Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D17352

Title: CAREER: Defining cell cycle phases and quiescence in any cell type

Creator: Christopher Plaisier - **ORCID:** 0000-0003-3273-5717

Affiliation: Arizona State University (asu.edu)

Principal Investigator: Chris L Plaisier

Data Manager: Chris L Plaisier

Project Administrator: Chris L Plaisier

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: NSF 22-586

Template: NSF-BIO: Biological Sciences

Project abstract:

The long-term goal of this research and education plan is to seamlessly integrate single-cell experimental profiling and computational modeling to characterize the cell cycle and quiescence and introduce high school students to the power of computational modeling in biomedical applications. The cell cycle is a highly conserved biological process that culminates in a cell dividing into two daughter cells. Cells can exit the cell cycle into a quiescent G0 state, and re-enter the cell cycle to divide with the proper signals 1,2. Actively proliferating cell populations are not synchronized to the same cell cycle phase, leading to all cell cycle states being present in the population at any given time. Early studies characterizing the cell cycle used population-level transcriptomic profiling of forcibly synchronized cells to observe the tightly regulated transcriptional programs of the cell cycle 3. A significant limitation of the synchronized population-level experiments is that they do not capture the signature or marker genes for quiescent-like G0 states. The lack of a transcriptional signature and marker genes for quiescent cells is a gap in the field that affects all single-cell studies that are confounded by the biological signal produced by the cell cycle and quiescence. Single-cell RNA-seq (scRNA-seq) has solved this problem by allowing the characterization of all the cell cycle and quiescent states in a population of cells without synchronization. The PI's first publication as an independent researcher used scRNA-seq to characterize all the cell cycle phases of human neural stem cells (hNSCs) and discovered a quiescent-like transcriptional state called Neural G0 4. Gaps in our knowledge about quiescence are the focus of this proposal and will be addressed by: Isolating, characterizing, and classifying quiescent-like G0 cells from

each dermal layer. Creating a causal and mechanistic transcriptional regulatory network from scRNA-seq profiles of hNSCs to discover regulators of the Neural G0 transcriptional signature. Use Boolean network modeling to predict the dynamics of TF regulatory networks for quiescent-like G0 states and validate them experimentally.

Last modified: 01-19-2024

CAREER: Defining cell cycle phases and quiescence in any cell type

The projects from this proposal will produce:

- Flow cytometry data
- scRNA-seq profiles
- Software
  - Single-cell classifier of cell cycle phases and G0
  - Single-cell transcriptional regulatory network inference
  - Boolean network modeling
- Curriculum materials from educational outreach activities
  - High school STEM class unit
  - Online graduate class through ASU

Standards for each inlcude:

- Flow cytometry data - data and capture parameters will be recorded using the MIFlowCyt standards.
- scRNA-seq profiles - data and all relevant metadata will be recorded using the minSCe standards.
- Software will be developed with open-source standards, a GPLv3 license, the code will be placed in github and documented compliant with PEP 8, and tutorials and vignettes will be provided.
- Curriculum materials from educational outreach activities
  - High school STEM class unit - will be develoed to meet the standards of the NGSS and deposited on OER.org.
  - Online graduate class through ASU - will be developed and deployed through ASU online.

The PI Dr. Plaisier will oversee the generation of data and the depositing of data and metadata into relevant repositories.

Standards for each inlcude:

- Flow cytometry data - all data and metadata will be deposited into flowrepository.org.
- scRNA-seq profiles - all data and metadata will be submitted to SRA and GEO.
- Software - code will be placed in github and documented compliant with PEP 8, and tutorials and vignettes will be provided in github.com as well. Packages will be made available through CRAN and PyPI, and as preinstalled docker containers from DockerHub.
- Curriculum materials from educational outreach activities
  - High school STEM class unit - will be deposited on OER.org.
  - Online graduate class through ASU - will be developed and deployed through ASU online.

All data and metadata will be made available alongside all analysis scripts. The PI has an open-source policy for all software that will be developed under GPLv3 software licenses.

The methods used to disseminate will also be used to archive, store, and preserve the data and metadata through public repositories:

- Flow cytometry data - all data and metadata will be deposited into flowrepository.org.
- scRNA-seq profiles - all data and metadata will be submitted to SRA and GEO.
- Software - code will be placed in github and documented compliant with PEP 8, and tutorials and vignettes will be provided in github.com as well. Packages will be made available through CRAN and PyPI, and as

preinstalled docker containers from DockerHub.

- Curriculum materials from educational outreach activities
  - High school STEM class unit - will be deposited on OER.org.
  - Online graduate class through ASU - will be developed and deployed through ASU online.

Planned Research Outputs

Dataset - "Flow cytometry data"

Flow-cytometry will be used to observe single cells in various cell cycle phases based on four stains described in the project description.

Dataset - "scRNA-seq"

Single-cell RNA-seq data will be collected for multiple goals in this proposal

Software - "Software for single cell analysis"

We will develop mutliple different software artifacts during the course of the grant to facilitate single cell analysis of the cell cycle.

Interactive resource - "Curriculum"

Curriculum will be developed to teach students about machine learning.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Flow cytometry data	Dataset	2027-12-31	Open	FlowRepository		Creative Commons Attribution 4.0 International	MIFlowCyt	No	No
scRNA-seq	Dataset	Unspecified	Open	NCBI Gene Expression Omnibus		None specified	minSCe	No	No
Software for single cell analysis	Software	Unspecified	Open	GitHub		None specified	None specified	No	No
Curriculum	Interactive resource	Unspecified	Open	OER Commons		None specified	None specified	No	No