

Plan Overview

A Data Management Plan created using DMPTool

Title: Enhancing Stress Tolerance in Cereal Crops using Transgenerational Epigenetic Memory

Creator: Christopher Benson

Affiliation: Pennsylvania State University (psu.edu)

Principal Investigator: Christopher Benson

Data Manager: Christopher Benson

Funder: United States Department of Agriculture (usda.gov)

Funding opportunity number: 34666

Grant: A7101

Template: USDA-NIFA: National Institute of Food and Agriculture

Project abstract:

Poa annua is a remarkably versatile hybrid grass species with biotypes ranging from perennial to annual. Despite its recent evolutionary origin, *Poa* can be found on all seven continents of the globe including extremely inhospitable locations such as Mount Kilimanjaro and the Antarctic Mainland. My preliminary work suggests that *Poa*'s extraordinary plasticity and tolerance to abiotic and biotic stress is at least partially due to retention of epigenetic marks from one generation to the next, often referred to as transgenerational plant memory. In this research proposal, I suggest using methods in enzyme-linked immunosorbent assays, methylation-sensitive AFLPs, bisulfite conversion, DNA and RNA sequencing, and phenotypic evaluation to elucidate the influence of epigenetic marks on transgenerational memory *Poa annua*. Recent reports have suggested the possibility of harnessing epigenetic memory for stress tolerance in what is now

called epigenetic breeding. Using *Poa annua* as a system to study epigenetic adaptability is a novel approach to discover genes that are under intense epigenetic pressure and contribute toward *Poa*'s extreme adaptability. The results of this work will guide epigenetic breeding in related cereal crops to enhance biotic and abiotic stress-tolerance. Influence of the study would range from the plant health to environmental preservation and improved agricultural economies by mitigating the risk of crop loss and reducing agricultural inputs. As such, this research proposal is in direct accordance with Farm Bill priority areas 1, 4, and 6.

Last modified: 07-17-2018

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

Enhancing Stress Tolerance in Cereal Crops using Transgenerational Epigenetic Memory

Expected Data Type

Describe the type of data (e.g. digital, non-digital), how it will be generated, and whether the data are primary or metadata. Research examples include: lab work, field work and surveys; Education examples include: number of students enrolled/participated, degrees granted, curriculum, and training products; Extension examples include: outreach materials, number of stakeholders reached, number of activities, and assessment questionnaires.

Data for this project will be collected as digital and non-digital types. For lab work in objective 1, enzyme-linked immunosorbent data will be read on an optical density meter and exported to excel where I can calculate percent methylation. Data for methyl-sensitive AFLP will be collected by the Penn State Genomics Core Facility and sent to me through email. Measures will include electropherograms of DNA fragment length. I will then score data based on a presence or absence of cut sites to determine methylation status of the reduced set of DNA fragments. All of these data will be primary data.

Data in objectives two and three will involve both primary and metadata. DNA and RNA-seq typically involves generating raw FASTQ files that can be processed and assembled into genomes and transcriptomes. Count data is typically processed to produced differential expression matrices and methylation status if determined using computational analysis. At the end of processing FASTQ files, I will have multiple transcriptomes and differential expression datasets for the RNA-seq portion. Processing the DNA sequencing data will yield an annotated *Poa annua* genome with methylation status of independent nucleotides.

Objective four data will be collected as phenotypic measurements, and site-specific methylation status. No metadata will be involved. Phenotypic data will be collected as plant measurements and analyzed in excel. Some imaging will also take place and images will be analyzed for leaf area in ImageJ. Also, data regarding methylation at specific candidate loci will be collected using methylation-specific PCR and involve data in the format of bands on an agarose gel during electrophoresis.

Data Format

For scientific data to be readily accessible and usable it is critical to use an appropriate community-recognized standard and machine readable formats when they exist. If the data will be managed in domain-specific workspaces or submitted to public databases, indicate that their required formats will be followed. Regardless of the format used, the data set must contain enough information to allow independent use (understand, validate and use) of the data.

Sequencing data will be made publicly available to NCBI. Raw FASTQ files will be made accessible in the standard Sequence Read Archive (SRA) format. Gene expression data collected during the RNA-seq will be made available through the recognized Gene Expression Omnibus (GEO) format. Transcriptome sequences will also be made available through the accepted Transcriptome Shotgun Sequencing (TSA) portal, and whole genome sequencing data will be made available through the Whole Genome Shotgun Submission (WGS) portal in NCBI. All of these formats are publicly recognized standard for submitting sequencing data. Other data types in this project do not have a community-standard format but will be made accessible in publications and attached supplemental data. All data collections, regardless of format, will be made available with enough information to validate and replicate my conclusion.

Data Storage and Preservation

Data must be stored in a safe environment with adequate measures taken for its long-term preservation. Applicants must describe plans for storing and preserving their data during and after the project and specify the data repositories, if they exist. Databases or data repositories for long-term preservation may be the same that are used to provide Data Sharing and Public Access. Estimate how much data will be preserved and state the planned retention period. Include any strategies, tools, and contingency plans that will be used to avoid data loss, degradation, or damage.

All sequencing data will be safely and permanently preserved in the recognized NCBI database. This includes SRA, GEO, TSA, and WGS datasets (Approximately 80Gb of data). All collected data of this proposal will eventually end up in digital format on the public lab computer. To protect integrity and safety of data, automatic daily backups of the hard drive are stored on two external disks that are safely kept in separate and safe locations. All collected data will be referenced in publications or made available as supplemental tables in publications. This will protect the long-term storage and easy accessibility and interpretation of data that is collected during the proposal. All of these measures will assure avoidance of data loss, damage, and degradation.

Data Sharing and Public Access

Describe your data access and sharing procedures during and after the grant. Name specific repositories and catalogs as appropriate. include a statement, when applicable, of plans to protect confidentiality, personal privacy, proprietary interests, business confidential information, and intellectual property rights. Outline any restrictions such as copyright, confidentiality, patent, appropriate credit, disclaimers, or conditions for use of the data by other parties.

As discussed in previous sections, all sequencing data will be safely and permanently preserved in the recognized NCBI database. This includes SRA, GEO, TSA, and WGS datasets. Remaining data will be made publicly available through supplemental figures and tables in publications. These measures preserve the long-term safety and accessibility of collected data. There is no risk of confidentiality, privacy, or business issues during the terms of the project. There are no restrictions such as copyright, confidentiality, patent, appropriate credit, disclaimers, or conditions for use of the data by other parties.

Roles and Responsibilities

Who will ensure DMP implementation? This is particularly important for multi-investigator and multi-institutional projects. Provide a contingency plan in case key personnel leave the project. Also, what resources will be needed for the DMP? If funds are needed, have they been added to the budget request and budget narrative? Projects must budget sufficient resources to develop and implement the proposed DMP.

The Huff lab already has access to all of the appropriate infrastructure to assure safe storage and management of data including two automatically backed up external hard drives and access to public databases such as NCBI. The PI will be responsible for implementation of the data management plan and will assure that safety measures are met. As a contingency plan, the PI mentor will also assure that data are safely stores and managed. Together, these measures guarantee that the data management plan will be employed for the duration of the study.

Monitoring and Reporting

Successful projects should monitor the implementation of the DMP throughout the life of the project and after, as appropriate. Implementation of the DMP must be a component of annual and final reports to NIFA (REEport) and include progress in data sharing

(publications, database, software, curriculum, outreach materials, etc.). The final report should also describe the data that was produced during the award period and the components that will be stored and preserved (including the expected duration) after the award ends. The DMP should be compliant with the Research Terms and Conditions that govern NIFA-funded project. The DMP is not intended to be a replacement for other grant reporting requirements.

I understand the guidelines of the data management plan terms and conditions and will report on its implementation in the annual and final reports to NIFA.
