Linking Quantitative Traits with Molecular Evidence for Alzheimer's Disease

Yinuo Zhang College of Arts and Sciences Class of 2023, supported by Xiaohui Yao, Shan Cong, and Li Shen
Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine,
University of Pennsylvania, Philadelphia, PA, USA
*Please address correspondence to Li Shen PhD (Li.Shen@pennmedicine.upenn.edu)

1. Introduction

The relationship between genetic variations of single nucleotide polymorphism (SNP) and brain imaging quantitative traits (QTs) is integral in the understanding of Alzheimer's Disease (AD). Genome-wide association studies (GWAS) is useful in identifying SNP values as risk factors of developing AD; however, there are still more ways to analyze GWAS results through longitudinal analysis as well as multivariate analysis. In this project, I focused on visualizing GWAS results through PLINK and MultiPhen in R.

2. Materials And Methods

I performed univariate GWAS analysis and longitudinal GWAS analysis using PLINK: https://zzz.bwh.harvard.edu/plink/dataman.shtml [1]. The binary genotype files were obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI): http://www.adni-info.org/. The phenotype files were obtained from Quantitative Templates for the Progression of Alzheimer's Disease (QT-PAD): http://www.pi4cs.org/qt-pad-challenge. Figure 2 shows the focus of this project on the relationships between brain imaging, biomarker measurement, genomics and clinical outcome. I also performed multivariate GWAS analysis using the R package MultiPhen: https://cran.r-project.org/web/packages/MultiPhen/index.html [2].

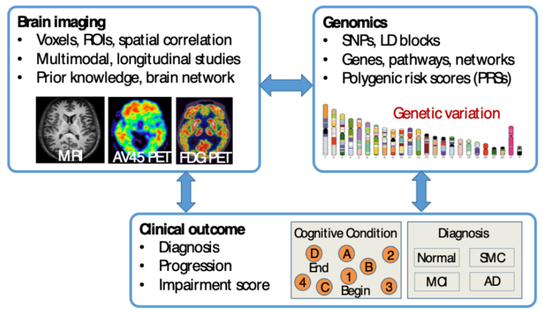

Figure 1. Graph showing the design of this project through GWAS analysis in revealing the genetic basis of AD by connecting brain imaging, genomics, and clinical outcome with one another.

3. Results

Figure 2 shows 19 SNPs identified with association with AD for 14 phenotypes in the longitudinal GWAS analysis with p value lower than 10^-6 between baseline and month 12. The most significant values in figure 2 is associated with the following genes: TOMM40, APOE, and APOC1. Figure 3 shows 32 SNPs identified with association with AD for cognitive measurements as the phenotypes with p value lower than 10^-6 at baseline in the multivariate GWAS analysis. The most represented genes are AGPAT4, KRAS (left gene), and FBXO28 (left gene) . The two figures are examples of heat maps generated for different timepoints in the longitudinal GWAS analysis and multivariate GWAS analysis.

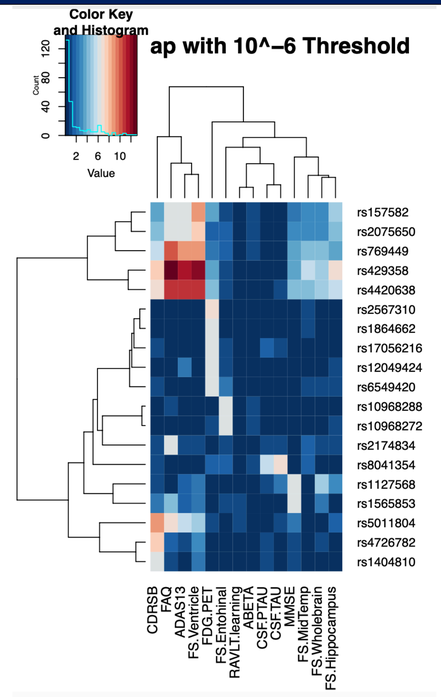

Figure 2. Longitudinal GWAS result for month 12. X axis represents phenotypes, y axis represents genes.

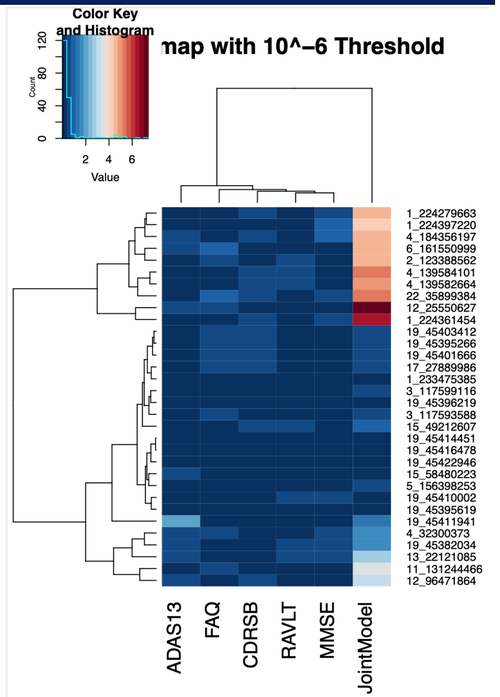

Figure 3. MultiPhen result for baseline. X axis represents cognitive measurements, y axis represents genes.

4. Conclusions

This project bridges genetic data with clinical diagnosis using intermediate phenotype data (16 quantitative traits). Multiple genes, with APOE4 on chromosome 19 being the most significant, were identified as potential AD targets among 16 phenotypes for the longitudinal GWAS study and the multivariate GWAS study. These identified genes warrant future implication as potential targets for potential diagnosis, prevention, and cure for AD.

References

[1]Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007)
[2]Lachlan Coin, Paul O'Reilly, Yotsawat Pompyen, Clive Hoggart and Federico Calboli (2020). MultiPhen: A Package to Test for Multi-Trait Association. R package version 2.0.3.

ACKNOWLEDGEMENTS: This work was supported by Penn Undergraduate Research Mentoring Program (PURM), the Shen Lab at the Perelman School of Medicine at the University of Pennsylvania, and funding from NIH R01 EB022574, R01 LM011360, U19 AG024904, R01 AG019771, and P30 AG010133; and NSF IIS 1837964. The complete ADNI Acknowledgement is available at http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.