SYBO-D-15-00024

Pathobiochemical signatures of cholestatic liver disease in bile duct ligated mice

Kerstin Abshagen; Andreas Hoppe; Maria Thomas; Isabell Müller; Matthias Ebert; Honglei Weng; Herrmann-Georg Holzhütter; Ulrich M Zanger; Johannes Bode; Brigitte Vollmar; Steven Dooley

BMC Systems Biology

Dear Dr. Abshagen,

Your manuscript "Pathobiochemical signatures of cholestatic liver disease in bile duct ligated mice" (SYBO-D-15-00024) has been assessed by our reviewers. They have raised a number of points which we believe would improve the manuscript and may allow a revised version to be published in BMC Systems Biology. Please consider this like a major revision.

Das ist ein milestone; wir sollten uns wirklich bemühen und dann haben wir das paper ordentlich publiziert.

Their reports, together with any other comments, are below. Please also take a moment to check our website at <http://sybo.edmgr.com/> for any additional comments that were saved as attachments.

If you are able to fully address these points, we would encourage you to submit a revised manuscript to BMC Systems Biology. Once you have made the necessary corrections, please submit online at:

<http://sybo.edmgr.com/>

If you have forgotten your username or password please use the "Send Username/Password" link to get your login information. For security reasons, your password will be reset.

Please include a cover letter with a point-by-point response to the comments, describing any additional experiments that were carried out and including a detailed rebuttal of any criticisms or requested revisions that you disagreed with. Please also ensure that all changes to the manuscript are indicated in the text by highlighting or using track changes.

Please also ensure that your revised manuscript conforms to the journal style, which can be found at the Instructions for Authors on the journal homepage.

A decision will be made once we have received your revised manuscript, which we expect by 12 Jul 2015.

I look forward to receiving your revised manuscript and please do not hesitate to contact us if you have any questions.

Best wishes,

Julio Vera

BMC Systems Biology

<http://www.biomedcentral.com/bmcsystbiol>

Reviewer reports:

Reviewer #2: The manuscript "Pathobiochemical signatures of cholestatic liver disease in bile duct ligated mice"

by Abshagen et al. aims at identifying markers for chronic liver disease progression

by measuring physiological parameters as well as the expression of selected genes

in mice following disrupted bile ligation over time up to 14 days.

The authors performed extensive correlation analyses between the expression data and

the various hematological measurements and combined them in a

"consensus correlation" value to predict, which genes and factors explain best the

different disease phases.

Finally, they propose a decision tree based on the main markers they found

to predict the disease stages.

The manuscript is a nice example of a collaborative effort in

experiment, analysis and interpretation, yielding interesting new insights into

the chronic liver disease. Overall, the results are well explained and compared

to the literature. In particularly the decision tree approach is a good idea to

translate the experimental data into new biological insight of chronic liver

disease progression.

Nevertheless, the manuscript needs a major restructuring of the data analysis

and a clear focus in its presentation. The reader is flooded with gene lists,

information and interpretation, while the motivation and explanation of

the analysis is underrepresented in the main text.

The data analysis has been completely restructured. The respetive method section has been extended and all information for reproducing the analysis is provided (source code, data sets). The important information form the supplements has been merged in the methods and supplementary information is limited to the correlation analysis with alternative correlation measures for comparison.

All information is there,

but scattered throughout the main text, methods section and supplement.

In particular Figs. 8 and 9 are overloaded with information and are hard to understand.

The manuscript would win substantially, if the authors focused on selected genes

and results, better explained their analysis approaches in the main text and

redrew their figures in a more intelligible format with readable fonts and structured

information.

The information has been reduced employing cluster methods /dimension reduction.

Major Points

- The authors should consider dimension reduction methods such as multidimensional

scaling or principal component analysis, when discussing the samples' expression

patterns, use statistical tests, when assessing the significance of gene regulation over time

The dimensionality was reduced by clustering and only looking at the main clusters.

TODO: statistical tests?

and display ROC curves, when discussing their decision tree model. This would

reduce lengthy description of the data and provide better overview on the dynamic behavior of the system.

TODO: implement decision trees with ROC curves

- For the qPCR data the authors used a single Gene, Gapdh, for normalization, which possibly

results in noisy dCT (delta CT) values. There is no guarantee that the

expression of this gene remains constant across the samples, given the severe impact

of BDL and the measurement time of 14 days. Usually, dCT values are

normalized to two control genes, also e.g. 18S and/or Hprt1. The authors need to check

and show the behavior of Gapdh and that their normalization approach does not affect their

results.

- The authors state a delta delta CT of approx. 17 between IL28b and Gapdh (Fig. 6K, 5d).

This is a rather large difference and would mean that IL28b appears around the qPCR cycle 5 already.

This should be checked. What is the primer efficiency for IL28b? Maybe it is unusually high.

- The benefit of the consensus score needs to be better discussed, e.g.

the choice of different weights seems rather arbitrary.

In supplement2, 1.2 Consensus correlations the authors state that

"the correlation of time average has 4 times higher weight".

Where does this number come from?

Why are the correlations separated in positive and negative parts and

then the smaller part are ignored? A lot of information is lost this way.

The score should be normalized to be able to compare in between factors.

- Why do the authors use a consensus measure to perform clustering (Fig. 7)?

Why do they not use all correlation data for each factor as a matrix,

and then apply clustering method on this matrix instead?

- The authors used the Pearson Correlation, which can be easily influenced

by outliers. In particular in combination with the noisy qPCR data this can lead to

spurious correlations. The authors should check for consistency of their results by

using either more robust correlation measures like Spearman correlation

or by low-pass filtering their data before performing the analysis.

- The description of the separator approach in the manuscript is very wordy and formulas are necessary

to understand what has been done.

- Figures 8 and 9: Significance and consensus scales are not really readable.

It is hard to distinguish between the not significant ones (0.1) and the significant ones (<0.05).

For example, few different colors would be more efficient than yellow gradient.

Also the correlation scale should be from -1 to 1.

- Overall, figure legends need to be improved, as they lack sufficient annotation

to understand what is displayed.

Minor Points

- All gene symbols should be written consistently with small letters and a capital

first letter throughout the text and figures, e.g. use Gapdh instead of GAPDH.

- in the derivation of the consensus correlation the authors use a rather unusual -log100 transformation.

is this true or a typo?

- The relative expression of the genes in Fig. 6 should be displayed in log2 scale and

the domain ranges should be the same for all genes to make the changes in gene expression comparable.

- Most likely, the authors used a log2 scale in Fig. 5, but annotation of the

color bars is missing and needs to be added.

- Page 9 lines 17 and 19, Page 14 line 19: Figs. 7 and 8 are actually Figs. 8 and 9

- Place Figure 7 after Figure 8, as the former is based on the latter

- The abbreviations in the leaf names of the circular tree in Fig. 7 are nowhere explained in the

main text and need to be added.

--------------------

Editorial Requests

--------------------

Please note that all submissions to BMC Systems Biology must comply with our editorial policies. Please read the following information and revise your manuscript as necessary. If your manuscript does not adhere to our editorial requirements this will cause a delay whilst the issue is addressed. Failure to adhere to our policies may result in rejection of your manuscript.

Ethics:

If your study involves humans, human data or animals, then your article should contain an ethics statement which includes the name of the committee that approved your study.

If ethics was not required for your study, then this should be clearly stated and a rationale provided.

Consent:

If your article is a prospective study involving human participants then your article should include a statement detailing consent for participation.

If individual clinical data is presented in your article, then you must clarify whether consent for publication of these data was obtained.

Availability of supporting data:

BioMed Central strongly encourages all data sets on which the conclusions of the paper rely be either deposited in publicly available repositories (where available and appropriate) or presented in the main papers or additional supporting files, in machine-readable format whenever possible. Authors must include an Availability of Data and Materials section in their article detailing where the data supporting their findings can be found. The Accession Numbers of any nucleic acid sequences, protein sequences or atomic coordinates cited in the manuscript must be provided and include the corresponding database name.

Authors Contributions:

Your 'Authors Contributions' section must detail the individual contribution for each individual author listed on your manuscript.

Further information about our editorial policies can be found at the following links:

Ethical approval and consent:

[http://www.biomedcentral.com/about/editorialpolicies#Ethics](http://www.biomedcentral.com/about/editorialpolicies" \l "Ethics)

Standards of reporting:

[http://www.biomedcentral.com/about/editorialpolicies#StandardsofReporting](http://www.biomedcentral.com/about/editorialpolicies" \l "StandardsofReporting)

Data availability:

[http://www.biomedcentral.com/about/editorialpolicies#DataandMaterialRelease](http://www.biomedcentral.com/about/editorialpolicies" \l "DataandMaterialRelease)