

Data Management Plan – Thrash & Kujawinski

(Meta)transcriptomic and 16S rRNA gene sequence data will be collected and curated according to the Genomic Standards Consortium (http://genc.org/gc_wiki/index.php/Main_Page). Minimum Information about a (Meta)Genome Sequence (MIGS/MIMS) standards (http://genc.org/gc_wiki/index.php/MIGS/MIMS). These outline a standardized format for the minimum information required to accurately describe 16S rRNA gene, genomic and metagenomic data, including metadata, with the goal of facilitating inter-study comparisons and transparency (Field *et al.*, 2008). 16S rRNA gene sequences and (meta)transcriptomic reads will be stored and made available on the NCBI Short Read Archive (SRA). Project storage will occur on the Louisiana State University High Performance Computing cluster Super Mike II, until results are published at which point, in addition to the dissemination strategies outlined above, data will be transferred to secure, archival storage at the LSU HPC.

Spectral data from mass spectrometry analyses will be stored both in existing laboratory databases at the Thrash and Kujawinski labs. All data will be available upon request. We will plan to submit data to the MetaboLights repository (<http://www.ebi.ac.uk/metabolights/>), sponsored by the EMBL-EBI in the United Kingdom, as Dr. Kujawinski has done in the past. However, future political impacts of “Brexit” on this resource may make such submission untenable.

All software and scripts created or modified for data analysis purposes will be submitted and maintained on GitHub (<http://github.com/drjcthrash/>).

Educational data will be maintained as part of overall Department of Biological Sciences CURE section research (headed by Dr. Wischusen), for which a dedicated webpage is currently under construction. LSU Biology has a Communications Coordinator (Dr. Paige Jarreau) who oversees collection and dissemination of CURE findings and educational research outcomes. In addition to formal publication, protocols for specific CURE courses will be posted on the CURE website, as well as survey data (e.g., that in the Project Description) as part of an overall outreach effort to share educational data with others.

Links to all data and code storage will be maintained centrally on the Thrash Laboratory webpage (<http://thethrashlab.com>). Dr. Thrash will maintain responsibility for management, retention, and submission of data to these resources throughout the duration of the project.

References cited:

Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P *et al.* (2008). The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* **26**: 541-547.