

## Data Management Plan

This project will generate transcriptomic data, carbon use efficiencies and iron quotas for model strains under different growth conditions, mutant bacteria cell lines and associated growth, iron and carbon parameters for the mutants vs wild type, and experimental data on bacterial interactions in co-culture.

All data collected through the proposed research will be archived and publication related information and material will be made available immediately. Publicly available data will be delivered in an interoperable format that enables wide-spread data sharing and facilitates secondary use.

All data and metadata originating from the proposed experimental work will be submitted to The Biological and Chemical Oceanography Data Management Office (BCO-DMO) website where it will be housed and available for scholarly use by the academic and scientific community (<http://www.bco-dmo.org/>). All data will conform to established measurement standards and community best practices.

All transcriptomic expression matrices made available through NCBI Gene Expression Omnibus (GEO). The NCBI Sequence Read Archive (SRA) will be the repository for all raw sequencing data. Any phylogenetic analyses not supported by NCBI (eg. .aln, .tre) will be deposited in FigShare.

Newly developed exconjugant or transformant bacterial cell lines will be cryopreserved in quintuplet, entered into a LIMS system associated with a spatially inventoried and weekly maintained cryotank, with subsequent availability to any interested researchers through request to either JCVI or UCSD. Ready-to-clone libraries of the developed ULOOP vectors including pre-designed backbones and all project-driven domesticated parts and detailed protocols will be made available through AddGene (<https://www.addgene.org/>) and protocols.io (<https://www.protocols.io/>), respectively.

The timely deposition of these data and the associated experimental and environmental data in several public databases has been accounted for in the JCVI budget though costs for data processing, deposition, and storage (IT equipment). JCVI is an active participant in the Genome Standards Consortium, which issues policy recommendations on data deposition such as the Minimum Information for Metagenomes Standards (MIMS). In accordance with these policies, experimental details and data-processing information will be included with data depositions, along with the measured environmental parameters. If new scripts for statistical analyses are developed, it will be in the open-source programming language. All publications will include GitHub links to the exact coding notebooks used for cradle-to-branch analysis and data presentation of the published analysis. This allows for reproducibility in addition to knowledge sharing with regards to analysis tools. Currently these are shared as Google CoLab or Jupyter notebooks, though we anticipate adaptation to newly developed coding notebooks if and when they are developed.