Plan Overview

A Data Management Plan created using DMPTool-Stage

Title: Bacteriophage genomics

Creator: Philippos Tsourkas

Affiliation: University of Nevada, Las Vegas (unlv.edu)

Principal Investigator: Philippos Tsourkas

Data Manager: Philippos Tsourkas

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 17-537

Template: NSF-BIO: Biological Sciences (2015-)

Last modified: 07-01-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Bacteriophage genomics

Data and Materials Produced

Describe the types of data, physical samples or collections, software, curriculum materials, and other materials to be produced in the course of the project. (For collaborative proposals, the DMP must cover all the various data types being collected by each collaborator.)

My lab is currently in possession of 40 *Paenibacillus larvae* phage genomes awaiting annotation. The genomes were obtained from collaborators at Brigham Young University using standard bacteriophage culturing techniques. We expect to receive more *Paenibacillus larvae* phage genomes from our collaborators over the duration of the award.

We will isolate and sequence bacteriophage DNA as part of an undergraduate course on discovery of bacteriophages of the U.S. Southwest. Bacteriophages will be isolated using standard bacteriophage culturing techniques. The DNA will be sequenced using Illumina sequencing technology and assembled using the Geneious 10.0 software.

We will use the TargeTron technology to generate edited *Paenibacillus larvae* phage genomes for the purposes of our experiments.

We will generate a novel tool for bacteriophage genome annotation.

Standards, Formats and Metadata

Describe the standards to be used for all the data types anticipated, including data or file format and metadata.

All DNA sequencing data, complete genomes, and edited genomes will be in .fasta format

Roles and Responsibilities

Describe the roles and responsibilities of all parties with respect to the management of the data (including contingency plans for the departure of key personnel from the project).

All data and data management will be the responsibility of the PI

Dissemination Methods

Describe the dissemination methods that will be used to make data and metadata available to others during the period of the award, and any modifications or additional technical information regarding data access after the grant ends.

All complete, annotated, novel bacteriophage genomes will be uploaded to GenBank immediately prior to publication and made publically available to the scientific community. All complete, annotated, novel bacteriophage genomes will also be published in the journal Genome Announcements.

The completed software will be freely available for download by the scientific community from a dedicated website indefinitely . A community wiki will be created for the software. The software will be actively maintained, updated and expanded indefinitely.

Policies for Data Sharing and Public Access

Describe the PI’s policies for data sharing, public access and re-use, including re-distribution by others and the production of derivatives. Where appropriate, include provisions for protection of privacy, confidentiality, security, intellectual property rights and other rights.

All complete, annotated, novel bacteriophage genomes uploaded to GenBank will be publically available by the scientific community indefinitely. There are no HIPAA or IRB related issues with the data

Archiving, Storage and Preservation

Where relevant, describe plans for archiving data, samples, software, and other research products, and for on-going access to these products through their lifecycle of usefulness to research and education.

All DNA sequencing data and complete bacteriophage genomes will be stored indefinitely on the PI's computer and on three external hard drives as backup storage.