Geography and Environment Predict Immune Allele Frequencies in Infectious Disease Burdened Populations

Latifa Jackson*, Max Shestov*, Aydin Tozeren
Drexel University, School of Biomedical Engineering, Science and Health Systems

*These authors contributed equally
LJ: Ljf27@drexel.edu
MS: ms3389@drexel.edu
AT: at62@drexel.edu

Introduction
Human immunological interactions with their environment are the substrate for natural selection. One approach to studying natural selection in humans has been to examine single genes in a population to directly assess selection caused by some environmental effect (i.e. HBB and malaria, SLC24A5 and UV exposure) Another strategy we build upon has demonstrated success in identifying variation in allele frequencies between populations taking into account genetics and environmental factors. The interplay of disease burden and environmental features are compelling external forces that impact individual survival.

Methods
To test whether there are correlative relationships between disease burden, environmental conditions and genetic variation in population allele frequencies, we used genes contained in immune gene curated by NCBI Gene. Immune associated genes were then projected onto the genome to identify cluster regions of genetic importance for immunity. We surveyed human polymorphism data from 75 populations from the Human Genome Diversity Panel (HGDP), HapMap Project and the 1000 Genomes datasets. Disease prevalence and pathogen load data for were used to characterize the epidemiological environments. We finally mapped this information to a global scale map using geographical information systems implementation in ArcGIS.

Results
We find evidence of three genomic regions that have a statistical overabundance of immune genes situated in close proximity to each other. These were located on chromosomes 9, 11 and 19. We have identified population signatures showing that there is a correlation between geographic location and allele frequencies for adaptive immune mechanisms.

Conclusions
We believe this method allows us to distinguish those immune polymorphisms that vary in human populations and are greatly impacted by environmental factors. These findings suggest that we can use the component parts of our model to predict those polymorphisms that may not significantly vary in human populations, but are significantly correlated to disease burden and environment.