Arkesh Das
CMSE 410
Advisor: Dr. Jianrong Wang

Assessing the Influence of False Discovery Rate Methods on Genetic Associations in an Immune Response GWAS (Project Pre-proposal)

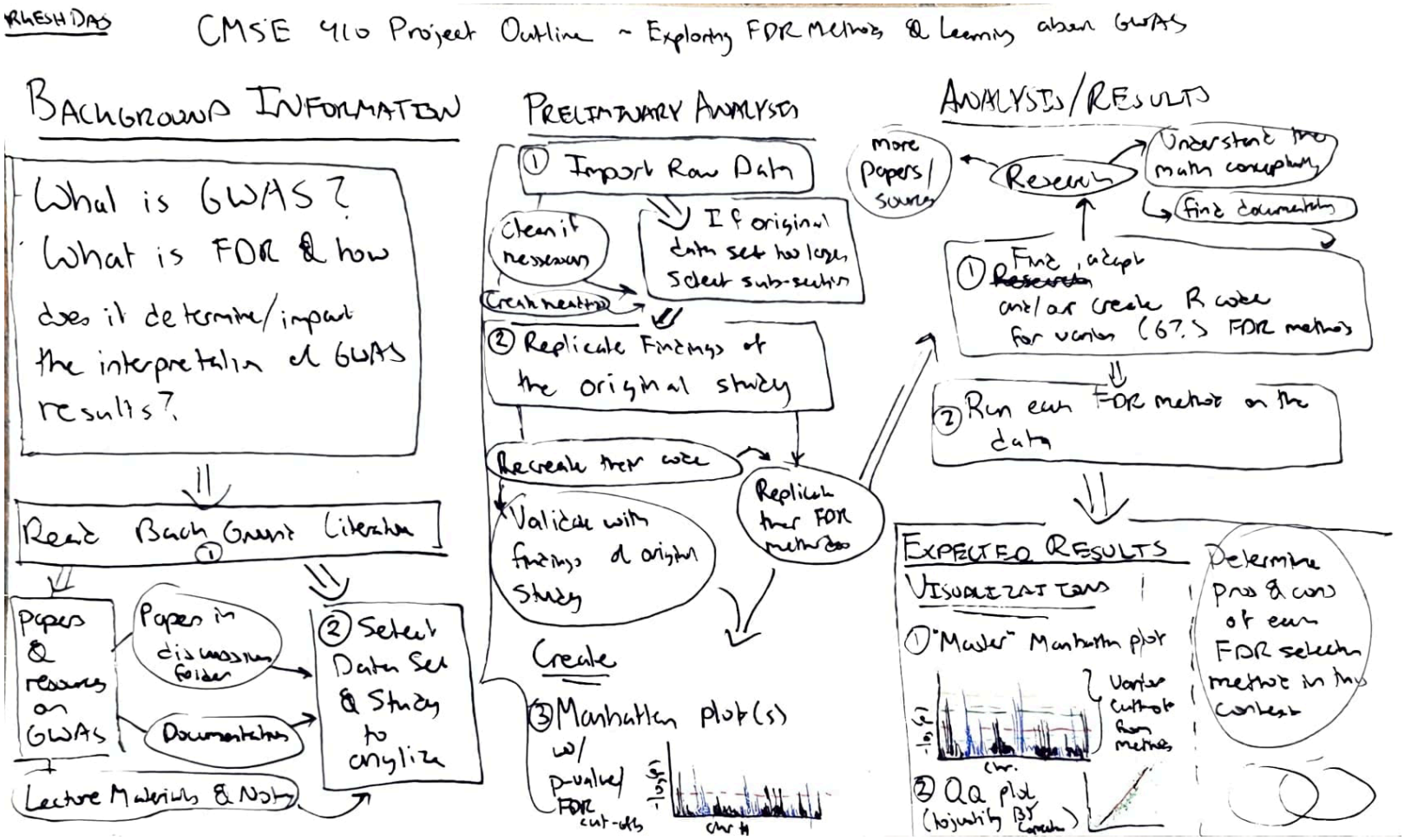
When analyzing the results of a genome-wide association study (GWAS), it is important to determine an appropriate p-value threshold for identifying statistically significant genetic associations. Because a GWAS may result in hundreds of thousands, if not millions of statistical tests, a conventional p-value cut-off would result in a large number of false positives. Therefore, determining the appropriate way to calculate and control the false discovery rate (FDR) is essential. Establishing an appropriate FDR in the context of a particular GWAS helps balance the risk of type I and type II errors. incorrectly declaring a non-associated variant as significant—with the need to detect genuine associations, ultimately ensuring that findings are both credible and reproducible.

To explore this problem, I found a GWAS study that was conducted by researchers as a part of the Milieu Intérieur project, which seeks to determine the genetic factors responsible for immune function and variation in humans (Scepanovic et al. 2018). The researchers in this study used the Benjamini-Hochberg (BH) procedure to control the FDR in their GWAS analysis. While the BH method helps reduce the incidence of false discoveries by adjusting p-value thresholds across the multitude of tests, it may fall short when accounting for the dependence structure among the genetic variants. Because many genetic markers are correlated due to linkage disequilibrium, BH may either be too conservative (resulting in a loss of true associations) or too liberal (leading to an excess of false positives). This trade-off represents a significant limitation in current methodology employed by the researchers in this study.

I hope to replicate the findings and analysis conducted by this research group by importing and cleaning their data and creating my own code in R. Once I have done this, I will research various other FDR methods, and try to implement them as well. Some methods that I have already decided to take a look at are the use of a simple p-value cut-off, the use of the Bonferroni Correction, the Benjamini–Yekutieli (BY) Procedure, and the Storey & Tibshirani (q-value) procedure). I will then create various visual representations, such as manhattan plots to determine the success and drawbacks of each method.

My main goals with this project are to learn more about GWAS and FDR. I hope that by exploring these different statistical methods, their strengths and weaknesses, as well as how to implement them, I will gain a greater understanding of them. In the context of this study, a better FDR would theoretically strike a better balance between type I and type II errors and would lead to more reliable identification of genetic variants linked with diseases of the immune system. These improvements could accelerate the discovery of new biomarkers to predict an individual's immune responses to different conditions or haplotype-specific therapeutic targets.

Graphical Abstract/Flow-chart of Project



References

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-wide association studies. *PLoS Computational Biology, 8*(12), e1002822. https://doi.org/10.1371/journal.pcbi.1002822

Scepanovic, P., Alanio, C., Hammer, C., et al. (2018). Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. *Genome Medicine, 10*, 59. https://doi.org/10.1186/s13073-018-0568-8

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America, 100*(16), 9440–9445. https://doi.org/10.1073/pnas.1530509100

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: Biology, function, and translation. *American Journal of Human Genetics, 101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005