Plan Overview

A Data Management Plan created using DMP Tool

Title: Assessing how disruption of the methanogenic community and their syntrophic relationships in tidal freshwater marshes via saltwater intrusion may affect CH4 emissions

Creator: David Berrier

Affiliation: Virginia Commonwealth University (vcu.edu)

Principal Investigator: David Berrier

Data Manager: David Berrier

Funding opportunity number: 15751

Template: VCU Rice Rivers Center Data Management Plan Template

Last modified: 07-08-2024

Copyright information:

Assessing how disruption of the methanogenic community and their syntrophic relationships in tidal freshwater marshes via saltwater intrusion may affect CH4 emissions

Types of data produced

Provide a description of the data you will collect or re-use, including the file types, dataset size, number of expected files or sets, and content. Data types could include digital spatial data, text, spreadsheets, images, 3D models, software, audio files, video files, reports, etc. Consider the following: o What data will be generated in the research? o What data types will you be creating or capturing? o How will you capture or create the data? o If you will be using existing data, state this and include how you will obtain it. o What is the relationship between the data you are collecting and any existing data? o How will the data be processed? o What quality assurance & quality control measures will you employ?

The field site soil redox and salinity will be measured prior to constructing microcosms. After the microcosms are constructed, methane flux, carbon dioxide flux, butyrate degradation rates, sulfate concentrations, organic matter content, C:N ratio and redox data will be collected from each microcosm during the experiment. Gas flux measurements will be determined using gas chromatography, sulfate and butyrate concentrations will be determined using ion chromatography, organic matter content will be determined gravimetrically, C:N ratios will be determined using a Perkin Elmer Series II CHNS/O Analyzer 2400, and redox and salinity will be measured using a salinity redox probe. In addition total DNA/RNA will be extracted from each microcosm and the 16s gene sequence of both bacteria and methanogens will determined using Illumina sequencing. Data other than the 16s gene sequence data, all other data will be entered into excel and stored using XML formats and/or CSV formats. The environmental data will be analyzed by R and python using ANOVA, regressions analysis, t-tests, and correlated to sequence data using multivariate statistics (e.g. mantel tests). Using multiple reps and including blanks for every variable measured will maintain quality control.

Data and metadata standards

Describe the format of your data; think about what details (metadata) someone else would need to be able to use these files. Describe the structural standards that you will apply in making data and metadata available. For example, for most ecological data, documentation should be structured in Ecological Metadata Language (EML). For GIS / geospatial data, FGDC compliant metadata must be generated for each GIS layer created and submitted to the Rice data holdings. An example of metadata could also be as simple as a "readme file" to explain variables, structure of the files, etc. Consider these questions: o Which file formats will you use for your data and why? o What contextual details (metadata) are needed to make the data you capture or collect meaningful? o What form will the metadata describing/documenting your data take? o How will you create or capture these details? o Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)

Environmental data, gas flux data, and butyrate degradation data will be entered into excel and stored as both EML and CSV for ease of use in the statistical program R and python. Sequence data from

Illumina sequencing will be stored in FASTQ format. The data folder will contain a readme folder the indicates the methods used to collect data on each variable, and the units for the data as well as any labeling schemes. Sequencing conditions (e.g. primer concentration and illumina protocol) will be included in the readme file)

Policies for access and sharing

It is the responsibility of researchers and organizations to make results, data, derived data products, and collections available to the research community in a timely manner and at a reasonable cost. In the interest of full and open access, data should be provided at the lowest possible cost to researchers and educators. Describe how and where you will make these data and metadata available to the community. Data and metadata should be made available in a timely and rapid data distribution; make sure you address how soon your data will be available. Indicate what data will be made available and preserved. Will data be accessible on a web page, by email request, via open-access repository, etc.? Consider these questions: o What data will be made available from the study and preserved for the long-term? o How and when will you make the data available? (Include resources needed to make the data available: equipment, systems, expertise, etc.) o What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable? o What related information will be deposited? o What is the process for gaining access to the data? o What resources will be needed to reuse the data? Examples include software or equipment. o How long will the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?

The environmental data, gas flux measurements, butyrate degradation rates and genetic sequence data will be made available after the resulting manuscript and publication. Until then data will be stored on the VCU microbial ecology R drive and on the Franklin lab Dropbox. All raw data will be made available to Rice Center and if deemed appropriate will be stored on the University's Rice Rivers Center Google Drive. There should be no fee to access the data. Data may also be hosted on a website such as http://datadryad.org/, however the data will always be accessible by e-mailing either Dr. Rima Franklin (rbfranklin@vcu.edu) or myself, David Berrier (berrierdj@vcu.edu).

Policies for re-use, distribution

Describe policies surrounding the re-use of your data. If you will not be making the data available for re-use immediately, explain why. Remember that the RRC specifies that you must make your data available no later than two years after your research is complete. If there are other policy issues regarding data access and re-use (ethical or privacy issues, for instance) elaborate on them here. Consider these questions: o When will you make the data available? o Does the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use? o Will any permission restrictions need to be placed on the data? o How long will the original data collector/creator/principal investigator retain the right to use the data before making them available for wider distribution? o Are there any embargo periods for political/commercial/patent reasons? If so, give details. o Are there ethical and privacy issues? If so, how will these be resolved? o Who will hold the intellectual property rights

to the data and how might this affect data access? o What and who are the intended or foreseeable uses / users of the data?

There should be no restrictions on access or reuse of the data. The data will be made avaliable after the publication or two years, which ever is first. The metagenomic sequences, and it's correlation with environmental variables may be used by future wetland ecologist to understand methane production in relation to microbial species composition.

Data storage & preservation

Describe your long-term strategy for storing, archiving and preserving the data you will generate or use. Consider the following: o What is the long-term strategy for maintaining, curating and archiving the data? o Which archive/repository/database have you identified as a place to deposit data? o What procedures does your intended long-term data storage facility have in place for preservation and backup? o How long data will/should be kept beyond the life of the project? o What data will be preserved for the long-term? o On what basis will data be selected for long-term preservation? o What metadata/documentation will be submitted alongside the data or created on deposit/transformation in order to make the data reusable? o What related information will be deposited?

Data will be on the Franklin Lab Dropbox and the VCU microbial ecology R drive for long-term storage. The storage of the data on these two different drives in addition to storage on the Rice Center google drive will insure that the data will be preserved. The data will be preserved indefinitely as part of the Franklin lab's database.