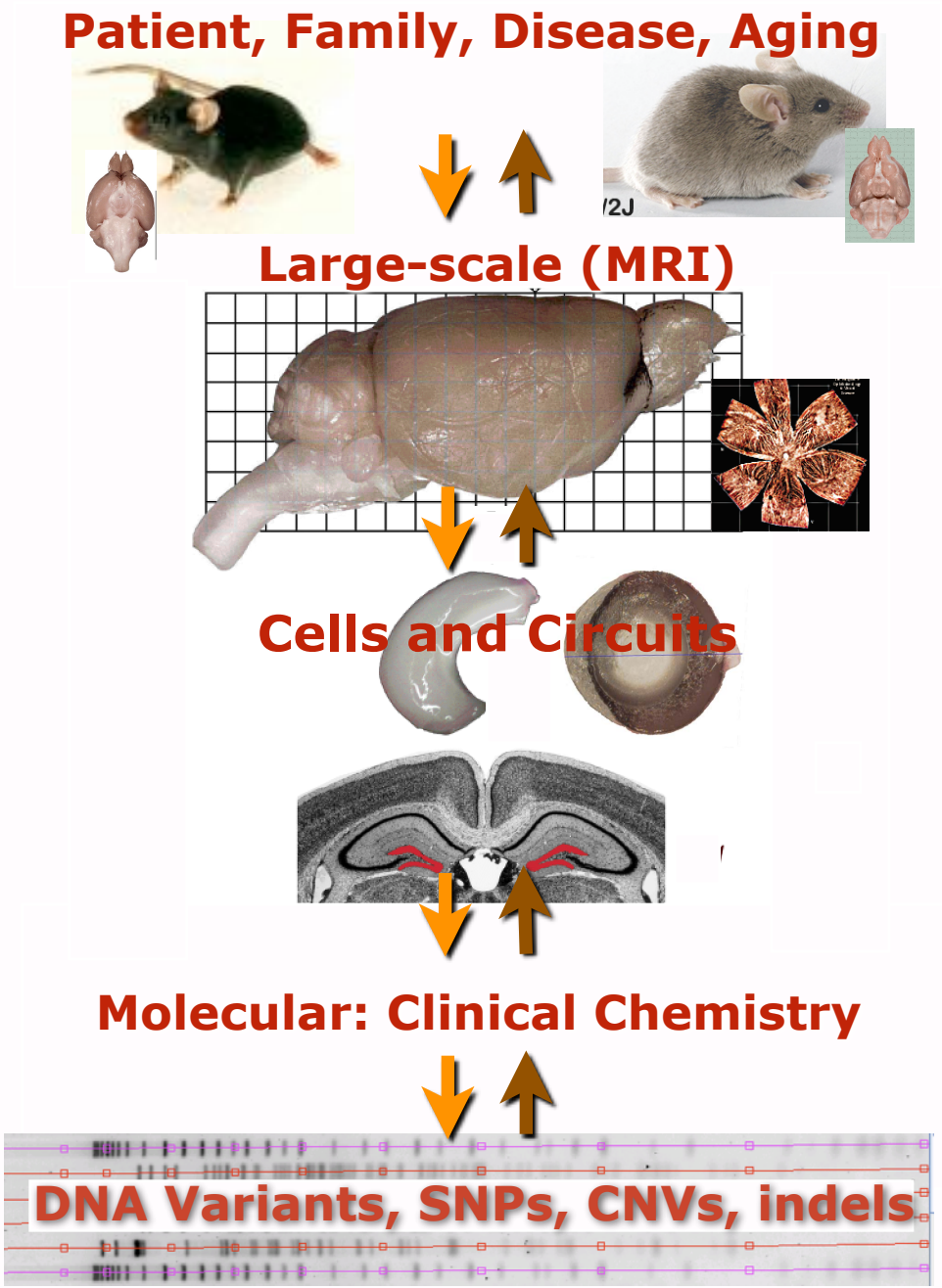
Robert W Williams PhD
University Tennessee

*My background:*
*Neuroscience and* statistical genetic. Studies of eye, brain, metabolism, addiction, aging.
*Methods:* Imaging, stereology, DNA sequence, transcriptome, phenomes
*Web apps for systems biology, genetics, neuroscience:*
*GeneNetwork.org, Mouse Brain Library (MBL.org), Addiction genomics (OPAR.io)*
*Statistical methods for better genome-phenome modeling, translation to human*



# *Research Interests:* Resources, data, workflows for experimental and quantitative P4 medicine

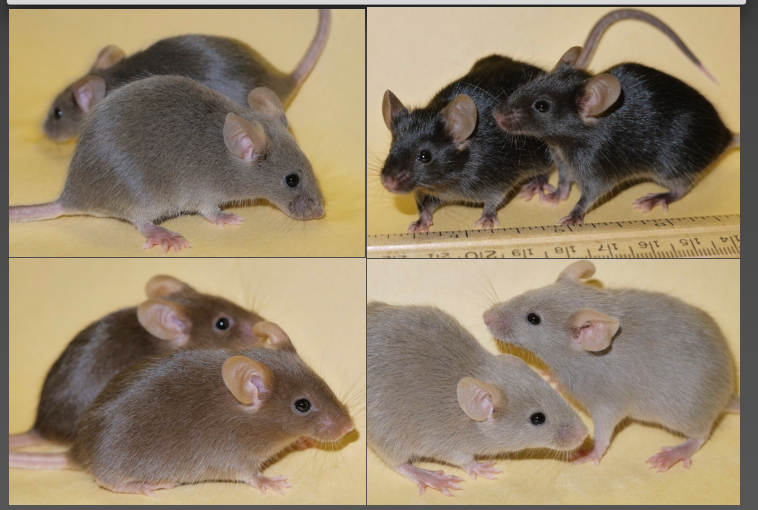
## *What types of data:*

Mainly quantitative phenotype data (usually poorly and non-uniform metadata), 180 matched genomes.

Wide range of omics data: BXD family of isogenic strains is the most deeply quantified cohort.

Data to 1973 rescued and in GeneNetwork with all new data.

Data coupled with a statistical genetics web stack.

Funded by UT, NIDA, NIGMS, NIAAA, NIA, NEI, NCI, NIMH

# Reaction to Task Statement

*Tools & Practices* to improve data capture, quality, management, preservation, and access, while reducing risk and cost:

- Establish low-cost secure enclaves of PHI, EHR, genomes
- Develop encrypted but computable file structure for human
- Data loss and cost is NOT the core problem. Challenge is initial data and metadata capture. Genomic commons success.

*Methods to get researchers to manage data and costs:*

- Motivate use of IPFS, GUIX-Docker, Galaxy, OSF, blockchain
- Push against static publication to Jupyter, R/shiny, and other publication methods that support data that breathes & breeds

*Implementation barriers*:

- 1. Great OSS tools, but as volatile as ever
- 2. Focus on narratives rather than data and analysis

https://www.nature.com/articles/d41586-018-07196-1