Peer Review File

Article information: http://dx.doi.org/10.21037/atm-20-6519.

Reviewer A

Novelty

Comment 1: There is an obvious lack of novelty, as this type of study has already been conducted in mouse and rat models of spinal cord injury (either transection or contusion), by RNA-seq data analysis or even microarray data analysis before that. Here, the authors claim that a comprehensive analysis of gene expression spatio-temporal changes has not been reported, which is not the case. One of the published studies is actually a resource article depicting molecular changes in the acute and subacute phases following thoracic contusion injury (Chen et al., Plos One, 2013). This should be referenced in the present paper and the results compared or at least discussed.

Reply 1: We included this reference and added more discussion in the revised version.

Comment 2: Furthermore, the present study gives a temporal but not spatial analysis, as the gene expression changes are not analysed in different tissues or anatomical localization around SCI. In fact, several other transcriptomic studies exist, with ever more details given on specific cell populations - for example Anderson et al., Nature, 2016 focusing on the positive role of astrocytes in regeneration; or Li et al., Nature, 2020 focusing on the role of microglia in resolving inflammation. Analysis of cell contribution to SCI progression is essential to test for axon regeneration and functional recovery. While some studies are indeed referenced, discussion of existing results is lacking.

Reply 2: We agreed with this reviewer and now added additional discussion to address this issue.

Design

Comment 3: Some details regarding the design of the study are missing, eg, number of rats per condition, exclusion criteria, how sham surgeries are obtained, localization of tissue of interest (slices proximal or distal to the injury site, inclusion of the injury site or not, slices pooled or

not). A schematic of the design would be appreciated. Resulting n number for biological replicates in RNA-seq is missing.

Reply 3: We added detailed descriptions in the method section as follows:

We performed laminectomy without any contusion injury as sham control. For experimental groups, we applied the Basso, Beattie and Bresnahan (BBB) locomotor scale at 1, 4, and 7 days post injury and excluded all animals that show hindlimb movement (BBB ≥ 3 at 1d post injury). Post hoc analysis included 4, 4, 4, and 4 animals for sham, 1d, 4d, and 7d post injury, respectively. Spinal cord tissues (0.25 mm rostral and caudal towards the injured epicenter, with injury site included) were harvested and pooled together. As suggested by this reviewer, we now added a schematic design in Fig. S1. We did not repeat RNA-seq results in this study.

Comment 4: Regarding DE analysis, is there a justification of such a stringent FDR-corrected p-value (0.001)? Many DE genes are probably missed in this case.

Reply 4: In literature, there were typically 3 commonly used criteria for FDR-corrected p value (0.05, 0.01 and 0.001). We selected such criteria for stringency.

Comment 5: Regarding RT-qPCR, n number for biological replicates is missing.

Reply 5: We used 3 biological replicates.

Results

Comment 6: Regarding RNA-seq data analysis, Supp Table S1 should be giving a full list of DE genes between the 4 timepoints. What is the table actually showing? It displays log2(FC), between which conditions? Does this table contain genes that could be used as controls, eg, expression of GFAP or Vimentin as markers of astrogliosis?

Reply 6: Table S1 contains 3 parts which represent comparisons between 0d and 1d (part 1), 0d and 4d (part 2) and 0d and 7d (part 3). Therefore, this table shows comparisons between individual time points as labeled in the column sample_1 and sample_2 in each page. We were sorry for any misunderstanding.

Comment 7: Data interpretation lacks accuracy. Why do the authors justify the increase in the number of down-regulated genes by increased cell death? Gene expression data as given by

FPKM should be normalized by the whole library size, so should account for reduced number of cells linked to cell death. Hence, this suggestion is not relevant or should be discussed in more details with the RNA-seq data in hand (check for markers of cell death for example).
Reply 7: We agreed with this reviewer that, without additional data, it was imprecise to claim that the increase of down-regulated genes attributes to increased cell death. We now removed such saying in the revised version.

Comment 8: One minor suggestion is to give the full name of the genes uniquely enriched at each timepoint, as they seem to be the genes of interest newly identified, yet no detail is given about their expression or function.
Reply 8: We did that in the revised version (see updated abbreviation table).

Comment 9: Regarding RT-qPCR, there should be at least one other housekeeping gene to be used as a control in addition to GAPDH.
Reply 9: Taking gapdh as the internal control, there is no significant difference in 18s at different time points, indicating that gapdh is a suitable internal control.

Comment 10: Supplementary Table S5: what are the criteria for selection of genes of interest?
Reply 10: We selected genes of interests based on the highest degree calculated by the gene-act-network analysis, along with those showed FDR p value < 0.05 by pathway analysis.

Discussion

Comment 11: As already mentioned above, discussion should be provided regarding existing results, since this type of study has already been published before. The authors claim that multiple studies have analyzed gene changes following SCI at the genomic and proteomic levels, but not transcriptomic, however it is not the case.
In addition, since SCI affects different cell types, a more detailed discussion should be added regarding the relation between gene expression changes in the acute and subacute phases and the specific cellular events taking place. These events are well described histologically and molecularly. Discussion should account for the heterogeneity in cell types affected by SCI, and in SCI models themselves.

Reply 11: Thanks for pointing this out. We rewrote the discussion by adding more comparisons between different studies and added more discussions relating to the heterogeneity in cell types affected by SCI.

Comment 12: The newly identified hubs deserve a more relevant discussion: what are Gngt2, Mgst2, Pla2g3? Is there anything known about these genes in the context of CNS degeneration and regeneration? About neural development? How do the authors plan to test functionally the implication of these genes to SCI pathophysiology? To SC repair?
Reply 12: We added more discussions of their functions related to CNS degeneration/regeneration and/or development. We will first identify their expression changes in distinct cells and then perform functional manipulations.

Format

Comment 13: All references in text are missing from the current manuscript, so it is difficult to evaluate literature review in the introduction and discussion.
Reply 13: We were sorry for this formatting mistake and now corrected it.

Comment 14: Some figures are unreadable (Fig3 and Fig4), so here again difficult to link the results to the figures.
Reply 14: We enlarged fonts of Fig. 3 and 4.

Reviewer B

Comment 1: One of the difficulties in assessing the contributions of the research presented is the lack of putting what is found in context about what is already known about spinal cord injury. For instance, a quick search through GEO turns up 452 rat series with spinal cord injury. Are any of these at similar time points? How do the results compare, contrast, or complement each other? In addition, the introduction is very basic, and does not contain any in-text citations about what is known about the response to spinal cord injury at early time points, whether or not a genome-wide transcriptional study has been performed. It is important to put the results and conclusions of this particular study in the context of what is known, which is currently not done in the manuscript.

Reply 1: Thanks for pointing this out. We now significantly rewrote the introduction/discussion for a comprehensive review of contemporary knowledge of the field.

Comment 2: Critical details are missing in terms of the RNA-seq analysis. How were the polymer, primer adaptor, and ribosomal RNA removed? Was MT RNA removed as well? What are the specific details for reproducibility (software, versions, parameters, gtf files, genome version).

Reply 2: We used cutadapt to remove the polymer, primer adaptor and ribosomal RNA. We did not particularly remove MT genes. We added detailed information to for reproducibility.

Comment 3: What is the motivation for using SOAP aligner? There are more commonly used aligners such as STAR, HiSat2, kallisto that are used today.

Reply 3: We feel sorry for this mistake. Indeed, we used Tophat but not SOAP for alignment (see below).

Comment 4: Why was a fold-change of 2 used? This will remove significant genes at a higher expression level.

In one place in the text, SOAP is used for alignment; in another, Tophat is used. The methods section is lacking critical details, and is marred by the inconsistency throughout.

Reply 4: We used 2-fold change for stringency.

Thanks for pointing out the alternative using of SOAP with Tophat. We used Tophat but not SOAP in this study and now significantly rewrote method part with more details.

Comment 5: How were the KEGG pathways identified? What sort of enrichment analysis was performed? How were the gene networks constructed? There is very little information presented here, and the approach cannot be adequately evaluated, including an analysis of their “in-house” pipeline.

Reply 5: We significantly rewrote method part. For pathway analysis, we applied Fisher’s exact to calculate p value and FDR according to KEGG database. We added a reference for detailed analysis (Draghici S, et al. A systems biology approach for pathway level analysis. Genome Res. 2007 Oct;17(10):1537-45).

Comment 6: The authors mention that “NF-kappaB pathway needs further research as a potential therapeutic target for clinical spinal cord injury.” However, the reality is there is a lot of research in this area, and none of it is cited here.

Reply 6: We rewrote this part and added related citations.

Comment 7: The resulting RNA-seq data needs to be made available in a public repository such as GEO, so the results can be validated. An accession should be provided for the data.

Reply 7: We are now depositing the original data into GEO.

Comment 8: It is very odd that there are references, but none of them are in-text citations. This should be rectified.

Reply 8: We corrected it.

Comment 9: On line 18, RNA Sequencing is defined as RNS-Seq, while on line 21, it is defined as RNA-Seq. Maintain the latter and remove the former.

Reply 9: We fixed it.

Comment 10: On p. 4, line 13, a 10-g rod was used, which induces a mild injury; however, the authors describe this as a severe contusion.

Reply 10: Thanks for pointing this out. We corrected with a 25g-rod in the revised version.

Comment 11: How many rats (n) were used at each time point, and at each location?

Reply 11: Post hoc analysis included 4, 4, 4, and 4 animals for sham, 1d, 4d, and 7d post injury, respectively.

Comment 12: Fig 1B is not a Venn diagram – it is an Euler diagram.

Reply 12: We corrected it in legend.

Comment 13: Fig 1A and 1C leads to a question about whether the data was properly normalized, since it seems to be skewed towards up-regulation. Since details about the data normalization are not provided, this cannot be properly assessed.

Reply 13: We don't think that the data were improperly normalized. We applied the same criteria to screen both up-regulated and down-regulated genes.

Comment 14: Given that the number of differentially expressed genes is relatively small, it is important to include the number of genes found for each of the significant KEGG pathways. If there are only four genes in a category, that does not necessarily mean that it should be concluded that the pathway is enriched, despite what the p-value shows. In fact, an FDR or q-value score should be used for these enrichments to correct for multiple hypothesis testing.

Reply 14: We agreed with the reviewer. Indeed, in supplemental table 4, we listed number of genes in each KEGG pathway, along with the FDR value. As the case of 4 genes, there were only 15 genes in that pathway and we don't think absolute numbers are criteria for the inclusion of particular pathways.

Comment 15: Figure 3 is very difficult to read – the text is too small to pick out what the genes are. In addition, no details are given to how the graphs are constructed. Different colors besides red and green should be used here, to make them color-blind safe.

Reply 15: We enlarged fonts of Fig. 3 and used different colors (magenta and blue) now.

Comment 16: For supplemental table 1, a p-value and q-value/FDR must be given – FC is not enough for differential expression.

Reply 16: We now listed the p values in supplemental table 1.

Comment 17: Table S3 is cut off, so I am unsure about what it contains

Reply 17: We fixed this issue.

Comment 18: Where do the PathwayID field come from in supplemental table 4? What program was used to generate these? What is the significance of these enrichments? What are the FDR values?

Reply 18: The Pathway ID was generated using fisher pathway. All listed enrichments were statistically significant between individual conditions (e.g. 0d vs 1d), according to FDR values. However, we did not claim that a higher enrichment fold corresponds to bigger importance. FDR values as adjusted p value reflect significance of predicted involvement of selected particular pathway(s) among other pathways.