Bayesian hierarchical model for joint estimation of SNP effects with integration of prior biological knowledge

Conference Paper in Genetic Epidemiology · November 2015
Impact Factor: 2.6

4 authors, including:

Miguel Pereira
Imperial College London
4 PUBLICATIONS 18 CITATIONS

SEE PROFILE

Christian X Weichenberger
Europaeische Akademie Bozen - Accademia …
32 PUBLICATIONS 373 CITATIONS

SEE PROFILE

Available from: Miguel Pereira
Retrieved on: 10 July 2016
Title: Bayesian hierarchical model for joint estimation of SNP effects with integration of prior biological knowledge

Miguel Pereira¹, John R. Thompson², Christian X. Weichenberger³, Cosetta Minelli¹

¹National Heart and Lung Institute, Imperial College London, UK
²University of Leicester, Leicester, UK
³EURAC, Bolzano, Italy

Abstract

Genome-Wide Association Studies are classically analyzed by estimating SNP effects individually and adjusting for multiple testing. However, SNPs identified so far explain a small proportion of the predicted heritability for most traits. The joint analysis of SNPs has been suggested to improve SNP detection in GWAS, but its implementation is limited by computational efficiency. Another way to improve SNP detection is through integration of prior information that gives more weight to SNPs supported by biological evidence, which can be achieved within the Bayesian framework.

By extending the computationally efficient approach proposed by Yi et al. (PLoS Genet 2011), we used a Bayesian hierarchical model to exploit SNP correlation structure, defined by linkage disequilibrium D’, and incorporate prior biological knowledge. SNPs in strong LD were grouped in LD blocks, and prior knowledge was retrieved with a bioinformatic tool (DINTOR). We piloted the model by testing the association of BMI with 2,614 SNPs within 30 LD blocks, of which 6 contained one “true” SNP, in 1,829 individuals. SNP effects were estimated both with and without LD block structure and prior knowledge.

The Bayesian joint estimation of SNP effects ranked the true SNPs and true LD blocks higher than the classical approach. Adding the LD structure to the model improved the ranking of true LD blocks, which was further improved by integration of prior knowledge, with 5 of the true LD blocks ranking in the top 6.

These preliminary results suggest that both LD block structure and integration of prior knowledge can improve SNP detection. We are currently working on: 1) improving the incorporation of prior knowledge; 2) using LD r² instead of D’; 3) scaling up the analysis to genome-wide level.