# Broad Aims of PsychENCODE data analysis
(based on the gdoc)

- Establish uniform data processing pipelines
- Characterize Functional Genomics Data from Normal Brains
  - Comparison with non-brain specific resources (GTEx, ENCODE, Roadmap)
    - Discover brain specific spliced transcripts and enhancers
  - Enlarged group of eQTLs
  - Characterization of Brain Development
- Conduct integrative analysis to discover functional genomic elements and genotypes associated with psychiatric disease

# Building up to Broad Analyses

- Uniform collection & processing of data

- Developing analyses specific to particular individuals (ie postdocs and students) & sub-projects

- Joining these together into larger themes

# Example Focused Project #1:
# Analysis of Repetitive Element transcription in brain

- Use RNA-seq samples from different developmental stages and conditions

- Estimate transcription signatures of L1Hs, ALUY, HERVK functional transcripts
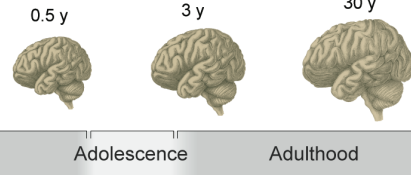
- Remove the effect of background transcription
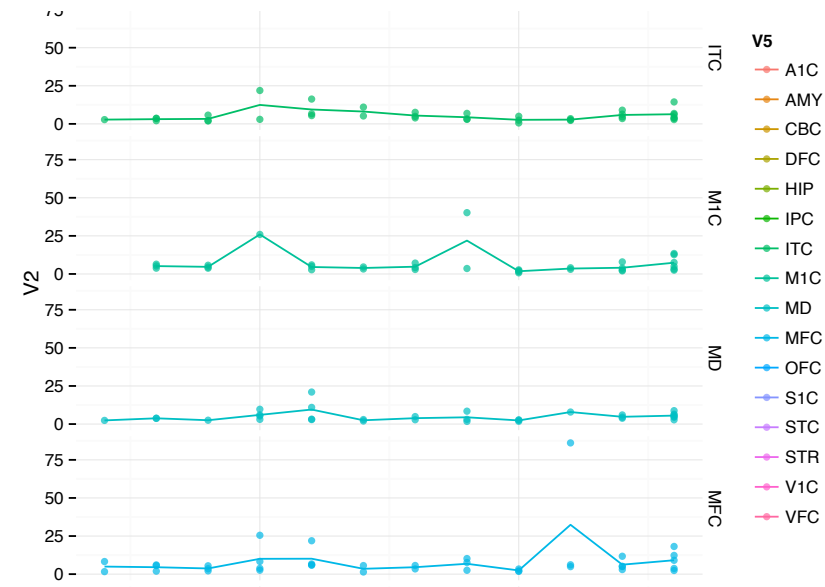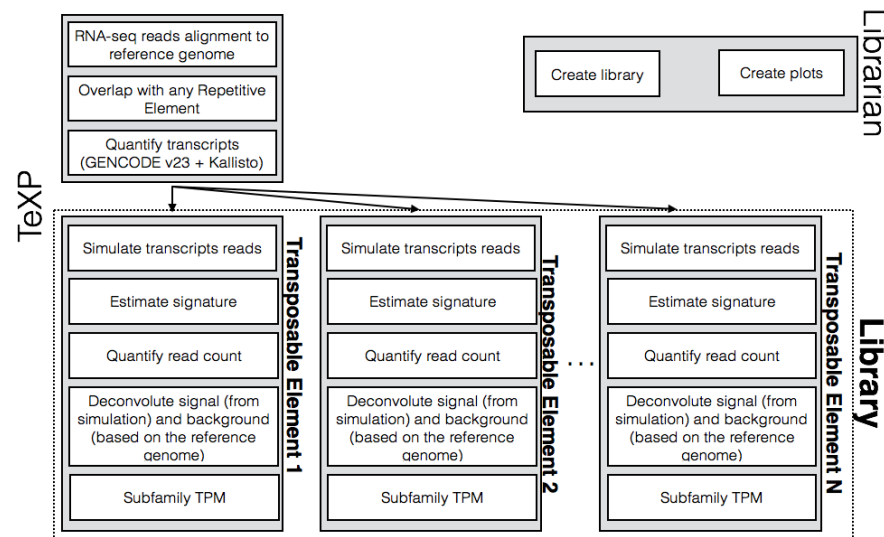
# Ex of Focused Project #2: Allelic analysis in brain



**Goal:** Assess allelic RNA expression in human brain

A: identify allelic genes across brain regions

B: identify correlates with neuropsychiatric disease

**BrainSpan**

5 adult brains x16 regions (incl. DLPFC)

**CommonMind**
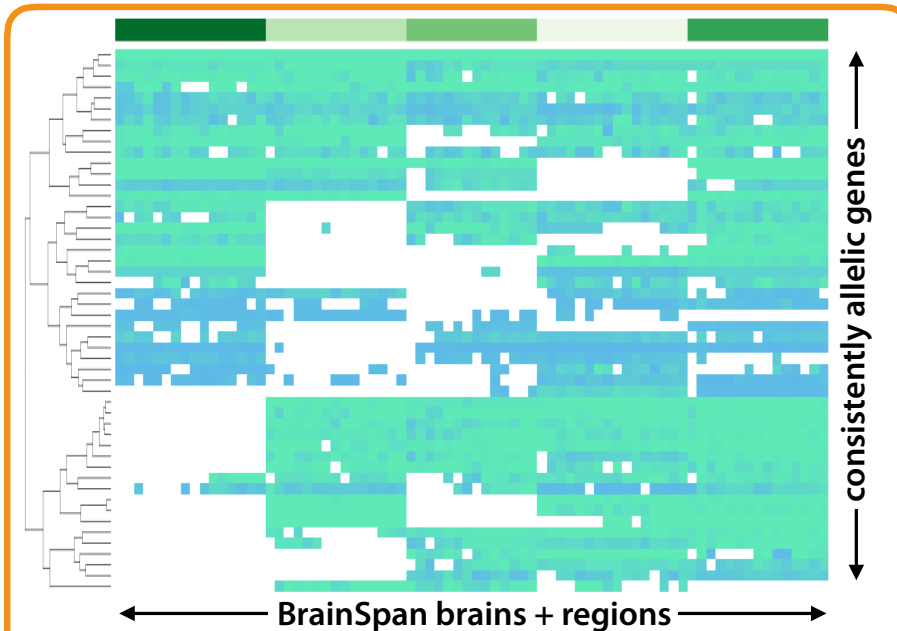
~600 adult brains x5 phenotypes

**1: genotype + RNA-seq**

**2: construct personal genomes & run our AlleleSeq tool**

1 Construct diploid personal genome

2 Merge and align reads from ChIP-seq/RNA-seq

3 Statistical comparison of read counts between alleles

**3: identify consistently allelic genes across individuals and brain regions**