, especially Jessica Green, Jordan Peccia, Ju-Hyeong Park, Janet Macher and Renzo Kottmann.

EM Glass, Y Dribinsky, A Wilke and F Meyer are at Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, USA;
P Yilmaz is at Max Planck Institute for Marine Microbiology, Bremen, Germany;
H Levin is at Building Ecology, Santa Cruz, CA, USA
R Van Pelt, D Wendel and R Knight are at Department of Computer Science, University of Colorado, Boulder, CO, USA;
JA Eisen is at Davis Genome Center, University of California, Davis, CA, USA;
S Huse, A Shipanova and M Sogin are at Marine Biological Laboratory, Woods Hole, MA, USA;
J Stajich is at University of California, Riverside, CA, USA;
R Knight is at Howard Hughes Medical Institute, Chevy Chase, MD, USA;
F Meyer is at Institute for Genomics and Systems Biology, Argonne National Laboratory, Argonne, IL, USA;
LM Schriml is at Department of Epidemiology and Public Health, Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA;
E-mail: marland@mcs.anl.gov
[1]These authors contributed equally to this work.

References

Barrett T, Clark K, Gevorgyan R, Gorelenkov V, Gribov E, Karsch-Mizrachi I *et al.* (2012). BioProject and BioSample databases at NCBI: facilitating capture and organization of metadata. *Nucleic Acids Res* **40**: Database issue D57–D63.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Dunn RR, Fierer N, Henley JB, Leff JW, Menninger HL. (2013). Home life: factors structuring the bacterial diversity found within and between homes. *PLoS One* **8**: e64133.

Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM *et al.* (2011). The Genomic Standards Consortium. *PLoS Biol* **9**: e1001088.

Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B *et al.* (2012). Defining seasonal marine microbial community dynamics. *ISME J* **6**: 298–308.

Hospodsky D, Qian J, Nazaroff WW, Yamamoto N, Bibby K, Rismani-Yazdi H *et al.* (2012). Human occupancy as a source of indoor airborne bacteria. *PLoS One* **7**: e34867.

Kembel SW, Jones E, Kline J, Northcutt D, Stenson J, Womack AM *et al.* (2012). Architectural design influences the diversity and structure of the built environment microbiome. *ISME J* **6**: 1469–1479.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al.* (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L *et al.* (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* **29**: 415–420.