A Bioinformatics Pipeline for Identifying Functional Explanations of SNP-Phenotype Associations on a Transcriptional Level

◆ **Stephan Hu**, Dr. Xi Rao, Dr. Yunlong Liu

Background: Genome-wide association studies (GWAS) have identified thousands of associations between single nucleotide polymorphisms (SNP) and traits of interest. These associations do not offer biological or functional explanations for differences in phenotype, and sorting through thousands to millions of SNPs makes finding explanations difficult. This study works on filling the gap between these associations and their functional effect on phenotype by identifying variants that are associated due, at least in part, to their effect on a transcriptional level.

Methods: GWAS analysis and RNA-sequencing was run on the post-mortem brain tissue of both heavy drinkers and non-drinkers. Genes that were associated with differential transcript production were overlaid with chromatin interaction data to identify potential enhancers. A number of properties of enhancers, such as their increased stability while bound, their location within topologically associated domains, and their location within transcription factor binding sites, were used to narrow down the list.

Results: Identified enhancers offer a potential functional explanation for the association between a SNP and trait.

Conclusion: A large gap currently exists between associations obtained from genome-wide association studies and a functional explanation of these associations. This study shows how a bioinformatics approach can fill this gap. This work can be extended from enhancers to include noncoding regulation such as miRNA binding and splice variation. The combination of these three will explain a large range of functional variation due to transcriptional differences. Future work should also consider methods of addressing functional differences on translation and post-translational levels. Although here it is used for an alcohol use disorder study, this protocol has the potential to be used in a wide range of statistical genomic settings to find functional explanations for associations between SNP and trait.

Stephan Hu is a third-year medical student currently undecided on his specialty of interest. He states that when it comes to research, having resilience and perseverance is just as important as being a critical thinker. Being able to work through numerous frustrations and setbacks, including many that have nothing to do with the actual science, was critical to the team's success.