Reviewer's report

Title: The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; Data release 1

Version: 0 Date: 24 Aug 2018

Reviewer: James John Ashton

Reviewer's report:

Background-

The authors present a succinct overview of the importance of developing multi-omic data in IBD in order to move to a systems biology and personalised medicine approach to disease. It would be interesting to discuss the recent publications from the RISK cohort where use of these data has enabled modelling to take a first step towards precision medicine in IBD. I feel this would place the 1000IBD project in context and emphasise it's importance.

What will be the arrangements for publication of any results developed from these data by institutions external to Groningen? Will co-authorship be required or just citation?

Methods- This is clearly an extremely ambitious and important project, with very important data likely to be generated by this work (and already developed). Whilst I do not include specific comments about the outputs yet several questions do arise which would be important to answer prior to using these data. Additionally the use of patient-based reporting systems has previous had flaws (reporting bias, subjective etc.) but also represents a very interesting way of collecting data.

What were the ages of the patients included? Were paediatric samples included?

Were the two new questionnaires developed for this study validated, reviewed externally or piloted prior to use? The importance of the quality of the clinical phenotypic data here cannot be understated

Will all raw fastq files from the WES, microbiome and transcriptomic sequencing be available to run through custom pipelines?

RNAseq and microbiome (MGS) data is notorious for being subject to batch effects, could the authors comment on how the 300 samples where grouped and how batch effect was minimised?
How quickly were biopsies snap-frozen after sampling? Previous studies have used RNAlater to preserve bacterial DNA and human RNA to prevent degradation, did the authors consider this?

Similarly how many of the stool samples were frozen within 15 minutes? The exact conditions for samples is important in determining the quality of the output data.

At what point in the disease course were patients recruited? Are any treatment naive? Are any followed longitudinally with multiple samples available?

Overall- This is a good report and the publication would benefit researchers to understand the access to these data, clearly further work is ongoing.

There are several typos that should be corrected including-

Page 10, line 43- GENERATION OF DIETART AND ENVIRONMENTAL DATA

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics
Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal