Post-GWAS data integration identifies risk factors for Alzheimer’s disease

Qiongshi Lu

University of Wisconsin-Madison
E-mail: qlu@biostat.wisc.edu

Abstract: Despite the findings in genome-wide association studies (GWAS) for late-onset Alzheimer’s disease (LOAD), our understanding of its genetic architecture is far from complete. Transcriptome-wide association analysis that integrates GWAS data with large-scale transcriptomic databases is a powerful method to study the genetic architecture of complex traits. However, it is challenging to effectively utilize transcriptomic information given limited and unbalanced sample sizes in different tissues. Here we introduce and apply UTMOST, a principled framework to jointly impute gene expression across multiple tissues and perform cross-tissue gene-level association analysis using GWAS summary statistics. Compared with single-tissue methods, UTMOST achieved 39% improvement in expression imputation accuracy and generated effective imputation models for 120% more genes in each tissue. A total of 69 genes reached the Bonferroni-corrected significance level in the transcriptome-wide association meta-analysis for LOAD. Among these findings, we identified novel risk genes at known LOAD-associated loci as well as five novel risk loci. Several genes, including IL10 and ADRA1A, also have therapeutic potential to improve neurodegeneration. Cross-tissue conditional analysis further fine-mapped IL10 as the functional gene at the CR1 locus, a well-replicated risk locus for LOAD. Extension of this framework to perform biobank-wide association analysis will also be discussed. Overall, integrated analysis of transcriptomic annotations and biobank information provides insights into the genetic basis of LOAD and may guide functional studies in the future.