Author's response to reviews

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

Authors:

Orsolya Lautner-Csorba (cs.orsi@gmail.com)
András Gézsi (gezsi.andras@gmail.com)
Ágnes F Semsei (f.semseia@gmail.com)
Dániel J Edélyi (erddan@gmail.com)
Péter Antal (antal@mit.bme.hu)
Géza Schermann (schermann.geza@gmail.com)
Nóra Kutszegi (norakutszegi@hotmail.com)
Katalin Csordás (csordaskatalin@gmail.com)
Márta Hegyi (marta.hegyi@gmail.com)
Gábor Kovács (kovi@gyer2.sote.hu)
András Falus (afalus@gmail.com)
Csaba Szalai (genomika.cs@gmail.com)

Version: 2 Date: 13 June 2012

Author's response to reviews: see over
Candidate gene association study in pediatric acute lymphoblastic leukemia evaluated by Bayesian network based Bayesian multilevel analysis of relevance

Orsolya Lautner-Csorba, András Gézsi, Ágnes F Semsei, Dániel J Edélyi, Péter Anta, Géza Schermann, Nóra Kutszegi, Katalin Csoordás, Mártza Hegyi, Gábor Kovács, András Falus and Csaba Szalai

13/06/2012

Dear Editors,

Thank you for handling and reviewing our manuscript.

Answers to the comments of Javier Leon:

Thank you for reviewing our manuscript, and your kind comments. Each of your suggestion was thoroughly considered.

Comment 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 (rs1294991) 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.

Answers: Thank you for your suggestions. As we wrote in our manuscript the risk-associated SNPs of the ARID5B (rs10821936, rs7089424 and rs4506592) were in LD with each other and in earlier studies the rs10821936 and rs7089424 SNPs were found to be associated with ALL (see ref. 3 and 6). Although the rs4506592 has not been reported earlier, but as it is LD with the other two SNPs it must be the same signal, as we indicated in our paper:

“We calculated the linkage disequilibrium (LD) coefficients between the different SNPs (Additional file 3), and found that in both genes, the significantly associated SNPs were in strong linkage with each other. This means that there is only one, but strong signal in each gene.”

According to your suggestion, we put a comment in the discussion:

“It must be added, however, that all of the associated SNPs in the ARID5B gene are in intron, and presently it is not known, how they influence the risk to ALL.”

The SNP in STAT3 is also in intron, we commented its association in the discussion:

“The decreased risk associated with the SNPs in the STAT3 gene in our study correlates with the results of other studies, e.g. the rs12949918 SNP was found to be associated with decreased susceptibility to different malignancies, like risk to B-cell non-Hodgkin lymphoma
or renal cell carcinoma [27]. According to in vitro studies, the SNP affects STAT3 mRNA levels, with the minor allele having a lower STAT3 expression [28]. Presently it is not known how it might influence the risk to hyperdiploid ALL.”

We commented the different mortality rate in the discussion:

“It must be noted, however, that the rate of died patients is lower in our study population (p<0.001), than in the whole Hungarian ALL population (see Methods), thus our results could be biased in this respect.”

Comment 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.

Answer:

In some cases there are data about such correlations, and we wrote them in the discussion:

“According to in vitro studies, the SNP affects STAT3 mRNA levels, with the minor allele having a lower STAT3 expression [28].”

And:

“In our study, children homozygous to the minor allele of the rs11667351 SNP had a very poor survival rate (40%, Additional files 5 and 6). In a study, this variant was associated with lower BAX mRNA in lymphocytes.”

In other cases we added this to the discussion:

“It must be added, however, that all of the associated SNPs in the ARID5B gene are in intron, and presently it is not known, how they influence the risk to ALL.”

Comment 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.

Answer:

We feel that we compared the two methods in the discussion:

“The BN-BMLA could confirm the association of SNPs in the ARID5B and IKZF1 to B-cell ALL with high posterior probability. Additionally, however, as explained in the detailed characterization of association relation, the results of the analysis gave additional information about the nature of the relations between the SNPs and the disease.”

And later:

“Evaluating the effect of SNPs on the survival rate of the patients resulted in some discrepancies between the two methods. The frequentist-based method detected only nominally significant associations, which, according to the accepted rules, have to be
rejected, while the BN-BMLA, especially in the case of the BAX gene found strong, convincing relevancy. It is generally accepted, that the frequentist-based methods can not properly handle the multiple testing problem. To avoid type I error, sometimes the frequentist methods are too conservative and unable to detect weak effects or interactions. The findings of the BN-BMLA are biologically plausible, but additional studies are needed to confirm these results. In the present paper we show the ability of BN-BMLA to evaluate a candidate gene association study. As can be seen from the results, the advantage is not that the BN-BMLA can detect more relevant variables, but the Bayesian networks offer a rich language for the detailed representation of types of relevance, including direct and indirect aspects.”

Comment 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).

Answer:
It is in reference [31], referred in the end of the paragraph:


Answers to the comments of Ignacio Varela:
Thank you for reviewing our manuscript, and your kind comments. Each of your suggestion was thoroughly considered.

Comment 1:
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.

Answer:
Thank you for your reasonable comment. In our study we used normal cells:
“Genomic DNA from children with leukemia was obtained retrospectively from whole, peripheral blood taken in remission phase using by…”

Furthermore, before isolating the DNA, we measured the normal cell count by FACS, which we added to the Methods:

“Before DNA extraction, the normal leukocyte cell count was checked by flow cytometry (FACS). We used only those samples where the normal cell count was over 5 G/l.”

We also added to the results:

“As the blood was taken retrospectively and in remission, and only those samples were used where the normal leukocyte cell count was over 5G/l, practically no tumour cells were in our samples, thus we investigated only normal cells and germline polymorphisms.”

Comment 2:
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.

Answer:

We included the discrimination criteria in the Methods:

“We stratified our patients in different risk groups according to the following criteria: Low risk (LR) group included children aged 1-6 years who have a white blood cell (WBC) count of less than 20,000/μl at diagnosis, good prednisone response and no T-ALL. High risk (HR) group included poor prednisone response, and/or evidence of t(9;22) (or BCR/ABL), and/or evidence of t(4;11) (or MLL/AF4). Medium risk (MR) group included children with no HR or LR criteria. Prednisone responses were determined after 8 and 33 days of induction treatment. The presence of 1,000 blast/μl or more in the peripheral blood on day 8 and >5% blast in the bone marrow on day 33 were defined as poor prednisone response.”

Thank you again for your valuable suggestions!

Regards,

Csaba Szalai