Pathway analysis of genome-wide association study (GWAS) for myopia in a Chinese population

Patrick Y.P. Kao,1 Maurice K.H. Yap,1 Kim Hing Leung,2 Chia-Chuen Khor,3 Seang Mei Saw,4 and Shea Ping Yip2

1Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University, Hong Kong SAR, China;
2Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China;
3Division of Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore;
4Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore.

Purpose
Myopia is a common disease of human eye. The etiology of myopia is complex. Genetic and environmental factors are both important risk factors. A number of genes of myopia have previously been identified, providing evidence of the complex nature of the disease. However, common diseases, such as myopia, are manifested by a larger number of genes each of which only confers very modest effect. In an effort to look for novel genetic risk factors that have small effects, we propose to carry out a pathway analysis. The analysis aims for searching modestly significant genes within common pathways which together may be related according to interaction or literature knowledge.

Methods
We have recruited cases and controls from Hong Kong (HK) and Singapore (SICC) respectively. Samples were genotyped with Illumina Human 610 Quad BeadChips. Quality checking of the samples and SNPs, and checking for global and local population stratification were carried out. After QC, 2,075 samples were included for analysis, including 94 HK cases, 95 HK controls and 1,886 SICC controls (totaling 94 cases and 1,981 controls). Batch effects were corrected using Eigenstrat. Imputation of data was then carried out using 1000 Genomes Phase I integrated haplotypes (Dec 2013) as the reference population. Association analysis was carried out using SNPTTEST.

Pathway analysis of genes having significant SNPs was carried out using MetaCore.
P-values of hypergeometric distributions of gene sets in each reported network (representing that the probability of selecting the genes is due to chance) were used as criteria for selection of networks. Tagging SNP were selected from the genes in the most significant networks for follow-up using Tagger.

Results
After association analysis of imputed data, there were 55,956 and 28,138 SNPs having $p \leq 0.01$ and $\leq 0.005$ respectively. The two gene lists were imported into Metacore for pathway analysis. The SNP list with $p \leq 0.01$ and $\leq 0.005$ respectively were mapped onto 2,135 and 1,274 genes. Top 3 networks with the smallest hypergeometric p-values were extracted for both datasets, which resulted in a gene set of 142 genes. Of these genes, 110 were found to have expression in eye according to information in the Unigene database. To follow-up the 110 genes, HapMap Chinese (CHB) tagging SNPs for the genes and significant SNPs were extracted, which accounted for 8,247 SNPs. These SNPs will be selected for genotyping with an independent sample set of a larger size to look for and confirm the exact associated genes causing myopia.

Conclusion
A pathway analysis of genome-wide association study (GWAS) for Chinese population was carried out. In total, 110 genes with eye expression were suggested to be among the most interesting genes represented by SNPs with small p-values. Follow-up genotyping will be carried out to explore which genes are responsible for etiology of myopia.

Acknowledgements:
This work was supported by a General Research Fund (PolyU 5637/12M) from the Research Grants Council of Hong Kong and a grant (87TP) from the Hong Kong Polytechnic University. Maurice Yap was also supported by the Endowed Professorship Scheme (KB Woo Family Endowed Professorship in Optometry) of the Hong Kong Polytechnic University.