Introduction: Strong genetic associations have been found for a number of psychiatric disorders. However, understanding the underlying molecular mechanisms remains challenging.

Rationale: To address this, PsychENCODE has developed a comprehensive resource and integrative models for the functional genomics of the human brain (resource.psychencode.org).

Results: The resource is arranged in a pyramidal structure.

Bottom – The base of the resource is a large number of datasets generated by PsychENCODE including ~3800 bulk transcriptome, chromatin (ATAC-seq and ChIP-seq), genotype and Hi-C datasets and single-cell transcriptomic data from ~32K cells for major brain regions such as prefrontal cortex and cerebellum. We have merged these with publicly available data from GTEx, ENCODE, Roadmap Epigenomics, and published single-cell analyses. Then via uniform processing, we created a harmonized resource, allowing us to survey functional genomics data on the brain over 1866 individuals.

Middle – Based on this uniformly processed dataset, we created a number of derived data products. These include: (1) lists of brain-expressed genes, co-expression modules, single-cell expression profiles for many brain cell types ; (2) ~79K brain-active enhancers with associated Hi-C loops and TADs; and (3) ~2.5M eQTLs comprising ~238K linkage-disequilibrium-independent SNPs, and a large set of other types of QTLs associated with splice isoforms, cell fractions and chromatin activity (i.e., isoQTLs, fQTLs, cQTLs). Using these, we found >88% of the cross-population variation in brain gene expression can be accounted for by cell-fraction changes. Furthermore, a number of disorders as well as aging are associated with distinct changes in cell-type proportions (e.g. an increase in proportion of microglia and astrocytes for autism). The derived data also enables consistent comparison between brain and other tissues. In particular, using spectral analyses, we found that the brain has distinct expression and epigenetic patterns, including a greater extent of non-coding transcription.

Top – The top level of the resource consists of integrative networks for regulation and machine-learning models for disease prediction. The networks include a full gene-regulatory network for the brain, linking TFs, enhancers and target genes. It is based on merging the QTLs, generalized element-activity correlations and Hi-C data. Using this network, we link more disease genes to GWAS variants for psychiatric disorders than previously. For instance, for schizophrenia, we link 321 genes to the 142 reported GWAS loci. We then embedded the regulatory network into a deep-learning model to predict psychiatric phenotypes from genotype and expression. Our model gives a ~6X improvement in prediction over additive polygenic risk scores. Moreover, it achieves a ~3X improvement over additive models even when the gene expression data are imputed, highlighting the value of having just a small amount of transcriptome data for disease prediction. Finally, it highlights key genes and pathways associated with disorder prediction, including a variety of immunological, synaptic and metabolic pathways, recapitulating *de novo* many results from more targeted analyses.

Conclusion: Our resource and integrative analyses have uncovered genomic elements and networks in the brain, which in turn have provided insight into the molecular mechanisms underlying psychiatric disorders. Our deep learning model improves disease risk prediction over traditional approaches and can readily be extended with additional data types (e.g., miRNA and neuroimaging).

Summary figure caption: A comprehensive functional genomic resource for the adult human brain. The resource is hierarchically organized into three layers. The bottom layer includes the raw sequencing datasets for various observed traits such as schizophrenia. The middle layer consists of the derived datasets, including brain-related functional genomic elements and QTLs. The top layer contains integrated models, which use these data to link genotype to observed phenotypes.