Updated:
02/10/2023

EXAMPLE DATA MANAGEMENT AND SHARING PLAN (in compliance with SF-424 Forms H)

EXAMPLE FOR SINGLE CELL GENOMIC DATA FROM MICE AND HUMANS

Commented [JB1]: *These example DMS Plans are provided for educational purposes to assist applicants with developing Plans but are not intended to be used as templates and **their use does not guarantee approval by NIH. Do not copy/paste this Plan without modifying it to reflect the types of data that are expected to be generated through your project.***

Note that the example DMS Plans may reflect additional expectations established by NIH or specific NIH Institutes, Centers, or Offices that go beyond the DMS Policy. Applicants will need to ensure that their Plan reflects any additional, applicable expectations (including from NIH policies, ICO policies, or as stated in the FOA).

In addition, these examples may reflect resources or policies that are in place at other institutions but that are not necessarily available at the University of Iowa. If needed, investigators can contact Research Data Services (lib-data@uiowa.edu) if they have questions regarding how to best complete their DMS Plan.

Commented [BJY2]: *Example from NIMH: https://www.nimh.nih.gov/sites/default/files/documents/funding/managing-your-grant/resource-sharing-docs/Mouse%20Genomic_Template_v2.docx*

ELEMENT 1: DATA TYPE

A. Types and amount of scientific data expected to be generated in the project:

As detailed in the Research Strategy Section, we propose the generation of a spatially mapped single-cell atlas of the developing mouse brain and include specific deliverables. Our primary deliverable for each modality will be a matrix of cells × (counts in peaks for ATAC, UMIs in genes for RNA, or methylation status for DNAm) along with a dense metadata table with information for each cell. This includes the animal sex, developmental time point, punch of origin with x,y,z coordinates, assigned cluster and inferred cell type, assigned subcluster and inferred cell type, as well as a number of QC metrics (total reads, passing reads, reads in peaks, TSS enrichment, cell barcode combination, date of preparation for each stage, sequencing platform, likelihood of being a doublet, and any other relevant metrics that arise during the project).

The amount and type of data from human cells will depend on the results from the mouse studies. Data sharing plans will be updated when appropriate (likely at the start of year 4 of the grant award).

B. Scientific data that will be preserved and shared, and the rationale for doing so:

The data described in section A will allow researchers to reproduce our publications and will allow them to collect additional data in a similar way to extend our results.

C. Metadata, other relevant data, and associated documentation:

In addition to a detailed methods section for any publications associated with this work, we will provide a detailed step-by-step protocol as a Supplementary Protocol document and maintain active protocols.io protocols for each technology and workflow.

We will additionally release protocol links as metadata to be associated with single-cell data deposited to the Neuroscience Multi-omic Archive.

In addition to providing detailed protocols, our laboratory has hosted visiting scientists to train on the data analysis pipelines developed and deployed by the lab. We welcome the opportunity to continue these training efforts.

ELEMENT 2: RELATED TOOLS, SOFTWARE, AND/OR CODE

All code and software that will be written to analyze the data will be deposited on GitHub for public access and be provided as Supplementary files for any publications. Code will be available no later than when a publication has been submitted.

ELEMENT 3: STANDARDS

We will use the standards that are adopted or defined by NeMO.

ELEMENT 4: DATA PRESERVATION, ACCESS, DISTRIBUTION AND ASSOCIATED TIMELINES

A. Repository where scientific data and metadata will be archived:

Mouse single-cell datasets: All single-cell epigenomics and transcriptomics data will made available through NeMO after initial data processing.

Human single cell data will be deposited to NeMO.

Upon publication we will host processed data matrixes and associated metadata as compressed downloadable archives at NeMO and, when appropriate, as supplementary information in journal publications.

We will release datasets associated with the technological advances proposed in the application once protocols are established and initial analysis performed, at which point data will be released along with a preprint prior to manuscript submission.

B. How scientific data will be findable and identifiable:

Data will be findable for the research community through searches at NeMO. NeMO assigns unique identifiers for each sample.

C. When and how long the scientific data will be made available:
The research community will have access to data as soon as NeMO is able to release it. NeMO will control the deletion of the data sets.

ELEMENT 5: ACCESS, DISTRIBUTION, OR REUSE CONSIDERATIONS

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

There are no special considerations related to accessing or distributing the mouse data to be generated in this award.

B. Whether access to scientific data will be controlled
Access to the human data sets in NeMO is controlled. The NIMH Data Archive Data Access Committee serves as the data access committee for NeMO.

C. Protections for privacy, rights, and confidentiality of human research participants:

The data are deidentified, and there is data access committee (see 5B) that is used by NeMO to evaluate the proposed use of the sensitive data from human subjects. The data archives which will hold the data are funded by NIH, so they have a Certificate of Confidentiality.

ELEMENT 6: OVERSIGHT OF DATA MANAGEMENT AND SHARING

The following individuals will be responsible for data collection, management, storage, retention, and dissemination of project data, including updating and revising the Data Management and Sharing Plan as necessary each year at the time of the Research Performance Progress Report.

[Name, position title, host institution, ORCID, email]

Validation Schedule (this section is required by NIMH)

Data will be validated using the existing pipelines at NeMO. We will submit each dataset once we reach a specific data freeze milestone. Upon each data freeze we will perform an initial phase of analysis that will

culminate in the production of the cell × property matrix and associated metadata, at which point the dataset will be released.

These milestones and target timelines include:

End of 1st quarter of year 2: A full spatial single-cell ATAC-seq map of an entire mouse brain at P14.

End of 4th quarter of year 2: An accompanying spatially-mapped single-cell RNA-seq dataset for a full mouse brain at P14, integrated with the ATAC dataset.

End of 2nd quarter of year 3: A full spatial single-cell ATAC-seq map of each time point.

End of 2nd quarter of year 3: A full spatial single-cell DNA methylation map for P14, integrated with RNA and ATAC datasets.

End of 4th quarter of year 3: The complete spatial single-cell RNA-seq dataset for all time points, integrated with the ATAC data.

End of 4th quarter of year 3: A full spatial single-cell map for all modalities, ATAC, RNA, and DNA methylation, integrated across modalities.