Author’s response to reviews

Title: Regulatory network analysis reveals the oncogenesis roles of feed-forward loops and therapeutic target in T-cell acute lymphoblastic leukemia

Authors:

Mengxuan Xia (D201677491@hust.edu.cn)
Qiong Zhang (zhangqiong@hust.edu.cn)
Mei Luo (luomei_luom@qq.com)
Pan Li (lip16@mails.tsinghua.edu.cn)
Yingxue Wang (344254557@qq.com)
Qian Lei (leiqian@hust.edu.cn)
An-Yuan Guo (guoay@mail.hust.edu.cn)

Version: 1 Date: 02 Nov 2018

Author’s response to reviews:

Dear Editor,

Please find enclosed our manuscript "Regulatory network analysis reveals the oncogenesis roles of feed-forward loops and therapeutic target in T-cell acute lymphoblastic leukemia" (ID: MGNM-D-18-00206).

Thank you very much for giving us the opportunity to send back a revised manuscript. We would like to express our sincere thanks to the editors and reviewers for the constructive and positive comments. We have revised the manuscript according to all the reviewers’ suggestions.

We have added a list in the following to provide a point-by-point description of the reviewers’ comments.

We hope that you will accept our revised manuscript for publication.

Sincerely Yours

Prof. Dr. An-Yuan Guo
Federica Eduati (Reviewer 1):

1a) I am not convinced of the robustness of the results, given that the paper is based on differential analysis of miRNA derived from the comparison of 48 T-ALL patient samples vs 2 healthy donors (for both CD34+ and CD4+CD8+ normal T-cell subsets), therefore with a very unbalanced sample size and with only 2 healthy samples. Additional computational analysis is required at least to show that differentially expressed miRNAs (DEMs) have comparable expression in the 2 healthy samples (to make sure that the differential expression is not driven by only one sample). Approaches like the one described by Zambelli et al. NAR, 2018 [https://www.ncbi.nlm.nih.gov/pubmed/29390085] could also prove useful to understand if the results of the differential expression analysis are confirmed or if they are rather due to individual samples.

Reply: Thanks for your thoughtful suggestion. As you mentioned, the unbalanced sample size among different groups in the miRNA datasets may cause some bias for the bioinformatics analysis and influence the robust of the final results, especially in the miRNA comparison. According to the suggestion, firstly, we examine the expression profilings of differentially expressed miRNAs (DEMs) in the normal samples, and confirmed that most of the DEMs had similar expression trends in these samples (see the following figure). Meanwhile, the tool RNentropy, which the reviewer suggested, was employed to assess the robustness of DEMs in our analysis. The DEMs in the core regulatory network appeared the robust profiling based on the outcomes of RNentropy. Most of the DEMs showed that all the replicates have local p-value < 0.01, suggesting the results were reliable and not caused by the bias of individual samples. Additionally, we supplemented this process to the method section of revised manuscript with the following words: "To avoid the bias caused by the unbalanced sample size in the different groups, RNentropy was employed to evaluate the robust of miRNAs expression, and the miRNAs with convergent expression profiles were used for further analyses".

Expression heatmap of differentially expressed miRNAs in normal cells

1b) Additionally, given availability of miRNA expression data for T-ALL cell lines (from the same publication of the 48 T-ALL patients and 2 healthy donors (18)), I am wondering why the author did not follow the same approach used to define differential gene expression (DEG), i.e. comparing T-ALL cell lines vs normal, T-ALL patient vs normal and looking at the intersection. In Results authors actually state that they used "T-ALL samples (patients and cell lines)", where the cell lines data used, and if yes how?

Reply: Thanks for the kind comments. miRNA expression data derived from GSE89978 [www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89978]. Actually, we didn’t used the cell line data because the original article of the dataset also only used this data to confirmed the differential expressed miRNAs between T-ALL and normal (PMID: 28801656). According to
your consideration, we followed the same approach used in the DEG detection, which identified the DEMs use the intersection part of DEMs from T-ALL patient vs normal and cell lines vs normal comparisons. In this analysis, miR-21-5p was identified as a DEM as its fold change is 2.9, whereas the fold change of miR-19b-3p in the cell lines vs normal comparison was 1.9. However, considering the heterogeneity of samples and the bias underlying the microarray technology, the miR-19b-3p was considered as a candidate DEM for further analysis.

2a) The authors claim that their network analysis is essential to identify feed-forward loops and therapeutic targets, however no validation is made of the role of the regulatory mechanisms of the feedforward loops, but they rather discuss the relevance of individual elements of the module (or of gene sets), which were identified simply by differential expression analysis (without need for network interpretation, apart for the definition of the hub nodes). For example, the authors mention that downregulation of CDC25A induce cell cycle arrest and miR-21-5p knock down induces apoptosis, but what is the their role as part of the FOXM1-miR-21-5p-CDC25A FFL? This is a coherent FFL and this kind of module has been known to function as a surveillance mechanism, is there any evidence this might be the case for this particular FFL?

Reply: We are very grateful to your comments. Although we didn't perform experiments for the FOXM1-miR-21-5p-CDC25A FFL, we predicted it base on published reports which had validated the regulation of this loop, such as (PMID: 20354770) and (PMID: 28096467). In the FFL, the up-regulated expression of the TF FOXM1 could increase the expression of both CDC25A (PMID: 23240008) and miR-21-5p, whereas the miR-21 could target CDC25A (PMID: 19826040), which in turn formed a dynamic balance for the FFL involved in the cell proliferation process, suggesting the miR-21 may mediate a negative feedback loop collaborated with the TF FOXM1 to regulate their targets CDC25A. Consequently, although the miR-21 repressed the expression of CDC25A, the CDC25A maintained at a high expression level promoted by TF FOXM1, which may trigger the abnormal cell cycle and cell proliferation in the progress of T-ALL.
T-ALL, while the core subnetworks with highly connected nodes were focused on the specific issues, such as the FFL FOXM1-miR-21-5p-CDC25A in abnormal cell proliferation. As the reviewer mentioned, the roles of FFL should be discussed and presented in the paper, which could help to understand the underlying regulatory mechanism. In the revised manuscript, we further exploited the possible role of the FFL FOXM1-miR-21-5p-CDC25A in the discussion section, and presented in the reply of the comment 2a as well.

3) The paper is missing important details in the Methods section, especially in the description of "Generation of network and analysis of hub components":

a) How is the network reconstructed? Reference to in-house scripts and previous publication is not sufficient, essential details to understand the nature of the inferred network should be provided here. For example, was the network derived purely from databases and prediction algorithms (as described in (13)) or did they also use expression correlation to filter the false-positive regulatory interactions (as described in (20))? Additionally, providing open source script to the community would be highly appreciated and would guarantee reproducibility of the results.

Reply: Thanks for your comments, and we had described the detailed methods in the revised manuscript: "Briefly, the miRNA-TF-target regulatory network was constructed by the following three components: 1) TF-targets regulation. TF-targets regulatory information was obtained from the ChIP-Seq experiment datasets of public databases (ENCODE, hTFtarget (http://bioinfo.life.hust.edu.cn/hTFtarget/) and TRANSFAC) and predictive TF-targets regulation of UCSC; 2) miRNA-targets regulation. Experimentally validated miRNA-targets regulations were incubated from miRTarBase and TarBasev7.0, while the predictive regulations were collected form Targetscan and miRanda; 3) Regulatory network detection". According to the regulatory relationship, we formed node table files and edge table files and visualized the information by Cytoscape. Actually, we are working on a webserver which contains the scripts used to construct the regulatory networks, and it will be online soon.

b) How was the subnetwork (Fig 4) derived and how was the pathway/GO term enrichment done? The description in methods "We selected top 5 TF and miRNA and combined with GO terms results to find core TFs and miRNAs" is not sufficient to understand and reproduce the author's findings.

Reply: We are sorry about unclear statement of this result, and we had stated it in the revised manuscript. The pathway/GO term were enriched by the DEGs using DAVID Functional Annotation Tool, which was supplemented in the method section. The subnetwork (Figure 4) was built by the DEGs enriched in the GO terms and the corresponding regulatory DEMs and
TFs. Based on the subnetwork, the top 5 TFs and miRNAs that contained the maximum connection combined with their targets were extracted to find the core TFs and miRNAs. Finally, after verifying their function and regulatory relationship with their target, we predicted FOXM1-miR-21-5p-CDC25A and SOX4-miR-19b-3p-RBBP8 as core FFLs in the regulation of cell cycle and cell division.

c) Usage of GSCALite tool is also poorly described (e.g. was the Spearman correlation between gene expression and drug sensitivity computed pan-cancer of specifically for ALL?)

Reply: Thanks for your thoughtful reminding. We had described it detailly in the revised manuscript with the following sentences: "The half maximal inhibitory concentration (IC50) information of 746 drugs and the corresponding genes expression profiles of 1861 cell lines were collected from GDSC and CTRP databases. The spearman correlation coefficient between the IC50 and genes expression profiles were calculated in all cell lines using the GSCALite, and the drug-gene pair with the p-value < 1e-4 & |cor_sprm| > 0.4 was considered as significant one used for the further analysis".

Other minor comments:

- p7 53-57: I would not highlight as a result that the "T-ALL and normal samples were segregated well by the expression level of TFs" in Fig. 2B, since this is an obvious consequence of the fact that only differentially expressed (between T-ALL and normal) TFs are shown here.

Reply: Thanks for your kindly suggestion. As you suggested, we removed it in revised manuscript: “Among the 434 DEGs, 31 differentially expressed TFs (13 up and 18 down) were identified and their expression hierarchical clustering analysis was showed in Figure 2B.”

- p8 7-11: "Our network contained 486 edges, which consisted of 132 DEG and 12 miRNA", please rephrase with the current structure "which" seems to refer to "edges".

Reply: Thanks for your careful reviews on the manuscript, we had rewritten this sentence in revised manuscript. “Our network contained 486 edges that were consisted of 132 DEGs (14 TFs and 118 genes) and 12 miRNAs”

- p8 11-19 not clear (probably because the corresponding methods section is not complete - see comment 3a) and need grammar check, e.g. "FOXM1 was only one TF" should be
"FOXM1 was the only TF"; "SOX4 regulated about 66% genes" should be "SOX4 regulated about 66% of the genes". Similar grammar issues are also elsewhere in the manuscript.

Reply: Thanks for the reviews. We had carefully amended the grammar mistakes and unclear descriptions to improve the readability in the revised manuscript.

- p9 27 refers to Figure 6B, which has no panel marked with B (I understand that the authors refer to the bottom part of the figure).

Reply: Thanks for your reminding and we are sorry for the negligence. We have added the panels on the corresponding figure and carefully checked all figures.

- Results section "Potential drug indicators of genes in the regulatory network" could be expanded going deeper in the interpretation of the results. For example, how should positive vs a negative correlation be interpreted?

Correlation seems to have been computed pan-cancer and based on the GDSC and CTRP gene expression data (not clear, see comment 3c). If this is the case, are positively and negatively regulated genes up- or down-regulated in the T-ALL dataset and can this tell us more about expected sensitivity/resistance to the drug? And can this be linked to positive/negative regulators identified in the network (to give an actual network interpretation of the results)?

In the same section (p10 15-19) the authors state "we found MCM2 was upregulated in the T-ALL samples and showed negative correlation with 39 drugs, which implied MCM2 may serve as a potential therapeutic target". My interpretation would be that if MCM2 is negatively correlated with drug A, then up regulated samples (i.e. T-ALL) are more sensitive to the drug (lower IC50), but it is not clear to me why MCM2 could be seen as a potential therapeutic target based on the above observation.

Reply: Thanks for the careful review and thoughtful suggestion. Actually, we didn’t dig deeper into the research between genes and their correlation with drugs. Genes with positive or negative correlation to a certain drug have the probability to serve as a biomarker to predict the sensitivity or resistance of cells to this drug. MCM2 showed high correlation with 39 drugs, which indicated that MCM2 may participate and play critical roles in biological pathway influenced by these drugs. The regulation of MCM2 may influence the treatment effect of these drugs. Thus we inferred that MCM2 may serve as a potential therapeutic target. But it was hard to confirm the causality between the expression profiling of genes and the drugs. That is to say, the results of
"Potential drug indicators of genes in the regulatory network" in our manuscript just provided a resource, which demonstrated the drug-gene information underlying the regulatory network.

Velia Siciliano (Reviewer 2): Please include all comments for the authors in this box rather than uploading your report as an attachment. Please only upload as attachments annotated versions of manuscripts, graphs, supporting materials or other aspects of your report which cannot be included in a text format.

Please overwrite this text when adding your comments to the authors.

In the manuscript "Regulatory network analysis reveals the oncogesis roles of feed-forward loops and therapeutic target in T-cell acute lymphoblastic leukemia", the authors analysed transcriptional and post-transcriptional profiles of T-cell acute lymphoblastic leukemia (T-ALL) and normal T cell samples. The goal is to identify new dysregulated pathways and new therapeutic targets.

I find the manuscript of general interest for the immunology community that would benefit from the new insights suggested in the paper.

In general, I believe that the authors should improve the writing because there are sentences that leave me confused. For example, Page 5 line 2-6: "The research of FFLs could speed up the confirmation of oncogenes and decrease complexity of experimental verification, and expand the understanding for regulatory mechanisms underlying specific biological processes." I do not understand the meaning of this sentence.

Reply: Thanks for the comments, and sorry for the unclear statement. To avoid the similar statement, we had carefully revised the manuscript to improve the readability. The above descriptions were modified as follow: “The FFLs could reduce the complexity of regulatory networks and provide comprehensive clues to identify oncogenes and the underlying regulatory mechanisms”.

The authors explore an interesting topic which is the role of microRNA in network topologies. It would be good in the introduction/discussion to expand this aspect. For example, they should comment about our network topologies such as cross-regulated positive and miRNA based negative feedback loop that has demonstrated that miRNA have an essential role to buffer gene expression role. They should cite Siciliano et al Nature Communications 2013, and comment
about whether the microRNAs they identified could help in fine regulation of target genes to skew the direction of T cell towards cell cycle or apoptosis.

Reply: Thanks for your thoughtful suggestion, and we had appropriately cited the reference in the revised manuscript. According to the suggestion and Siciliano’s paper, we had deeply reconsidered the roles of the miRNAs related feedback loop underlying the development of T-ALL, and discussed the role of FOXM1-miR-21-5p-CDC25A involved in the process. For example: “The upregulated expression of the TF FOXM1 could increase the expression of both CDC25A (47) and miR-21, whereas the miR-21 could target CDC25A (48), which in turn formed a dynamic balance for the FFL involved in the cell proliferation process, suggesting the miR-21 may mediate a negative feedback loop collaborated with the TF FOXM1 to regulate their targets CDC25A. Consequently, although the miR-21 repressed the expression of CDC25A, the CDC25A maintained at a high expression level promoted by TF FOXM1, which may trigger the abnormal cell cycle and cell proliferation in the progress of T-ALL.”

There are also several typos to fix:

miR… the M should be always lower case letter, regardless if it is at the beginning of a sentence (i.e. see Pag5 line 9)

Pag 6 line 13 "And" at the beginning of the sentence should be removed

Pag 6 line 40 "thymocytes" is thymocytes

Pag 8 line 9 different should be differentially

Pag 13 line 11 "jurket" is "jurkat"

Reply: Thanks for your careful reviews, and we had fixed these typos in the revised manuscript.

Pag 8 line 50-51 "In our results, the identified 31 TFs (13 up and 18 down) were significantly expressing in the 434 DEGs" I do not understand the meaning of significantly expressing (were the TF significantly expressed? What does significantly mean here?"

Reply: We are sorry for unclear statement of this result. We rewrote in revised manuscript: “Among the 434 DEGs, 31 differentially expressed TFs (13 up and 18 down) were identified and their expression hierarchical clustering analysis was showed in Figure 2B.”
Pag 9 line 19-22 "In conjunction with the results above, the TFs and miRNAs acted as hub regulators in our regulatory network may regulate their targets genes to form key modules involved in the development of T-ALL." Not clear the meaning

Reply: Sorry for the unclear description. We had revised this in the manuscript:” In conjunction with the results above, the hub TFs and miRNAs combined with their target genes in our regulatory network may form key modules involved in the development of T-ALL.”