BIOINFORMATICS LETTER TO THE EDITOR Vol. 28 no. 1 2012, pages 145–146
doi:10.1093/bioinformatics/btr596

Genome analysis Advance Access publication October 28, 2011

Comments on 'An empirical comparison of several recent epistatic interaction detection methods'

Xiang Wan†, Can Yang† and Weichuan Yu*

Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China

Associate Editor: Jonathan Wren

Contact: eeyu@ust.hk

Received on September 20, 2011; revised on October 17, 2011; accepted on October 22, 2011

1 BACKGROUND

The detection of epistasis between different loci helps to provide in-depth understanding of complex diseases. Many methods have been proposed to find epistasis in genome-wide association studies (GWAS). A recent article (Wang *et al.*, 2011) published in *Bioinformatics* conducted an empirical comparison on five methods [TEAM, BOOST, SNPHarvester, SNPRuler, and Screen and Clean (SC)] in terms of power, type-1 error rate, scalability and completeness. This work will help users to better understand the pros and cons of each method. However, the lack of epistasis definitions may bring some confusions for users to select the appropriate method for their tasks because the tests of epistasis in these methods belong to two different categories, while the comparison is based on the same experimental setting. In this letter, we like to clarify the difference of two epistasis tests.

2 EPISTASIS

2.1 Definition

The concept of epistasis (Bateson and Mendel, 1909) was introduced around 100 years ago. It is generally defined as interactions among different genes. Recently, Phillips (2008) summarized three categories of epistasis with respect to the essential role of gene–gene interactions in the structure and evolution of genetic systems.

- Functional epistasis is a functional description that addresses the molecular interactions.
- Compositional epistasis (Bateson and Mendel, 1909) is referred as the blocking of one allelic effect by another allele at a different locus.
- Statistical epistasis (Fisher, 1918) is defined as the statistical deviation from the additive effects of two loci on the phenotype.

The definition of the statistical epistasis provides a quantitative way to measure how the phenotypic effect of one locus depends on the genotype at another locus (Phillips, 2008). This facilitates the mathematical analysis of epistasis.

2.2 Tests of statistical epistasis

Given a pair of SNPs, there are two major tests of epistasis in the literature. One is referred to as the 'two-locus interaction test', the other 'two-locus association test allowing for interaction' (Cordell, 2009). In the following, we provide the procedures of these two tests in case–control studies based on logistic regression. Here we use Y to denote the two-locus interaction disease status ($Y=1$ for cases and $Y=2$ for controls).

2.2.1 Two-locus interaction test Given two loci X_p and X_q, there are three steps to test two-locus interaction:

- Fit the logistic regression model with only main effects. The corresponding model is called main effect model:
$$\log\frac{P(Y=1|X_p=i,X_q=j)}{P(Y=2|X_p=i,X_q=j)}=\beta_0+\beta_i^{X_p}+\beta_j^{X_q}. \quad (1)$$
- Fit the logistic regression model with both main effect terms and interaction terms. The corresponding model is named full model:
$$\log\frac{P(Y=1|X_p=i,X_q=j)}{P(Y=2|X_p=i,X_q=j)}=\beta_0+\beta_i^{X_p}+\beta_j^{X_q}+\beta_{ij}^{X_pX_q}. \quad (2)$$
Please note that the superscript X_p of $\beta_i^{X_p}$ in both equations is merely a label and does not represent the exponent. The term $\beta_i^{X_p}$ represents the coefficient of X_p at category i. This representation extends to $\beta_j^{X_q}$ and $\beta_{ij}^{X_pX_q}$ as well. Here $i=0,1,2$ and $j=0,1,2$ represent the three genotypes of loci X_p and X_q, respectively.
- Let $\hat{L}_M$ and $\hat{L}_F$ denote the log-likelihoods of the main effect model and the full model, respectively. Conduct the χ^2 test on $2(\hat{L}_F-\hat{L}_M)$ with the degrees of freedom $df=4$.

This procedure completely matches the definition of statistical epistasis from Fisher (1918). In practice, computing $\hat{L}_F$ of the full model is very efficient, but computing $\hat{L}_F$ of the main effect model involves Newton iterations (Wan *et al.*, 2010a). Since a typical GWAS contains hundreds of billions of SNP pairs, it is computationally challenging to test two-locus interactions in GWAS. As a result, many methods choose not to fit the main effect model and

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First authors.

instead use a much simpler procedure to test two-locus associations allowing for interactions.

2.2.2 Two-locus association test allowing for interaction Given two loci X_p and X_q, there are also three steps to test two-locus associations allowing for interactions:

- Compute the maximum log-likelihoods $\hat{L}_{\emptyset}$ of the null logistic regression model,

$$\log \frac{P(Y=1)}{P(Y=2)} = \beta_0. \quad (3)$$

- Compute the maximum log-likelihoods $\hat{L}_F$ of the full logistic regression model in Equation (2).
- Conduct the χ^2 test on $\hat{L}_F - \hat{L}_{\emptyset}$ with $df = 8$.

The two-locus association test allowing for interaction is based on 8 degrees of freedom and measures the sum of individual effects and interaction effect.

3 COMMENTS ON RESULTS

3.1 The experiment of testing 'power'

Among the methods compared in Wang *et al.* (2011), BOOST (Wan *et al.*, 2010a) and the Screen and Clean method (Wu *et al.*, 2010) test two-locus interactions, whereas TEAM (Zhang *et al.*, 2010), SNPHarvester (Yang *et al.*, 2009) and SNPRuler (Wan *et al.*, 2010b) measure two-locus associations allowing for interactions. In the comparison, the authors observed that BOOST has a lower power in the experiment using models with main effects. This observation could be misleading as BOOST only measures the interaction effect. In fact, the interaction effect does not play a dominant role in those main effect models, which makes it more difficult to detect two-locus interactions than two-locus associations allowing for interactions. It is worthy pointing out that Model 1 in Wang *et al.* (2011) is a multiplicative model which is an additive model in the log scale. Therefore, there is no interaction effect in this model. Besides, there are some critical issues in the test of two-locus associations allowing for interactions, which need to be carefully addressed.

- The size of two locus associations allowing for interactions could be very large. Suppose one SNP has a very strong main effect and does not interact with any other SNPs. The test statistic of this SNP with any other SNP in the test of two-locus associations allowing for interactions can be significant due to its strong main effect. This leads to the high redundance.
- In order to extract SNP pairs with significant interaction effects from two-locus associations allowing for interactions, the post-processing is needed. Here single-locus effects should be handled appropriately.

BEAM (Zhang and Liu, 2007) is a very successful method to address these issues. BEAM classifies SNP markers into three categories: SNPs unassociated with the disease, SNPs associated with the disease independently and SNPs jointly associated with the disease (interaction). All three categories of SNPs are integrated into a Bayesian model. The Markov Chain Monte Carlo method is used to optimize the partition model. After that, the posterior probability of each SNP belonging to which category can be computed. The B-statistic and conditional B-statistic are developed to test the significance of associations allowing for interactions and the significance of interactions.

3.2 The experiment of testing 'Completeness'

In the experiment of testing 'Completeness', the authors claimed that BOOST made some wrong pruning. This claim can cause some confusion. BOOST is a two-stage method. In the first stage, BOOST designs an upper bound for the likelihood ratio test statistic to screen insignificant interactions. In the second stage, it uses the typical procedure of likelihood ratio test to measure the significance of selected pairs. The validity of the upper bound has been theoretically proven. BOOST checks every SNP pair against the upper bound in stage one. Thus, it is impossible to miss any significant interaction in BOOST. A possible explanation for this phenomenon may be that BOOST pruned SNP pairs based on the two-locus interaction test, while the authors used some other criteria to make their conclusion.

3.3 The experiment of testing 'Type-1 error'

In the experiment for 'Type-1 error', the authors showed that the type-1 error rate of BOOST is 0.065, which exceeds the nominal error rate 0.05. As BOOST uses the standard likelihood ratio test in stage two, the type-1 error rate of BOOST should agree with the nominal level of the standard statistical test.

Funding: RPC10EG04 from the Hong Kong University of Science and Technology.

Conflict of Interest: none declared.

REFERENCES

Bateson,W. and Mendel,G. (1909) *Mendel's Principles of Heredity*. Cambridge University Press, Cambridge.

Cordell,H. (2009) Detecting gene-gene interactions that underlie human diseases. *Nat. Rev. Genet.*, **10**, 392–404.

Fisher,R.A. (1918) The correlations between relatives on the supposition of mendelian inheritance. *Philos. Trans. R. Soc. Edinburgh*, **52**, 399–433.

Phillips,P.C. (2008) Epistasis-the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.*, **9**, 855–867.

Wan,X. *et al.* (2010a) BOOST: A fast approach to detecting gene-gene interactions in genome-wide case-control studies. *Am. J. Hum. Genet.*, **87**, 325–340.

Wan,X. *et al.* (2010b) SNPRuler: Predictive rule inference for epistatic interaction detection in genome-wide association studies. *Bioinformatics*, **26**, 30–37.

Wang,Y. *et al.* (2011) An empirical comparison of several recent epistatic interaction detection methods. *Bioinformatics*, [Epub ahead of print, doi:10.1093].

Wu,J. *et al.* (2010) Screen and clean: a tool for identifying interactions in genome-wide association studies. *Genetic Epidemiol.*, **34**, 275–285.

Yang,C. *et al.* (2009) SNPHarvester: a filtering-based approach for detecting epistatic interactions in genome-wide association studies. *Bioinformatics*, **25**, 504–511.

Zhang,X. *et al.* (2010) Team: efficient two-locus epistasis tests in human genome-wide association study. *Bioinformatics*, **26**, i217.

Zhang,Y. and Liu,J. (2007) Bayesian inference of epistatic interactions in case-control studies. *Nat. Genet.*, **39**, 1167–1173.