Peer Review File

Article information: https://dx.doi.org/10.21037/tau-23-130

Reviewer A

Comment 1: This study reports the identification of protein markers in the seminal fluid of a small cohort of male infertile patients with low sperm counts. Similar studies have been reported, and the idea is not new. However, the authors did not thoroughly discuss how their findings have added new insights. The authors also did not comment on whether their approach has validated any of the previous findings or has generated conflicting results. There is also a concern about the experimental results. In view of these concerns, the manuscript is not suitable for publication in its present form.
The idea of identifying biomarkers for low sperm count associated with male infertility in the seminal proteome is not new. How are the findings reported in this study compared with what is known in the field? Is the proteomic approach in this study more sensitive and thus provides a better resolution to the seminal proteome? Are the markers identified specific to races or geographical locations of the study cohort?
Reply 1:We do agree that the idea of search for seminal plasma biomarkers is not new, however the list of studies reporting on putative biomarkers and referring to spermatogenesis status (not to vasectomy or oxidative stress or asthenozoospermia) is rather short (Davalieva at al. 2012, Drabovich et al. 2011, Milardi et al. 2012, Yamakawa et al. 2013, Cui et al. 2018). We have collected samples representing different entities of spermatogenesis impairment (severe oligozoospermia – cryptozoospermia – non-obstructive azoospermia) and analyzed them in one step proteomic approach which is quite unique. In addition, some of the previous studies differed in their methodological approach. As seminal plasma is rich in proteins, the application of highly robust quantification methods enables good discrimination between true and false positive results and thus, proper identification and quantitation of proteins. We do not consider our results as racial or geographical-specific. Both, MS and Western blot analyses showed that the results have been heterogenic within one group what suggests to search for seminal plasma biomarkers in big cohorts than using e.g. 4 samples per group as most studies did. We have added a paragraph in the Discussion section comparing our results with the results of other relevant studies.
Changes in the text: see Page 13, lines 377-389.
Comment 2: The result that ADGRG2 was missed by the quantitative proteomic analysis has raised concerns about the sensitivity and resolution of the approach. In fact, only very few DEPs were identified.
Reply 2: Indeed, there were only few DEPs identified and the quantitative analysis for ADGRG2 was not perfect due to low number of peptides detected during mass spectrometry. Perhaps, the application of multiple reaction monitoring mode for this protein would give more satisfactory results. In routine mass spectrometry analyses like ours, data dependent analysis (DDA) mode is usually applied which prefers the most abundant proteins for their fragmentation. Low abundant proteins can be missed and in

this case, ADRGR2 is low abundant and its detection was at the sensitivity threshold thus, we could missed some of them. Looking at the raw MS data (Supplementary Table 1, row 254), it can be visible that MS results are rather compatible with Western blot images and that rather only samples having intense Western blot signals were also detected in MS analysis. The names of individual samples have been added to Western blot images to enable the comparison of the results.

Regarding only few DEPs identified, the application of developed quantitative approach could have narrowed the number of DEPs. However, such approaches have been implemented in many other studies and were proved to eliminate false results. It is possible, that less abundant proteins were missed during further analysis. For this purpose, a different proteomic approach should be applied in our view.

Comment 3: The String analysis is merely a database showing the proposed protein interactions. The results should not be present as evidence of protein interaction without experimental validation. Moreover, even if the interactions can be validated, how the interacting partners implicate in SO, C and NOA remain elusive. Therefore, I do not think that the in silico protein interaction analysis has added much to the conclusion.

Reply 3: Thank you for the comment. Our intention was to show that the identified DEPs do not rather create any functional network and seems to be a result of individual mechanisms. We have eliminated this analysis from Methods and Results sections, however, left the Figure for the. better interpretation of results

Changes in the text: see Page 6, lines 163-165, Page 8, Lines 203-204.

Comment 4: The background of the Western blot images, especially for ADGRG2 and SMS, is variable. This would cause an artefact in quantifying the band intensities.

Reply 4: Low concentration of ADGRG2 can explain to some extend the high background in Western blot images. The apparatus acquiring images adjusts the intensity based on the saturation from the luminescent signal. It is well known fact in Western blot analyses that if the protein amount is low to be detected, the background increases. We do not know what was the real reason of high background in our experiments however, the background was not too high to prevent from an analysis. For each row, the protein signal was normalized against total proteins visualized in stain-free gels. It is also worth of notice that in some cases up to three different proteins were detected in the same PVDF membrane resulting in different background intensity after immunodetection. The subtraction of the background was done manually using the Image Lab software (BioRad Laboratories) based on the signal intensity histogram and was done for each protein band individually.

Comment 5: In Figure 4, are the error bars representing S.D. or S.E.M.? It is surprising to see the significance of such marked variations in ADGRG2.

Reply 5: Thank you for the comment. The error bars indicate S.D. and the information has been added in the Figures description. Indeed, for ADGRG2 the signal intensity was very diverse however, it was rather compatible with MS results. Western blot is more sensitive and specific technique giving higher resolution in comparison to routine MS analysis. All Western blot calculations including normalization and background subtraction were manually revised to deliver reliable results. As already written, the protein signals were normalized against total protein in stain-free gels eliminating

loading error. Such variations in a protein abundance cannot be neglected and in our opinion, provide knowledge on seminal plasma proteome. Cell proteome is dynamic and is stable for only few proteins. Protein composition of any plasma can be even more dynamic, especially of seminal plasma as its content is a combination of different secretomes. We do not know anything about the secretory ADGRG2 protein and mechanisms of its secretion. The variances between control samples are a fact and at this point it is not known why. Thus, we have named our study as a pilot one.
It is worth to add that ADGRG2 has been also identified as downregulated in seminal plasma samples from patients with non-obstructive azoospermia, Sertoli-cell only syndrome and mixed atrophy in PhD dissertation of Raouda Sgaier, supervisor PD Dr med. Thorsten Diemer form Justus LiebigUniversity, Giessen. In that work, the variances in ADGRG2 expression were also visible in Western blot images. However, this data have not been published yet and the thesis are available online. It indicates that our analysis is robust.
Changes in the text: see Page 19, lines 635 and 640.

Comment 6: Also in Figure 4, the asterisk refers to the internal reference sample. Is it the same sample being used for all four groups?
Reply 6: Yes, the asterisk indicates the control internal sample. We tried to stick to one sample included for all the four groups and the most constant was the control sample K11. However, for SMS and PAEP proteins other samples displaying an average signal were chosen, K6 and K3 respectively. It has been corrected for Figures 4 and 5. Also, in Materials and methods section, the sentence concerning the reference has been corrected.
Changes in the text: see Page 7, lines 179-181, Figure 4, Figure 5.

Reviewer B

Comment 1: This manuscript aimed to identify proteins present in seminal plasma that correlate to specific causes of male infertility using mass spec and Western blotting. The experimental design and approach is strong, yet there are many limitations and the conclusions are quite weak.
-The main limitation is the sample size (10 or less samples per group). Was there a power analysis conducted to determine if this was a reasonable sample size?
Reply 1:We have not calculated the sample size with power analysis. The group had as many samples as there were available. We took care that the group was not less than four samples like it was published in the other reports..
Comment 2: Additionally, there are some questions regarding methods. Were the Western blots performed on the same samples as those used for mass spec? Lane numbers/labels are missing, what were the controls for the blots, how were samples normalized (load controls?), what is meant by “internal ref sample” and where are the size markers? This data is presented very poorly.
Reply 2: Yes, for Western blotting the same samples were used as for MS. For all four groups the same control sample was used as a reference sample serving also as a loading

control for normalization of total protein loading and then, the intensity of the signal from protein bands. The signal from the reference sample was set as value 1 and the signal from other samples within the group were proportionally calculated as values lower or higher than 1. The calculations were done using Image Lab software (BioRad Laboratories) dedicated for such analyses. The signals for the all investigated proteins were detected within the proper molecular weight range indicating no false positives. We have corrected the Figures according to the remarks.
Changes in the text: see Page 7, lines 179-181, Figure 4, Figure 5.

Comment 3: More broadly, if you refer to Figure 2, there really are no proteins unique to each type of male infertility investigated. This could be an interesting finding. To investigate the authors should have assessed the mRNA levels for the proteins found in control samples in the male infertility samples. At least the authors would know if this is a protein degradation issue or a potential genetic mutation.
Reply 3: Thank you for this comment. Severe oligozoospermia, cryptozoospermia and non-obstructive azoospermia are concerned to have similar etiology that is either genetic factor (chromosome Y microdeletions) or hormonal disorders. However, in many cases the etiology is not known what is visible in diverse responses of patients to spermatogenesis-stimulation therapy. Recently, oxidative stress has been taken into account as a factor impairing sperm development. We assume that these three groups of patients (or at least some of the patients within the groups) may be characterized by the same causative factor however, of different intensity or spreading different effect depending on other molecular mechanisms taking place in a developing germ cell. Investigation of alterations at mRNA level is a good point to study the mechanisms in details however, due to very low number or lack of spermatozoa in the semen samples it forces us to use testicular mRNA. On the other hand, as testicular biopsy is more advanced procedure for diagnosis, the access to such biological material in Poland is very limited. Our possible results aiming in this direction can be therefore implemented in future studies. Thus, the suggestion to treat the presented results as “a pilot study” (as suggested in the next paragraph) is an adequate option.
Comment 4: The authors also do not ever really discuss the different causes of male infertility in Intro.
Reply 4: Thank you for the comment. We have added a new paragraph introducing to the infertility.
Changes in the text: see Page 3, lines 56-65.

Comment 5: Overall, this study appears more of a “pilot” study in my view, and maybe if presented as such, this would be more suitable for publication once other issues are addressed. I might even suggest an edit to the title to state that this is a “pilot” study.
Reply 5: We do agree with this suggestion. The title has been edited accordingly.
Changes in the text: see Page 1, lines 1-2.

Reviewer C

Comment 1: This is a very interesting and well-designed study, the results of which can be very useful for the diagnosis of different types of alterations in spermatogenesis. However, although in the case of patients with azoospermia the authors specify that it is non-obstructive, the type of oligo or cryptozoospermia has not been classified. The results would be different if in the latter cases it was obstructive or non-obstructive. For this reason, the authors should report the type of these two pathologies.

Reply 1:Thank you very much for the suggestion on the anatomical etiology of the disorders.

Changes in the text: see Page 5, line 121.