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: 17 May 2012

Reviewer: Javier Leon

Reviewer's report:

This work explores a new statistical method (Bayesian network based Bayesian multilevel analysis of relevance or BN-BMLA) to analyze SNP associations, using as model a cohort of pediatric ALL. They used data of 543 children with ALL, and 529 healthy controls, which is a very big cohort for most hematological studies - this is a strong point of the work. They studied genotyped 66 SNPs in 19 candidate genes selected form previous literature by their association with ALL. Most of the associations found with the BN-BMLA were also found with the classical logistic regression method.

In general, the work does not give much novel information so as to SNPs associated to ALL risk and survival, and the work is rather confirmatory. However its major aim was to not to discover new associations but to validate the BN-BMLA method. The major conclusion of the paper is supported by the fact that this new method reveals SNP associations also found with the frequentist-based statistical method. For example, frequentist method informs that 6 SNPs in RID5B and IKZF1 genes show association with increased risk to ALL. When they analyze the data with the BN-BMLA, they found association with one SNP in ARID5B and another one in IKZF1. The work demonstrates that BN-BMLA might be a useful supplementary to the traditional frequentist-based statistical method to analyze SNP associations.

Discretionary Revisions

1. The risk-associated SNPs on IKZF1 mapped in the 3'UTR and is already reported. The significant risk-associated SNP of the ARID5B genes maps in an intron. This is unreported, but it is difficult to explain why an intronic SNP should result in increased risk for this leukemia. The same with an SNP (s1294991) in STAT3, associated with a decreased risk in hyperdiploid ALL and also mapping in an intron. In contrast the BAX and CEBP genes SNPs make good sense, as it maps in BAX promoter, and the impact of both genes in leukemia development has been established. A problem that makes it difficult the association with 10 year-survival is that the mortality rate of the studied cohort is lower than in the whole population. The authors might comment on this.

2. However, data on the expression of all these genes is lacking, so the presence of these SNPs cannot be correlated with mRNA expression levels. This could be
incorporated into the Discussion

3. The paper difficult to follow as the results of both methods are never summarized side-by-side in the ALL risk and survival analyses. Such comparison would be very helpful.

4. “In the scientific literature it is known that STAT3 is activated in the presence of active Notch.” A reference is lacking here, as the canonical activating kinases for STATs proteins are kinases of the JAK family, not Notch (which is not a kinase).

5. The Discussion is too long

**Level of interest:** An article whose findings are important to those with closely related research interests

**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