Danny Crichton

Essay #3: GWAS

12 Feb. 2009

Prof. Russ Altman

The Med School SNP: rs17070145 and Better Human Memory Performance

Scientists have exerted significant effort developing an understanding of human memory using patients like the amnesiac H.M. Today, the subjects have expanded to human genes. Recently, scientists using a Genome Wide Association Study (GWAS) believe they have found a SNP that is correlated with better memory recall. Led by Andreas Papassotiropoulos and Dietrich A. Stephan, their team used a combination of psychological tests and genotyping to detect SNPs for further study. Reporting their findings in *Science*, they a found significant correlation between rs17070145, a T/C SNP, and memory performance. The SNP is located in the *KIBRA* gene – which they experimentally link with memory (Papassotiropoulos 477-478). Interestingly, the team used GWAS to identify possible genes, and then performed a variety of statistical correlations and experiments to prove a causal link, avoiding some of the statistical issues generally associated with these studies. In evaluating their report, this paper will look at their applicant pool, study design, statistics and confidence builders.

Scientists must carefully design a GWAS study to avoid statistical bias. Due to genetic recombination, sample pools need to be evaluated to ensure that nearby SNPs are not in linkage disequilibrium. Unlike clinical studies, the best samples for GWAS studies are very homogenous – a design carefully followed in the Papassotiropoulos study. Their initial cohort was composed of 351 young adults from Switzerland (475). This sample size

is relatively low for a GWAS study – a number I will analyze later. They then evaluated each subject and found ten who were genetically distant from the rest of the cohort; these people were removed from the study. With a scientifically homogenous pool, this study began with a solid foundation to perform a GWAS.

The study design is relatively straightforward. The team tested the sample pool with a standard memory test – looking at a sequence of nouns and then attempting to recall them 5 minutes and 24 hours later (Supporting 1). After these tests were complete, the scientists then split the sample into four separate groups based on memory performance and genotyped them. From there, they correlated better memory with two SNPs – rs17070145 and rs6439886 (Papassotiropoulos 475). Finally, they used experimental techniques to find causal links between these two SNPs and memory performance.

This overview raises an important question – was the separation of groups the best way to detect SNPs? When phenotypic traits are highly polygenic, as in this study, researchers generally split the group into low performers and high performers, cutting average subjects entirely. This decision, combined with a higher threshold for candidate SNPs, may have identified more likely SNPs, but it may also have missed important other possibilities. Especially since this paper focuses on two specific SNPs in its experiments, further research is necessary to detect other possibilities.

Finding target SNPs, Papassotiropoulos and his team regenotyped all their subjects and correlated the two SNPs with memory performance. The strongest association was between rs17070145 and recall after 5 minutes. This association had a p-value of $4x10^{-6}$ (Papassotiropoulos 475). There were also associations between this SNP and the other memory tests, although with lower significance ($8x10^{-4}$). The other SNP of interest,

rs6439886, showed lower levels of statistical significance (and was eventually cut during replication). These numbers bring up one of the most challenging aspects of GWAS studies – the high number of samples needed to gain the discriminating power to prove a causal link. In this case, a link cannot be made because the p value is not low enough to counter the 500,000 possible SNPs on the study's Affymetrix 500K chip (Supporting 2). The authors of this study avoided this problem with their design. Their claim with the GWAS data is not that these SNPs cause memory performance differences, but instead used the data to detect candidate SNPs and then investigated those more thoroughly. As such, they are not even using the correlation as causal evidence, but rather using it as a first step.

An important component of GWAS studies are its confidence builders – replication of the study and further experiments. In this study, the authors used both techniques to tightly link rs17070145 with memory. Papassotiropoulos replicated his results with two additional cohorts: a similarly aged one from America and a second one from Switzerland (Supporting 1). These studies showed roughly similar correlations between rs17070145 and memory recall (Papassotiropoulos 475). Importantly, the study did not perform a GWAS study on these two cohorts, but rather searched for a link between rs17070145 and memory. As discussed previously, there may be additional SNPs missed by this study.

In addition to replication, the scientists compared the SNP to the gene it is located on – *KIBRA*. They noted that the SNP was located in the coding region, specifically in the truncated part of the gene that forms *KIAA0869*, which is highly prevalent in memory regions of the brain (Papassotiropoulos 477). Next, they looked at the possibility of whether some other SNP was causing the differences in phenotype. They discovered that the SNP was in a haplotype block, and that this SNP is in the causal gene (although this SNP

may not be the exact causal variant – any area for more research). Finally, the scientists performed fMRI scans on a small subsample of their cohort and found that those homozygous for C had more brain activity to answer the memory test questions, indicating they used more brain power and had less efficiency (Papassotiropoulos 478). Altogether, the study uses a variety of replications and experimental techniques to verify their claims.

Papassotiropoulos used a strong study design to prove his link between rs17070145 and memory performance. The GWAS study itself would not have been sufficient to establish a causal claim, but the addition of two cohorts, a haplotype analysis and fMRI experiments add significant and credible evidence to his argument. Nonetheless, significant areas of research remain to be studied to conclusively prove this link. Additional work is needed to identify why this SNP creates such memory performance differences, whether there are other memory-related SNPs in *KIBRA* or other genes, and which human populations have this variant.

Works Cited

Common Kibra Alleles Are Associated with Human Memory Performance. Andreas Papassotiropoulos, Dietrich A. Stephan, et al. Science 314 (5798), 475. [DOI: 10.1126/science.1129837]

Supporting Online Material for Common *Kibra* Alleles Are Associated with Human Memory Performance. Andreas Papassotiropoulos, Dietrich A. Stephan, et al. http://www.sciencemag.org/cgi/data/314/5798/475/DC1/1