Author's response to reviews

Title: Molecular subtypes of urothelial carcinoma are defined by specific gene regulatory systems

Authors:

Pontus Eriksson (pontus.eriksson@med.lu.se)
Mattias Aine (mattias.aine@med.lu.se)
Srinivas Veerla (srinivas.veerla@med.lu.se)
Fredrik Liedberg (fredrik.liedberg@med.lu.se)
Gottfrid Sjödahl (gottfrid.sjodahl@med.lu.se)
Mattias Höglund (mattias.hoglund@med.lu.se)

Version: 2 Date: 17 April 2015

Author's response to reviews: see over
Dear Editor

We thank the reviewer for taking their time to critically read our manuscript. Both of the reviewers argue that the original manuscript needed extensive editing and language control. We have performed an extensive rewriting of the manuscript and particularly made sentences shorter and arguments more to the point. The present manuscript is 12% shorter without deleting any of the original information, data, or conclusions. We hope that these changes will answer to the editor’s demands.

Answers to reviewer 1

Previously, the authors have described three major subtypes of urothelial carcinoma, and the present work they aimed to gain insight into UC biology by applying molecular stratification. Finally, they revealed that Urobasal, Genomically Unstable and SCC-like tumors represents three fundamentally different tumor types at the gene regulatory level. Several transcription factors were uncovered for each subtype separately. In addition, they further explored the protein expression of the associated differentiation factors across UC subtypes using IHC. Comprehensively, this is a well-written paper containing interesting results which merit publication. Nevertheless, for the benefit of the readers, several points need clarifying and several spelling mistakes and punctuations should be revised. There are given below:

1. Why you selected 0.4 works as the threshold of Quality Threshold Clustering (QTC) in Statistical Analysis? Any reference supported?
   **Answer.** The setting 0.4 is a compromise that works well according to our experience. In the end the setting is dependent on 1) the nature of the data and 2) what specific questions you ask. By using 0.4 we obtained a reasonable number of clusters with reasonable sizes. The critical end point is how genes in the clusters behave in a heat map.

2. The author have applied IHC to further investigate the expression of these differentiation factors, and further studies will be continued on the function tests for these findings.
   **Answer.** Yes we agree and this is now noted in the conclusion section.

3. Spelling mistakes and punctuations, such as 1) Urothelial Carcinoma (UC) arises from the urinary bladder epithelium that consist of three cell layers; (we can use “:” instead) basal cells, transiently amplifying cells, and differentiated umbrella cells; 2) PPARG and RXRA cooperate as heterodimers and induce differentiation of the urothelium upon activation and coordinates gene expression though (through?) the transcription factors ELF3, FOXA1, and members of the GATA transcription factor family.
   **Answer.** Se comment on top of the rebuttal letter.
Answers to reviewer 2

The data analysis and new information are comprehensive and important to the basic research in urothelial cancers. Extensive analysis has been performed; however, an extensive re-write is recommended to more clearly define the goals and the conclusions. A more grammatically correct and concise version of the manuscript, especially the discussion, would significantly increase the clarity of the significance and conclusions.

Answer. Se comment on top of the rebuttal letter.

Specifically:
Minor essential revisions

1) In the methods section
   a) provide a rationale for the cohorts chosen,
   Answer. We have used the Chungbuk dataset in previous investigations as validation data in previous investigations, e.g., PMID: 22553347. The TCGA bladder cancer dataset is the largest MI dataset for bladder cancer. The breast cancer TCGA dataset was chosen due to its thorough annotations. We believe that this information is not really necessary for the reader and will burden the materials and methods section with extra unnecessary information. The Materials and Method section provide exact information on the datasets that were used in the investigation.

   b) note that TCGA data are specific to primary tumors,
   Answer. The Lund 308 cohort of urothelial carcinomas does not contain any metastases, see PMID: 22553347

   c) clarify the other datasets’ derivations and limitations (ie the Lund 308 – are these primary, metastatic, ??),
   Answer. The Lund 308 cohort of urothelial carcinomas does not contain any metastases, see PMID: 22553347

   d) explain your specific exclusions.
   Answer. The exclusion of the “infiltrated” subtype from the Lund 308 cohort is now explained.

2) Results section:
   a. Section “Stat3 regulated gene signature...” states “we could establish...” and references a figure but provides no verbiage documenting how this can be established. Supporting explanation is needed.
   Answer. This sentence is now deleted

   b. Conclusions are often placed in the results section. Move to discussion/conclusions.
   Answer. All conclusions in the results part have now been deleted and moved to discussion/conclusions.
3) Discussion section:
a. Authors note that ‘RAR expression did not vary across the data’ but do not explain in context to subtype differences or comment regarding potential downstream differences. Additionally, ‘across the data’ is vague. Be specific to what is being compared, please.
Answer. The sentences describing the RAR expression has now been rephrased.

b. Statement ‘from a clinical perspective FOXM1 may be involved in, and create, the known aggressiveness’ is also vague and may be speculative, as the comment is not supported, nor is the statement further documented.
Answer: this section of the discussion have now been rephrased.

c. Perhaps break the discussion into biomarker-focused and subtype focused sections for more clear focus of the key points. Currently, this section is very hard to follow.
Answer. As noted at the top of the rebuttal letter both the results and the discussion part of the manuscript have been extensively edited.

d. Last sentence of discussion is not a section concluding statement.
Answer. Last sentence is now a section concluding statement.

4) Conclusion section:
a. Incomplete.
Answer. The conclusion section has been extensively edited.

b. Provide context for what these different subtypes mean in terms of treatment/biomarker ‘targetability’.
Answer. The aim of the present investigation was to identify gene regulatory systems operating in urothelial carcinoma. Treatment and targetability has been previously addressed by us in PMID: 22553347.

c. No limitations noted.
Answer. Limitations are now noted.

d. No mention of the comparison to breast cancer data.
Answer. This comparison is now mentioned.

e. No clear wrap-up.
Answer. The conclusion has now a wrap-up.

5) Comparison to breast cancer. Mentioned in methods, noted again in discussion, but not in conclusions.
Answer. The comparison is now noted in all sections.

Please provide more context regarding why only breast cancer, basal subtypes, is compared to your analysis, including any supporting data that would indicate this is the only other type of cancer with similarities. Otherwise, analysis is incomplete.
Answer. The reviewer has missed that we in addition to breast cancer basal-like also compare our data to breast cancer LumA, LumB, HER-enriched as well as lung squamous cell carcinoma. The major aim of this manuscript was to describe the gene regulatory systems in urothelial carcinoma. In doing so we could not avoid to see the similarities to some other well-known cancer types, and we highlight these similarities when we find them informative. We are fully aware that these analyses are not complete and refer to publications by Prat et al and Hoadley et al for more thorough comparisons.

6) Grammatical errors and vague sentences, unclearly defined sections, missing words, etc. decrease the ability to fully analyze the novelty, importance of the work. Example errors include incomplete sentences, non-parallel sentence structure, and incorrect word usage.
Answer. See comment on top of letter.

While the overall content may provide new information in the field, the manuscript needs to be rewritten to increase the clarity of the content, and the methods must be further defined (#1
Answer. See comment on top of letter.