**Gerstein lab experience with identifying QTLs and associating non-coding genetic variants to functional activities at cellular resolution**

The Gerstein lab has extensive experience in identifying QTLs of various types, as well as relating these QTL types to one another. In order to identify cis-eQTLs, we closely adhered to the protocols used by GTEx, as well as the standardized approach established by ENCODE to uniformly process RNA-seq data. Furthermore, we benchmarked our results with direct comparisons to available data files in the GTEx portal (gtexportal.org)[2](https://paperpile.com/c/kP6ZZc/8cox). We followed existing protocols to ensure compatibility between our results and those previously published. This also enabled us to compare our results and those published previously. Using these methods in PsychENCODE I [3](https://paperpile.com/c/kP6ZZc/G4HB), we identified a set of eQTLs several times as large as those in previous studies (targeting a saturating proportion of protein-coding genes). We found ~2.5 million eQTLs comprising ~238,000 linkage-disequilibrium–independent SNPs. The lists of significant eQTLs are available at resource.psychencode.org. In particular, we measured the similarity between PsychENCODE brain DLPFC eQTLs and GTEx eQTLs of other tissues using pi\_1 values and SNP-eGene overlap rates. We found that both pi\_1 values and SNP-eGene overlap rates are higher for brain DLPFC than for the other tissues. In a separate analysis, we intersected our eQTLs with other QTL types (such as cQTLs and cell fraction QTLs). We found that eQTL-cQTL intersections often suggested that an expression-modulating function of an eQTL derived from chromatin changes (an example which we discussed in our study was that related to mTOR).

We also have experience in identifying trans-eQTLs. We matched RNA-seq and genotype data to generate trans-eQTLs using QTLtools, and comprehensively controlled for false-positive trans-eQTLs. We identified ~562K trans-eQTLs and found that more than 80% of trans-eGenes overlap with cis-eGenes. We then explored possible mechanisms for trans-eQTL associations. We found that some cis-eGenes and trans-eGenes show the TF-target gene relationship. Trans-eQTL variants and trans-genes show higher interaction frequencies than randomly selected pairs. These findings indicated that trans-eQTL variants interacted with genes via interchromosomal interactions. We also found that if one trans-eQTL variant is associated with multiple trans-genes, those trans-genes show higher co-expression patterns than randomly selected genes.

# REFERENCES

1. [Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).](http://paperpile.com/b/kP6ZZc/l6Ri)

2. [GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).](http://paperpile.com/b/kP6ZZc/8cox)

3. [Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, (2018).](http://paperpile.com/b/kP6ZZc/G4HB)