Plan Overview

A Data Management Plan created using DMP Tool

Title: Deciphering the single-nucleus genomic regulatory structure of opioid use disorder in the human brain

Creator: Janitza Montalvo-ortiz

Affiliation: Yale University (yale.edu)

Principal Investigator: Janitza Montalvo-Ortiz

Funder: National Institutes of Health (nih.gov)

Funding opportunity number: PAR-20-225

Template: Digital Curation Centre

Project abstract:

Opioid use disorder (OUD) has reached epidemic levels in the United States, associated with increased rates of hospitalization and drug overdose death, the latter showing a significant steep rise of up to 29.4% in 2020. While genetic risk factors have been identified in recent large-scale genome-wide association studies (GWAS) of OUD, these explain only part of the variance and often map to noncoding regions. Epigenetic modifications have been implicated in the etiology of opioid use disorders (OUD) underlying the gene and environment interplay. We and others have found that alterations of DNA methylation (5mC), one of the most studied epigenetic mechanisms, are associated with OUD in both the human peripheral and postmortem brain. However, most of this work has been done in bulk tissues, which obscures the functional role of cellular diversity in human cells. Further, research is needed to assess additional and novel epigenetic regulatory layers to gain a better understanding of its contribution to gene regulation and its ability to interpret the functionality of GWAS genetic variants in the context of OUD. Here, I offer a novel framework to tackle these gaps and challenges: 1) conduct simultaneous profiling of DNA methylation, DNA hydroxymethylation, and 3D genome structure in single human nuclei, 2) identify OUD-dependent regulatory signatures within cell types and brain regions, 3) evaluate the crosstalk between the different epigenomic regulatory layers, and 4) construct gene programs to finely map OUD GWAS variants and polygenic signals to function. This comprehensive single-cell multiomics mapping of OUD will examine the dorsolateral prefrontal cortex (DLPFC), amygdala (BLA), and nucleus accumbens (NAcc), part of the addiction circuitry, of human postmortem brain samples collected from the UTHealth Brain Collection datasets and using the VA Brain Bank (NPBB) as a validation cohort. This work is highly innovative and will open new lines of research on the genetics and epigenetics of OUD by providing novel mechanistic insights into its gene regulatory structure in the human brain. This proposed study will identify and help inform molecular targets to be used as prevention and treatment efforts for individuals suffering from OUD.

Start date: 09-01-2023

End date: 08-31-2028

Last modified: 07-08-2024

Copyright information:

Deciphering the single-nucleus genomic regulatory structure of opioid use disorder in the human brain

Data Collection

What data will you collect or create?

This project will produce genomic, epigenomic, transcriptomic, and chromatin organization data from human postmortem brain (20 opioid use disorder cases and 20 healthy controls). Specimens will be received deidentified and the PI won't have access to identifiers. To further protect donors' identities, summary statistic data will be made available for sharing easily at the time of journal publication. Individual-level data will be made available in dbGAP with a standard dbGAP application.

How will the data be collected or created?

Deidentified human postmortem tissue specimens will be received from UTHealth Brain Bank. The tissue will be preprocessed at the Division of Human Genetics laboratory and processed in close collaboration with the Yale Center for Genome Analysis. Based on ethical considerations, only the following data produced in the course of the project will be preserved and shared: 1) Subject ID#, sex, age of death, postmortem interval, pH, ethnicity, DSM diagnoses for substance use disorders, and diagnoses for major psychiatric comorbid conditions, genotyping data, sequencing data and platform used, method for DNA extraction, inclusion/exclusion criteria, sample size, statistical packages to analyze data, and summary-level data.

Documentation and Metadata

What documentation and metadata will accompany the data?

To facilitate the interpretation of the data, data dictionary and data collection instrument, the number of samples collected, where the samples were collected, and study design (e.g., case/control) will be created, shared, and associated with the relevant datasets.

Ethics and Legal Compliance

How will you manage any ethical issues?

Samples will be deidentified. The PI does not have access to identifiers. Phenotype data will be made available in .txt format and will not require the use of specialized tools to be accessed or manipulated. Summary statistics will be made available in .txt format and will not require the use of specialized tools to be accessed or manipulated; or some other format that does not require the use of specialized tools.

In accordance with FAIR Principles for data, we will use open file formats (e.g., JPEG, MP4, CSV, TXT, PDF, HTML, etc.) and persistent unique identifiers (PIDs).

Information needed to make use of this data and metadata items will be included wherever applicable.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

Storage and Backup

How will the data be stored and backed up during the research?

All data that can be shared will be deposited in dbGAP and made available upon approval of the application.

How will you manage access and security?

All data that can be shared will be deposited in dbGAP and made available upon approval of the application. dbGAP stores data under a persistent identifier and is searchable. We will also provide a reference for sharing in all scholarly publications. No (non-biological) individual identifiers traceable back to the individual participant will be provided.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Shared data generated from this project will be made available at the time of publication (summary statistics) or the end of the funding period (individual-level data), whichever comes first.

What is the long-term preservation plan for the dataset?

The duration of preservation and sharing of the data will be a minimum of 5 years after the end of the funding period.

Data Sharing

How will you share the data?

Individual-level data will be shared only via dbGAP application. There are no anticipated factors or limitations that will affect the access, distribution, or reuse of the scientific data generated by the proposal.

Are any restrictions on data sharing required?

Access to the scientific data will require dbGAP application.

Responsibilities and Resources

Who will be responsible for data management?

PI Janitza Montalvo-Ortiz will be responsible for the day-to-day oversight of lab/team data management activities and data sharing. Broader issues of DMS Plan compliance oversight and reporting will be handled by the PI and Co-I team as part of general Yale University stewardship, reporting, and compliance processes.

What resources will you require to deliver your plan?

Access to high performance computing clusters and services at Yale University.

Planned Research Outputs

Dataset - "Single-nuclei multiomics of opioid use disorder", working title"

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?
Single-nuclei multiomics of opioid use disorder", ...	Dataset	Unspecified	Restricted	None specified		None specified	None specified	No	No