

A SNP Ranking Method Using Gene Relation Network Derived from Literature for Genome-Wide Association Study

Jaeyoung Kim¹, Miyoung Shin^{2,*} and Munpyo Hong³

¹Graduate School of Electrical Engineering and Computer Science, Kyungpook National University, Daegu, South Korea

²College of IT Engineering, Kyungpook National University, Daegu, South Korea

³College of Liberal Arts, SungKyunkwan University, Seoul, South Korea

ABSTRACT

Human diseases are often incurred by some variants of gene interactions caused by such reasons like environment factors or gene mutations. Recently, to identify the loci and the traits of disease marker genes, a large volume of SNP genotype data obtained from SNP chip experiments are being investigated. In particular, the task of prioritizing SNPs is of great significance to find SNP markers associated with certain diseases. For this purpose, many of existing approaches are calculating ranking statistics of SNPs only based on the difference of their genotypes between normal and patient samples. However, there exist other types of biological resources worthwhile to use for such tasks. Bio-literature is one of them that includes many valuable results about gene functions, gene-gene interactions, gene-disease relationships and etc. In this work, we examine a SNP ranking method which incorporates gene relation networks derived from literature, along with SNP genotype data, into the ranking process. For evaluation, we applied this method to identify SNP markers associated with prostate cancers and it was confirmed that the use of gene relation networks for prioritizing SNPs can lead to the better findings of disease related SNPs than when they are not used.

Keywords: Genome-Wide Association Study, Gene relation network, Bio-literature, Disease markers, SNP prioritization.

ACKNOWLEDGEMENTS

This work was supported by the Korea Research Foundation of Korea (KRF) grant funded by the Korea government (MEST) (No. 2009-0067724).

REFERENCES

1. B. Maher, "Personal genomes: The case of the missing heritability." *Nature*, vol. 456, pp. 18-21, November 2008.
2. C. M. Lewis, "Genetic association studies: design, analysis and interpretation," *Brief Bioinform.*, vol. 3(2), pp. 146-53, Jun 2002.
3. J. L. Morrison et al., "GeneRank: using search engine technology for the analysis of microarray experiments," *BMC Bioinformatics*, vol 6:233, doi:10.1186/1471-2105-6-233, 2005.
4. D. Cheng et al., "PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites," *Nucleic Acids Research*, vol. 36, pp. 399-405, 2008.

* To whom correspondence should be addressed

Email: shinmy@knu.ac.kr