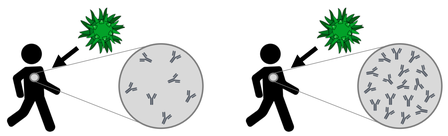
Genome-Wide Association Study of Immune Response to Twelve Common Pathogens

Maike Morrison[1], Dr. Linda Kachuri[2], Dr. Sara Rashkin[2], Dr. John Witte[2]
maike.morrison@utexas.edu
[1]Department of Mathematics, The University of Texas at Austin
[2]Department of Biostatistics and Epidemiology, The University of California San Francisco

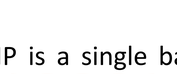
Background

- Immune response strength (amount of antibody produced in response to infection) varies between individuals (Figure 2).
- Understanding the genetics of variability in immune response strength can yield insights into pathogen-induced disease etiology, potential drug targets, and the biological mechanisms of immune response.



What is a GWAS?

In a Genome-Wide Association Study (GWAS), we fit linear regression models to identify the single-nucleotide polymorphisms (SNPs) most strongly associated with immune response strength (Figure 1).

A SNP is a single base pair, like the blue rectangles here:  At this SNP, an individual can have 0, 1, or 2 copies of the dark blue allele (■).

Our linear regression models use this information (G, for genotype) to predict y, the strength of immune response, in an equation such as this one:

$$y = \beta_0 + \beta_1 \times (\mathrm{G} \in \{0, 1, 2\}) + \beta_{\mathrm{Age}} \times (\mathrm{Age}) + \cdots + \epsilon$$

We include covariates such as age to account for variation in y due to non-genetic factors.

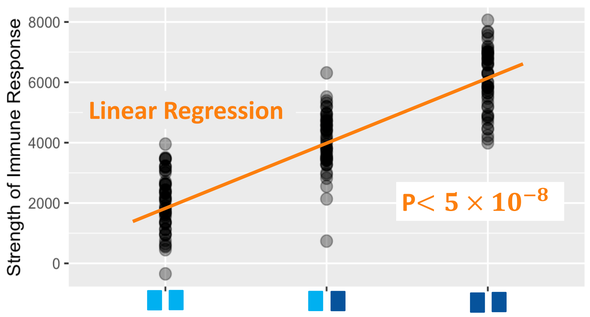


Figure 1. Simulated data for illustration of GWAS method. Each point represents an individual; x-axis is their genotype at a specific SNP; y-axis is their immune response strength. Here, P represents the statistical significance of the slope, β_1. 5×10^{-8} is the genome-wide significance threshold.

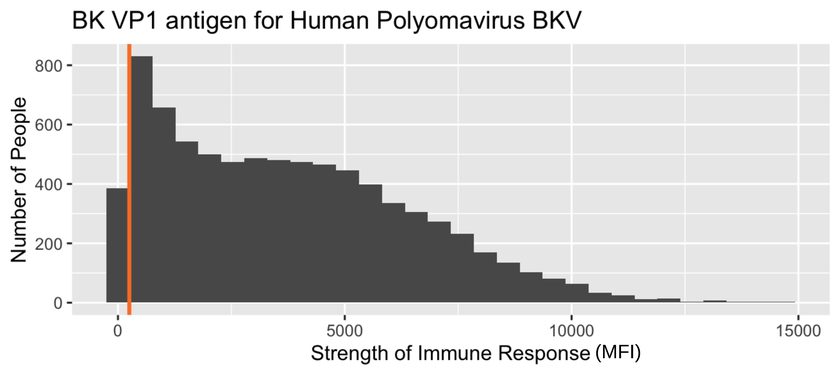


Figure 2. Distribution of immune response to Human Polyomavirus BKV (PBKV) in 7,944 adults of European ancestry. Immune response strength is measured using Multiplex Serology. Higher median fluorescent intensity (MFI) indicates higher antibody titer in an individual's serum and stronger immune response. If immune response strength exceeds 250 MFI (orange line), an individual is considered PBKV seropositive and is inferred to be infected with PBKV. 95.2% of the sample is PBKV seropositive.

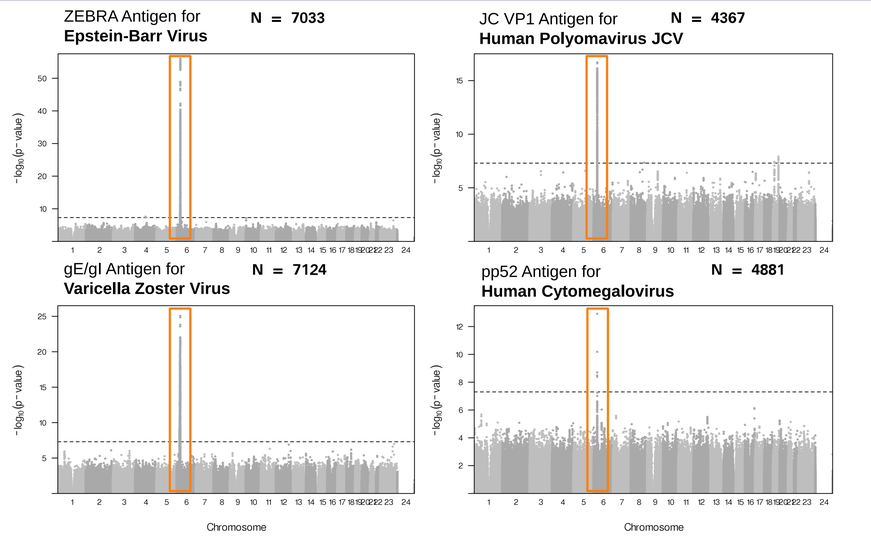


Figure 3. Sample of four Manhattan plots showing associations between immune response strength and SNPs in the Major Histocompatibility Complex (MHC). The MHC (peaks within orange boxes) is a set of genes encoding proteins that enable the immune system to distinguish self from non-self. These genes have well-established associations with immune response.

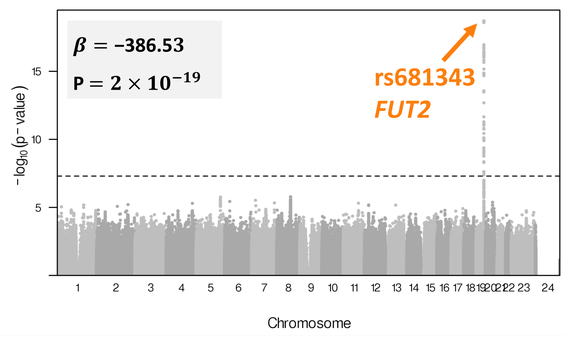
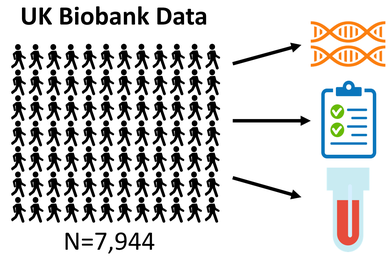


Figure 4. Manhattan plot of SNPs associated with immune response to Human Polyomavirus BKV (PBKV). This analysis used a sample of 7,349 individuals infected with PBKV. The lead SNP, rs681343, is a nonsense mutation in the gene *FUT2*. The *FUT2* gene creates an enzyme required for the soluble A & B antigen synthesis pathway (Figure 5). Our lead SNP is in perfect linkage disequilibrium with a common nonsense mutation which results in a nonfunctional enzyme.

Acknowledgements: Thank you to the Witte Lab and the UCSF SRTP program for guidance and support, and to the Amgen Foundation for funding.

References:
1. Ovsyannikova, I. G., Kennedy, R. B., O'Byrne, M., Jacobson, R. M., Pankratz, V. S., & Poland, G. A. (2012). Genome-wide association study of antibody response to smallpox vaccine. *Vaccine*, *30*(28), 4182-4189.
2. Li, Y. R., Li, J., Zhao, S. D., Bradfield, J. P., Mentch, F. D., Maggadottir, S. M., ... & Guo, Y. (2015). Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nature Medicine*, *21*(9), 1018.
3. Le Pendu, J., Ruvoën-Clouet, N., Kindberg, E., & Svensson, L. (2006). Mendelian resistance to human norovirus infections. In *Seminars in Immunology* (Vol. 18, No. 6, pp. 375-386). Academic Press.

Data & Methods



- The 22 antigens we analyzed target 12 viral, bacterial, and parasitic pathogens that are established risk factors for cancer and cardiovascular or neurodegenerative diseases.

Results

- Found associated SNPs for **17 antigens** against **10 pathogens**
- **99** independent ($r^2 < 0.05$), significant ($P < 5 \times 10^{-8}$) SNPs; **47** in genes
- Many of these genes associated with autoimmunity and vaccine response[1,2]
- Across all 22 antigens there were **45** independent, significant SNPs in the MHC (Figure 3)
- Identified a SNP in *FUT2* which suggests that secreted histo-blood group antigens could play a role in polyomavirus immune response (Figures 4, 5)

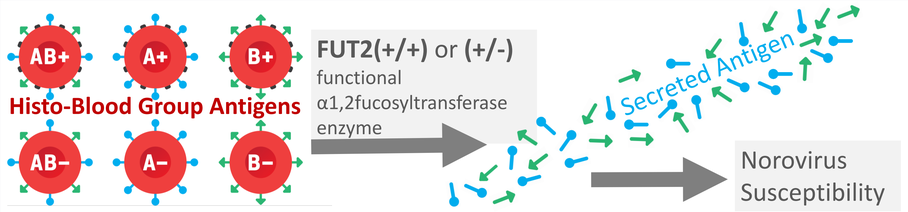


Figure 5. The soluble A&B antigen synthesis pathway, which is dependent on the FUT2 gene product α1,2fucosyltransferase. FUT2 has two Mendelian alleles: a dominant, functional enzyme (+), and a recessive, nonfunctional enzyme (-). FUT2(-/-) individuals (who have two nonfunctional gene copies) represent 20% of Europeans and are called non-secretors. This is because, without a functional enzyme, they are unable to secrete A and B blood group antigens into bodily fluids such as saliva and serum. Recent work has shown that non-secretors are highly resistant to norovirus infection[3].

Conclusions & Future Directions

- Different genes are associated with different antibodies.
- Some of these genes could elucidate the biological mechanisms underlying immune system interaction with specific pathogens.
- Future work could explore the genetic basis of pathogen-induced cancers & immune response, the biological function of immune-response-associated genes, and personalized prevention (vaccines) & intervention (drug targets) strategies.