Reviewer's report

Title: Candidate gene association study in pediatric acute lymphoblastic leukemia evaluated by Bayesian network based Bayesian multilevel analysis of relevance

Version: 1 Date: 29 May 2012

Reviewer: Ignacio Varela

Reviewer's report:

Dr. Lautner-Csorba et al. describe in the present manuscript a new method of analysis of gene association studies, based in Bayesian networks, that improves the traditional frequentist-based statistical methods. This new system offers a solution for the multiple hypothesis problem inherent to frequentist-based methods and provides of tools to detect new levels of information in the variables relationship allowing the discrimination of direct, indirect, strong or weak interactions among them. When the manuscript is considered of potential interest for the field and offers possibilities in the extraction of new knowledge from already generated data combined with meta-analysis of background knowledge; some aspects in the work presented in the manuscript would need to be clarified and addressed before acceptance can be advised from this referee.

Major Compulsory Revisions:

The authors used blood samples extracted from Childhood ALL patients to test the new method developed. As the study is dealing with patients of a type of blood cancer, it is reasonable to assume that most of the cells present on the blood correspond to tumour cells. If this is correct, it is also reasonable to think that most of these cells harbor multiples chromosomal abnormalities. It is well described in this type of tumour the presence of genomic alterations (Buitenkamp et al. Leukemia 2012, Mullighan et al. Leukemia 2009). One of these common alterations involves the deletion of one of the two alleles of the gene IKZF1 (one of the two genes that are identify in this study). As it is probable that an alteration in the copy number in the regions of interest would affect the statistical analysis perform in the present study (for example, amplification, deletion or lost of heterozygosity-LOH on the genes involved in this study would surely affect the Hardy-Weinberg equilibrium test and the frequentist-based methods), it becomes essential to test the validity of the data presented in the manuscript that the authors present evidence that these potential effects are not affecting the calculations. Has the Copy-Number/LOH status been tested on the samples used in the study? Have the authors purified normal cells from the blood samples? and if so, how and with which success? If the authors has tested these variables and discarded them, or if they considered that don't affect their analysis a detailed explanation of these circumstances should be included in the manuscript.
Minor Essential Revisions:

In the survival analysis of the patients, the authors discriminate between several risk groups, nevertheless, it is not stated how the patients have been classified on these groups. Are these groups created according clinical criteria (age of diagnosis and cell counting) or according to the data generated on the genotyping analysis? A detailed explanation of what these groups represent need to be included in the manuscript.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests