Introducing the NIDA Core "Center of Excellence" in Omics, Systems Genetics, and the Addictome

Laura M. Saba¹, Saunak Sen^{2,3}, and Robert W. Williams³

¹Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

²Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN, USA

³Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, TN, USA

ABSTRACT

Aims: The purpose of the NIDA P30 Core Center of Excellence in Omics, Systems Genetics, and the Addictome is to empower current and future NIDA researchers to examine the interwoven roles of genetic and environmental variation on drug abuse risk, relapse, and treatment by eliminating many of the technical barriers to analysis.

Methods: We are assembling sophisticated omics resources that will give investigators mechanistic and behavioral insights and provide training in omics, systems genetics, and advanced computational/statistical modeling. In the Transcriptome Informatics and Mechanisms research core, we are gathering large genome and transcriptome data sets for rodent models of addiction and upgrade methods and tools for quantitation and integration of these data to uncover molecular mechanisms of addiction. In the Systems Analytics and Modeling research core, we will use innovative systems genetics methods to understand the linkage between DNA differences, environmental risks, and the differential risk of drug abuse and relapse. Our Pilot core will catalyze new collaborations among early career investigators in the field of addiction research.

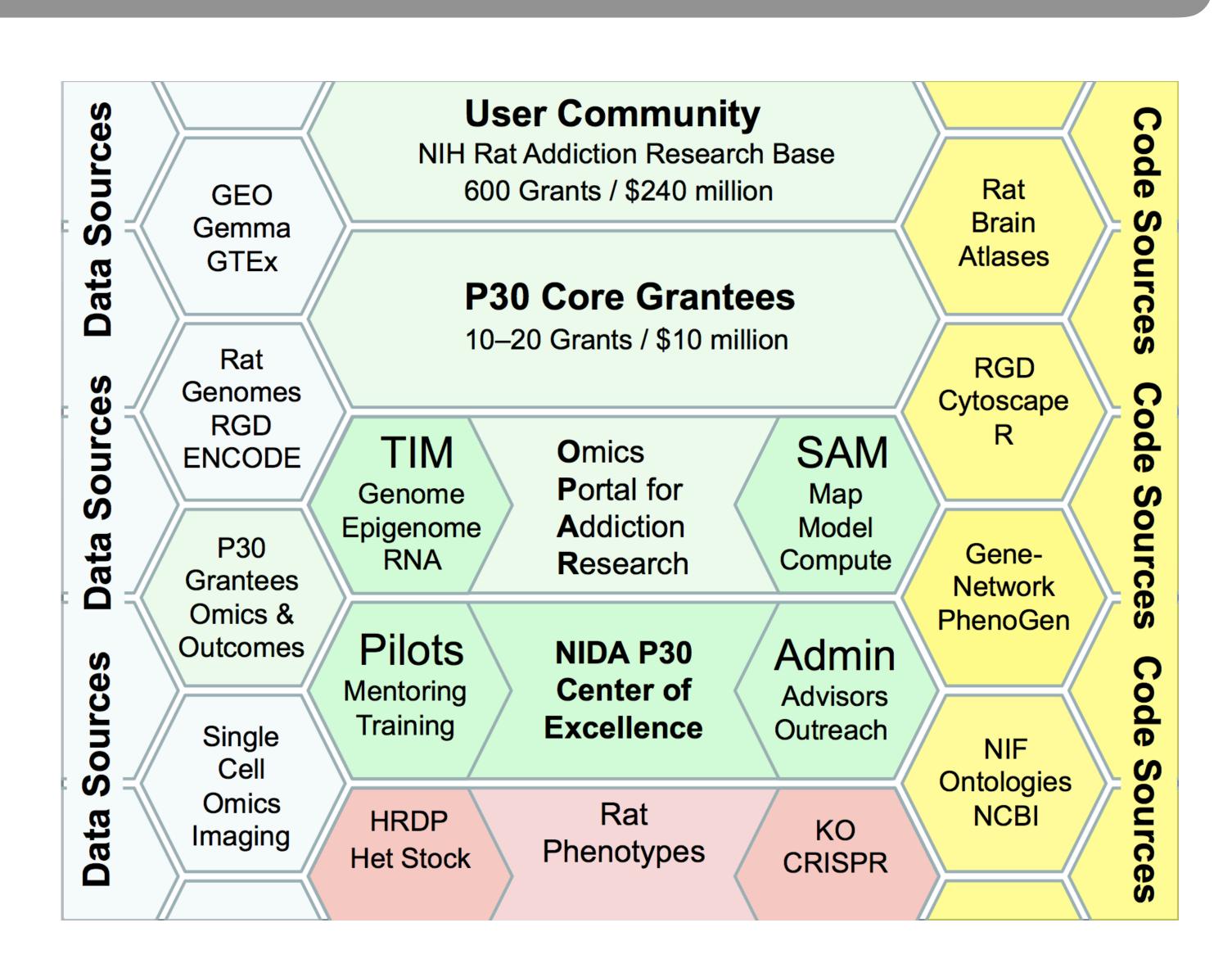
Results: We are building The Omics Portal for Addiction Research (OPAR) that integrates old and new transcriptomic datasets with a focus on rat genome and transcriptome datasets, with the long-term goal of improved translational relevance to human addiction. OPAR will eventually include analytical tools developed by the research cores. Researchers will have access to the latest rat genome, transcriptome, and epigenome data sets matched to powerful multicore GPU-based servers.

Conclusions: As we build this resource for reproducible research in addiction, we are seeking input from the research community on how to best meet their needs and wants. Our vision is that the Center can be used a catalyst for building an NIDA Addictome Portal that will include all genomic data relevant to addiction research.

INTRODUCTION

Many addiction research groups realize the power of precision medicine and want to incorporate this approach into their current research. However, often this requires that individual research groups generate, process, store, and combine massive genome, transcriptome, behavior, and health record datasets. This creates several barriers that research groups must overcome before they can capitalize on this type of research. Two of the top barriers are:

- 1. access to <u>high-end computational resources</u> and appropriate analytic frameworks
- 2. bioinformatics/biostatistics collaborators with the domain expertise to handle high-dimensional datasets (epigenomes, proteomes, images, metagenomes, single-cell assays) related to addiction.



METHODS

Systems Analytics and Modeling (SAM) Research Support Core

The SAM Research Support Core, led by Dr. Saunak Sen, has three goals:

- 1. To support researchers analyzing complex addictome data—essentially to amplify their efforts by embedding their data in the Omics Portal for Addiction Research
- 2. To integrate a wide range of software for open-access analytics to enable researchers without computational skills to address complex questions using vetted methods and interfaces
- 3. To develop new methods to dissect genotype-to-phenotype (G2P) relations in model organisms—rats in particular

Current projects under development that will be implemented in OPAR include:

- Smart search capabilities to connect investigator queries with relevant information across rat genetic websites and databases
- Fast linear mixed model computations for QTL analysis with population structure (e.g., Heterogeneous Stock and Hybrid Rat Diversity Panel)
- Matrix linear models for mapping multivariate traits
- Methods/workflow for single cell RNA-Seq data
- Development of data mining methods

Transcriptome Informatics and Mechanisms (TIM) Research Support Core

The TIM Research Support Core, led by Dr. Laura Saba, has three goals:

- 1. To provide databases in a format that is easily utilized for systems genetics research
- 2. To centralize the labor- and computationally-intensive tasks of processing RNA-Seq data
- 3. To share our statistical and computational expertise through improved analysis pipelines and effective visualizations of systems genetics analysis results

Current projects under development that will be implemented in OPAR include:

- Brain and liver transcriptome reconstructions from 43 strains of the Hybrid Rat Diversity Panel
- Quantitation and transcriptome reconstruction of over 80 heterogeneous stock rats in 5 different brain regions based on data generated as part of the NIDA Center for GWAS in Outbred Rats (http://www.ratgenes.org)
- Development of pipelines for RNA-Seq data processing in mass
- Evaluating protein-coding potential of novel transcripts based on RNA-Seq and DNA sequence information

RESULTS

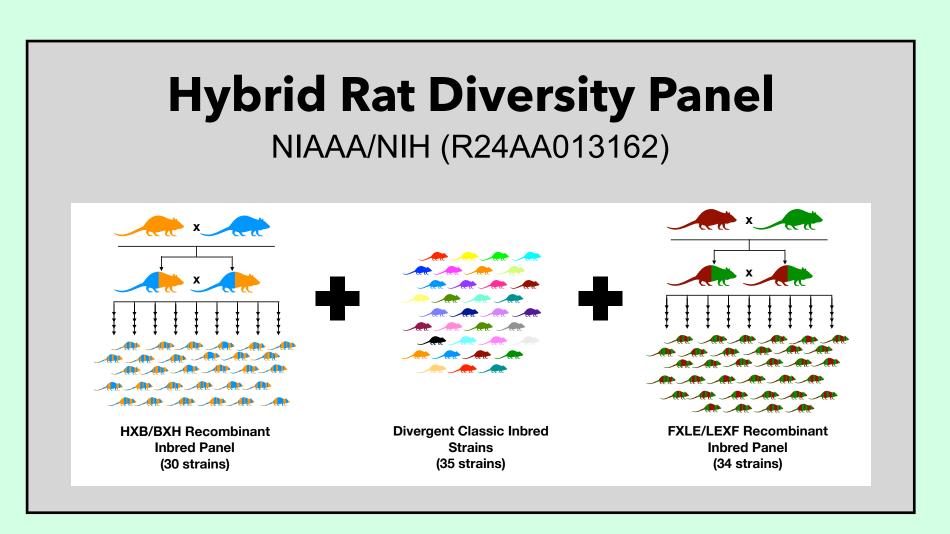
Omics Portal for Addiction Research (http://www.opar.io)

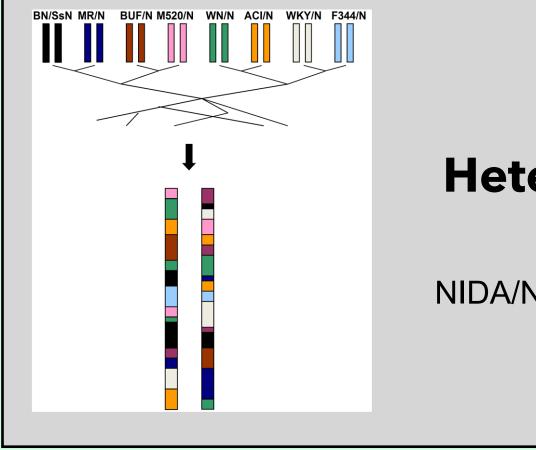
Access to Curated Data - The Rat Addictome

Past: Rat addiction related data sets (GEO & GEMMA)

Present: Two foundational rat populations (HRDP & NIH/HS)

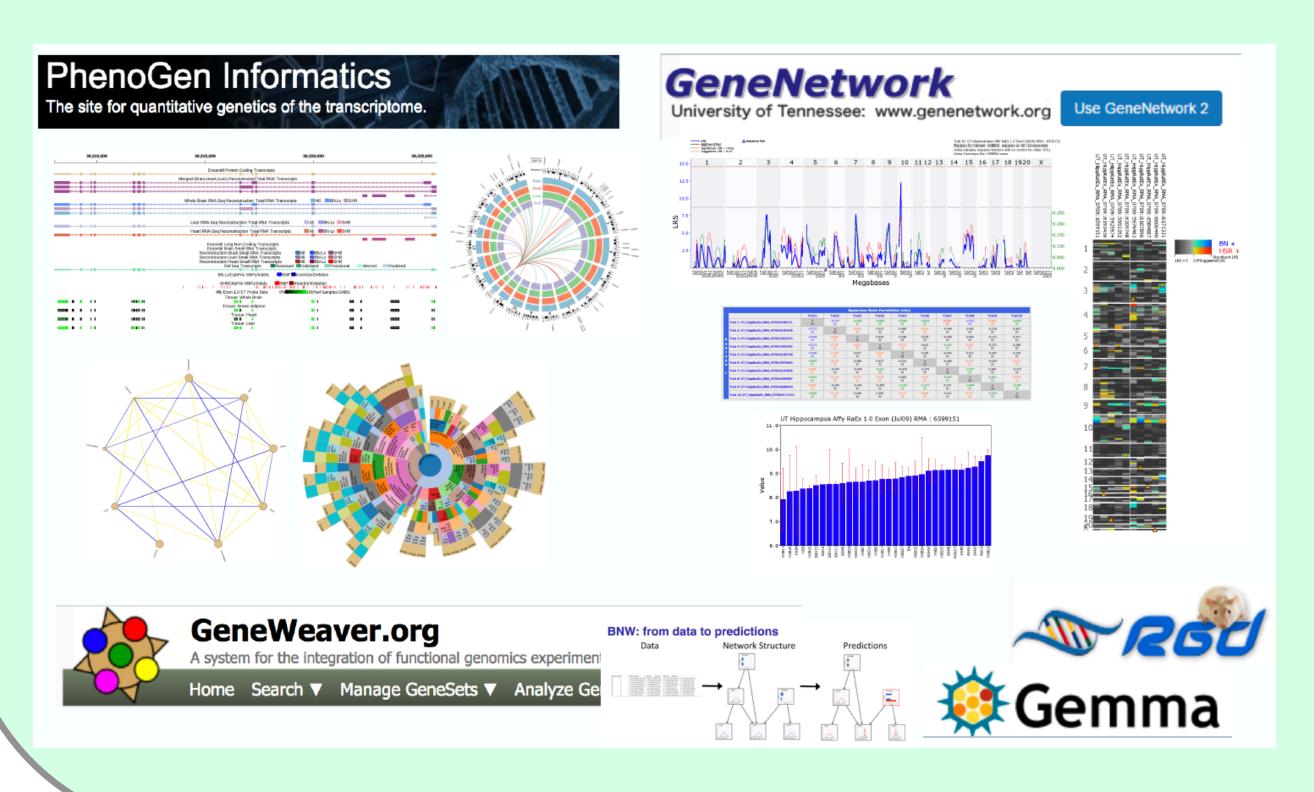
Future: Data direction from NIDA/NIAAA researchers





NIH Heterogeneous Stock NIDA/NIH (P50DA037844)

Access to the Best Tools



More Intuitive Search Capabilities

Where in the brain is P2rx4 expressed?

What rat genes are expressed in hippocampus within 10 Mb of my QTL?

What other genes and traits are associated with P2rx4?

Is P2rx4 genetically controlled?

Are there multiple isforms of P2rx4 in the nucleus accumbens?

What genes expressed in the VTA have a binding site for mir-23?

Training

Topics: OPAR Usage, RNA Processing and Analysis, Systems Genetics

Modes: Workshops at meetings, webinars, on-site training by request, short course in Colorado/
Tennessee

Services

Study Design Consultation: types of rat genetics populations, public resources, connections with other researchers

RNA-Seq Data Processing: quantitation/analysis of smaller studies, access to code/documentation for standard analysis pipelines, collaboration on bigger projects including those that will eventually become crucial data in OPAR

Pilot Grants

1 to 2 year collaborative grants of \$20,000 per year to gather preliminary data using rats on a novel phenotype or novel omics technology.

Next Call for Applications: April 2019

DISCUSSION/CONCLUSIONS

CONFLICT OF INTEREST

GRANT SUPPORT

As we build this resource for reproducible research in addiction, we are seeking input from the research community on how to best meet their needs and wants. Our vision is that the Center can be used a catalyst for building an NIDA Addictome Portal that will include all genomic data relevant to addiction research.

The authors have no financial relationships that relate to the topic of this poster/presentation.

This work is supported by NIDA (P30DA044223)