Author's response to reviews

Title: Genome-wide Associations of Signalling Pathways in Glioblastoma Multiforme

Authors:

Stefan Wuchty (wuchtys@ncbi.nlm.nih.gov)
Alexei Vazquez (vazquez@ias.edu)
Serdar Bozdag (bozdags@mail.nih.gov)
Peter Bauer (ppeettoob@gmail.com)

Version: 3 Date: 15 December 2012

Author's response to reviews: see over
Dr. Tim Sands,
Series Editor, BMC Medical Genomics

Dear Dr. Sands,

I am writing to submit a revision of the manuscript MS: 4505447597509390 ‘Genome-wide Associations of Signalling Pathways in Glioblastoma Multiforms’ by S. Wuchty, A. Vazquez, S. Bozdag and P. Bauer that we would like to be reconsidered for publication in BMC Medical Genomics. Please find attached a point-by-point response to the comments of the reviewers.

In the revised manuscript, we determined significant associations between genomic alterations and the expression change of signalling pathways in glioblastoma multiforms (GBM). Applying a machine learning approach, we mainly found that under-expressed pathways were associated to chromosomal regions with no signature alterations, while we found that over-expressed pathways strongly associated to signature alterations. We also found that identified genomic regions were highly enriched with genes having driving roles in gliomas, a result that indicates the biological relevance of obtained associations. Furthermore, such genes were found to be concentrated in certain signaling pathways, allowing us to define driver pathways. We believe that the revision of our manuscript dealt well with the points that the reviewers raised and is of interest for the readers of BMC Medical Genomics.

Best regards,
Stefan Wuchty
Reviewer 1:

“The present manuscript is a revised version of one that I reviewed for BMC Bioinformatics a while back. My comments at that point largely revolved around how the methods were presented, because there was a lot I did not understand of both details and rationale. The main point in this manuscript, as I understand it, is the combination of several existing methods into a pipeline for combining gene expression data and copy number variations to yield lists of loci or genes that are associated with specific tumor types.

There are marked improvements in the manuscript since last I saw it, both in terms of both language and in how the contents are presented. (That said, it is clear that the authors should switch on their spellcheckers - there is an "enrichmnet" of typos in some parts.)”

We thank the reviewer for his positive remarks. As for typos we took extra care to avoid those.

I am still unsure - and more so after reading the comments of reviewer 2 of the previous round - of where the benefits lie in the authors' method compared to earlier or alternative methods. The manuscript is fairly long in relation to its contents, and a lot of it is about the method details (it even has equations; cf. [B. R. Jasny, Heavy use of equations impedes communication among biologists, PNAS 2012]) so I expect the method details to be important. Yet I find myself guessing (or not guessing) whether specific steps were standard procedure, novel, necessary or something else again. (The authors gave me a good explanation of one step in the earlier reply: "we focus on those genes in a pathway that are important for the pathway in the disease case").”

As for the justified request of the reviewer we are as detailed and distinct as possible in the revised manuscript. Also, we took care to explain the rationale for each step in our procedure.

“I'm having a hard time formulating a coherent criticism of the manuscript. There is something about it that gives me a bit of a headache, and it could well be that I know too little about the subject. But when I read the abstract I think I start to see what it is that bothers me: The method feels like it's mostly there for a one-shot analysis of the data, but the biological results feel like they're there as a case study for the method. Or maybe it's something else. And maybe I'm being too harsh, but I think the authors should think hard about what the important parts of the manuscript are, and try to make things as clear to the reader as it is to them. The method looks sound, and the biological results appear to be consistent with and adding to existing knowledge, and if that's the case then they deserve to be presented in an accessible way.”

We agree with the reviewer that the main message of the paper was not as clear as possible. In the current revision, we focus on eQTLs of pathway/locus pairs in glioblastoma multiformes and stress the relevance of our results. In particular, we
show that there exists a marked difference of eQTLs that are linked to up- and down-regulated pathways and show the biological significance of the identified loci.

Reviewer 2:

“The article by Wuchty et al introduces a new approach to associate copy number altered loci and pathways. It does so by first identifying altered genes by pathway enrichment statistics, followed by application of random forest methodology to associate CNAs to the average expression levels of selected pathway genes. For glioblastoma multiforme, the method uncovers well-established chromosomal alterations, and for oligodendrogliomas, new regions are proposed. The paper is interesting as an example of how to gain biological interpretability from cancer genomics data, is clearly written, is metodologically sound (comment below), and should be of interest to specialized readers in the fields of neurooncology and computational methods. I recommend publication.”

We thank the reviewer for his positive comment!

“My main comment is that it is not sufficiently clear from the text what normal samples are used (normal brain? other tissues? cultured cells?) and how the choice of reference might significantly alter the results. It would be important to clarify this before publication.”

We agree with the reviewer and clarified the point in the manuscript. As a matter of fact the control samples were taken from epilepsy patients, a control that is widely accepted.

“It is not clear whether the simulation approach used to obtain p-values controls for multiple testing. Is a separate p-value calculated for each association, or is there some form of global correction, please clarify.”

We clarified this point in the manuscript. All p-values obtained (i.e. for all associations) were corrected for multiple testing with a FDR.

“Minor comment: the fact that genes have a high gsea score, here termed 'leading edge genes', does not imply that such genes 'drive the pathway', which is a recurring statement in the results section. More, generally, it is not clear what 'drive a pathway' means, so it would be better to use much more neutral language ('pathway members with a high expression relative to a reference panel of noncancer tissues' or similar).”

We agree with the reviewer. We changed those sentences to ‘leading edge genes drive the enrichment of pathways’.