Peer Review File

Article information: http://dx.doi.org/10.21037/jtd-20-2396.

Reviewer A

Comment 1: I suggest to the author to better explain "biomarkers research" and their importance in this disorder. I also suggest to introduce the problem of a current validated biomarkers (such as ACE2) that results low specific and sensible.

I also suggest to added the role of CD4/CD8 ratio during diagnostic work up of sarcoidosis patients.

Reply 1: We agree with your remarks and suggestions. We have made appropriate changes in the Introduction. Accordingly, we added citations from suggested articles.

Comment 2: However, many inflammatory cells and tissue cells with various, sometimes even opposite effects may be positive for CD31, CD38, CD44, and CD103", also in this case I suggest to the author to better explain the pathogenetic mechanisms that allow the development of sarcoidosis patients and the role of these proteins.

Reply 2: We agree with your comment. We have made appropriate changes in the Introduction. Accordingly, we added citation from suggested article.

Comment 3: I also suggest to added in the gating strategy, some references that can explain the choose of monoclonal antibodies and their role.

Reply 3: Thank you for the comment. We had added explanation for the choose of monoclonal antibodies and supplemented Figure 1. We added to references suggested article.

Comment 4: Can the authors added the results of BAL analysis performed for diagnostic work up? How many lymphocytes are present in the BAL? how many macrophages? CD4/CD8 ratio?

Reply 4: We fully agree with your comment. We added these data to Table 1 in Materials and methods.

Comment 5: Figure 1. Typical histograms for CD31, CD38, CD44, and CD103 expression on CD4+ and 160 CD8+ lymphocytes in blood and bronchoalveolar lavage fluid" need to be improved. The dot plot are shameful.

Reply 5: We fully agree with your comment, unfortunately flow cytometer had used is old enough and we do not have big choice. We did our bets changing the Figure 1. We had also changed the title of the Figure 1.

Reviewer B

Comment 1: What are the main results and message of this research? Authors concluded lymphocyte subsets are different between the patients and healthy controls. However, it is obvious. Because lymphocyte activation occurs in the disease.

Reply 1: Thank you very much for your comment. We appreciate your point. However, we do not agree with it. Since sarcoidosis manifestation is different, the difference in lymphocyte subsets is not obvious. In fact, a great part of the patients is asymptomatic at the time of diagnosis. Some of them without any lung parenchyma changes (stage I) or more or less fibrotic changes (stage III and stage IV).

The main results and message are presented in Conclusions: "… the lymphocyte subtype profiles were significantly associated with the clinical manifestation, radiological sarcoidosis stage, and smoking status. The lymphocyte profiles in BALF and lung tissue did not correlate, which may indicate the need for assessment of both profiles in further studies. Increase of CD3+CD4+CD31+ in BALF may serve as supporting evidence for a diagnosis of sarcoidosis. Increase of CD3+CD4+CD38+ in BALF and blood and CD3+CD4+CD44+ in BALF – are markers of acute immune response in sarcoidosis. CD4+103+ T cells in BALF and in blood are markers of a persistent immune response in sarcoidosis patients and are a probable prognostic factor of the chronic course of the disease".

Comment 2: Results and discussion paragraph are redundant, and it is hard to understand the point of the research.

Reply 2: Thank you very much for your comment. We agree, that there are a lot of results – number of markers from different compartments. So results and discussion are complex and may be difficult to read. We have made appropriate changes.

Comment 3: Why did authors choose these markers among markers of lymphocytes?

Reply 3: Thank you very much for your question. Chosen markers of lymphocytes were selected since at the time of initiation of this work we believe that they may have potential role in ongoing inflammation ant fibrosis development. Introduction was supplemented.

Comment 4: How did the authors diagnose the patients without pathology? Only 30 patients were diagnosed the disease in whole group.

Reply 4: Thank you very much for your comment. We appreciate your point. However, we disagree with it. All our sarcoidosis patients were diagnosed accordingly to international (WASOG) and national guidelines. All the patients were follow-up carefully and no one alternative disease for the study patients was revealed.

Comment 5: In table 1, the serum parameters are lacking. e.g. ACE, sIL-2R.

Reply 5: Thank you very much for your comment. We did not investigate these parameters in this study. Moreover, there are a lot of publications concerning the role of these parameters. They did not show good sensitivity or specificity as predictor markers for sarcoidosis. Introduction was supplemented with the information on serum markers.

Comment 6: The histological control samples seem to be inadequate. Because the samples were collected from the patients with sarcoidosis and alveolitis seem to be observed in some control samples (Figure 2G-L).

Reply 6: Thank you for your comment. We agree that the Figure 2 is confusing. It has been edited to represent comparison between sarcoid and control groups better, pictures were edited to equate magnifications and legend was changed accordingly. Additionally, controls used for immunohistochemistry were added to supplementary appendix.

Comment 7: Why did the authors divide the patients into smokers and non-smokers in Figure 6-7. Why did not in Figure 4-5?

Reply 7: Thank you for the comment. Unfortunately, patient population was not big enough – dividing sarcoidosis stages groups into smokers and non-smokers we would have too small groups for statistical calculations.

Comment 8: How about the correlation between expression levels of these markers and serum parameters?

Reply 8: Thank you very much for your comment. Correlations between blood, BALF, and lung tissue lymphocyte subsets are presented in separate paragraph of Results. We did not investigate serum parameters in this study.

Reviewer C

Comment 1: The title: the authors analyzed the marker expression on lymphocytes but no other cells it should be specify in the title.

Reply 1. We agree with your remarks and suggestions. We have made appropriate changes in the title.

Comment 2: The Table 1 lack any type of analysis.

Reply 2: We agree with your comment. We added these data to Table 1 in Materials and methods.

Comment 3: What is the characterization of control group? Fibreoptic bronchoscopy was performed in healthy people or patients with other than sarcoidosis diagnosis? What were the results of spirometry of plethysmography in controls?

Reply 3: Thank you very much for your comment. Fibreoptic bronchoscopy was performed in healthy volunteers. We added data to Table 1 in Materials and methods.

Comment 4: There is a mistake and discrepancy in methodology vs data presented in Figure 1 due to fluorochrome conjugated to antibody in flow cytometry. Line 139-141 vs axis x and y in Figure 1.

Reply 4: We are terribly, sorry, you are right. Data in Figure 1 are true, our mistakes were in the text.

Comment 5: How did authors count CD4/CD8 ratio if both CD4 and CD8 were conjugated to the same fluorochrome and were analyzed in a separate probe? Positive vs negative? Were the calculations the same for CD4 and CD8?
Reply 5: Samples were gated on CD3+ PerCP, then from one probe percentage CD3+PerCP and CD4+ PE was received, and from other probe – CD3+PerCP and CD8+PE.

Comment 6: Why immunochemistry for CD4 was performed on different visualization system than CD8, CD38, CD44 and CD131? Why the authors did not evaluated CD31 in immunohistochemistry?
Reply 6: According to our internal protocol we use DAKO visualisation system for CD44 immunohistochemistry for all tissues, except bone marrow samples. This rule is applied to achieve stronger staining after decalcification procedure in trephine biopsies. We did not have enough financial resources to evaluate all markers.

Comment 7: The authors should present exact p-value for data presentation.
Reply 7: Thank you very much for your comment. Exact p-values added to figures makes them too chaotic, because of that p-values added to the text.

Comment 8: The result description is really complex and hard to follow, for better study presentation it should be shortened.
Reply 8: Thank you very much for your comment. We agree, that there are a lot of results – number of markers from different compartments. So results and discussion are complex and may be difficult to read. We have made appropriate changes: we had shortened the text where it was possible, also divided Results to sections.

Comment 9. The axis in figures should be corrected – the y axis should be described as CD3+CD4+ cells or CD4 lymphocytes or analogously for CD8.
Reply 9: Thank you for the comment. Figures were changed accordingly to your suggestion.

Comment 10: The numbers in figures should have dots not commas.
Reply 10: Thank you very much for your comment. Figures were changed accordingly to your suggestion.

Comment 11: The authors describe in results lymphocyte count, lymphocyte numbers but I assume that the data are presented as percentage of positive cells in selected gate.
Reply 11: Thank you, you are right. The text was corrected whenever possible.

Comment 12: Table 2 – the authors in line 347 wrote compared to non-smoking controls. The total number of controls is 5 what is the number of smoking and non-smoking controls? The calculation in such small groups should not be statistically analyzed.
Reply 12: Thank you very much for your comment. All 5 controls (healthy volunteers) were non-smoking. The data presented in Table 2 are intermediate and were not used anywhere else.

Comment 13: The authors presented selected (most important) correlations. What was the criterion for selection of correlation? The strength (r value?) maybe the authors should considered presented all obtained correlations in supplementary file.
Reply 13: Thank you very much for your comment. We have made appropriate changes. We added correlations in the Supplementary file.

Comment 14: Discussion line 393 "The most interesting and promising results were obtained for …." That were all results presented in this manuscript. Authors should critically selected the most interesting results and highlighted them in principal findings in discussion.
Reply 1: We agree with your comment. We have made appropriate changes in the introduction part of the Discussion.

Comment 15: The discussion section should be rewritten: it is too long, contains dominating amount of information that should be presented in introduction and too many repeated results. Please, carefully select the most pronounced results, frequently do not repeat the values from

results and describe the clinical and biological meaning of presented data as well as compare the results with other authors. In fact the discussion from line 563 is properly written.

Reply 15: Discussion was shortened deleting the text, please see changes in the text.

Reviewer D

Comment 1: In the results paragraph on the abstract, it is not clear the which are the main lymphocyte profiles identified and which clinical correlations were found. Also, the conclusion in the abstract does not appear to reflect the main focus of the study.

Reply 1: We agree with your comment. We have made appropriate changes in the Abstract.

Comment 2: The paper contains a lot of valuable data and is very comprehensive, which makes it in some are difficult to follow, especially regarding the comparison between the different compartments (blood and BALF) and the subgroups of patients (smokers/non-smokers, HC, Löfgren and non-Löfgren).

Also, the results are described in different order between groups, that require the reader going back and forth between main manuscript and figures (specially figure 4 and 5). It would be better to have more consistency when presenting the data (sarcoidosis or controls first).

Reply 2: Thank you for the comment. It would be difficult not to agree, we had tried to stucturaise as much as possible. We had shortened the text where it was possible, also divided Results to sections.

Comment 3: In figure 2, the authors should consider a more organized arrangement with a consistent magnification of the pictures since it is confusing the comparison between the groups. The legend should explain better the purpose and findings on the images. Also, an isotype control for the stainings should be added in the supplementary materials.

Reply 3: Thank you for the comment. We agree that the Figure 2 is confusing. It has been edited to represent comparison between sarcoid and control groups better, pictures were edited to

equate magnifications and legend was changed accordingly. Additionally, controls used for immunohistochemistry were added to supplementary appendix.

Comment 4. In the statistical analysis, how did the authors evaluate the distribution among the groups and skewness? There are small numbers in some of comparison groups with high SD which makes the use of non-parametric analysis and the use of median and IQR more appropriate in this setting (for example, figure 4 and figure 5 and table 2). State if the authors corrected for multiple comparison and which type of analysis they used for comparison of paired blood and BALF samples. Two-sided P values would also be preferred.

Which type of statistical correlation was used in the analysis? Add exact P value, IC 95% and scatter dot plots (can be added in the supplementary materials)

The authors comment the calculation of odds ratios to estimate the relationships between the risk factors during the initial research and observation of disease progression. However, there is no information regarding this value in the results of the manuscript.

Reply 4: Thank you very much for your comment; you had pointed out an important aspect. We had recalculated all results using non-parametric analysis and made changes both to the text and figures. Considering calculation of "odds ratios to estimate" - it was our mistake – we are preparing other article on long term results, and mixed up the description, sorry for that and thank you for correction. We added correlations in the Supplementary file.

Comment 5: The use of terms expression (eg line 255) should be changed to frequency or percentage so it is not confused with MFI. Also, line 275 should be rephrased "more positive lymphocytes" into a more appropriate term.

Reply 5: Thank you, you are right. The text was corrected whenever possible.

Comment 6: How do the authors define "activated cells" and what is the role of CD31 as T cell activation marker? (line 281 and 320). As the authors refers in the discussion, CD31 is downregulated after T cell activation, then how could the patients in the smoking non-Löfgren

group had more activated cells? (lines 417 and 419). It would be good to use other common markers of T-cell activation and, instead of presenting the population as percentage or expression, present the populations as -low, -intermediate, or -high expression of these markers.

Reply 6: Thank you, you are right. The text was corrected whenever possible.

Comment 7: The discussion is largely a restatement of the results with little discussion of the significance and implications of the presented findings. Clarifying the key findings and their clinical or mechanistic implications (eg. Pulmonary function test indices) would significantly increase the impact of the work.

Reply 7: Thank you very much for your comment. We agree, that there are a lot of results – number of markers from different compartments. So results and discussion are complex and may be difficult to read. We have made appropriate changes: we had shortened the text where it was possible, also divided Results to sections.

Major points

Comment 1: While a significant amount of information is provided about the flow cytometry analysis there is some critical information that is either missing from the main text or difficult to locate.

Reply 1: We have made appropriate changes in the Introduction and improved Figure 1 (we are sorry, but instrument we are working on is old enough ant we were using three colors only). We supplemented the Introduction with suggested citations and excerpts from the Discussion.

Comment 2: Please include a representative gating strategy in the supplementary material showing the selection of lymphocytes, life-dead cells, single cells, and then specific T cell populations for both PBMCs and BALF. The graphs presented do not show the positive and negative populations, specifically for the markers CD31, CD38, CD44. It should be stated which scatterplots belong to Blood and BALF in the graphs. Also, in material and methods, the authors

describe PE-conjugated anti-CD31, -CD38, -CD44 and -CD103, while in the legends of the Figure 1 it is presented as FITC.

Reply 2: We have made appropriate changes in the Introduction and improved Figure 1 (we are sorry, but instrument we are working on is old enough ant we were using three colors only). Considering fluorochromes, we are terribly, sorry, you are right. Data in Figure 1 are true, our mistakes was in the text. We supplemented the introduction with suggested citations and excerpts from the Discussion.

Comment 3. Since the authors are analyzing the immune profiles of 2 different compartments, it would be also be of value to add markers for central memory, effector memory and tissue resident memory T-cells, and stratify the proposed biomarkers according those subgroups.

Reply 3: Thank you for proposal for the future research.

Comment 4. It is not clear if they are referring to total number, percentage or expression since they use it similarly throughout the paper. The authors should define a term and be consistent with it through the paper.

Reply 4. Thank you, you are right. The text was corrected whenever possible.

Minor comments:

Comment 1: Check the p value in line 259.

Reply 1: p for stage II vs. control 0.028, for stage III – 0.054.

Comment 2: What do you mean by "relevant group of smokers" in line 280?

Reply 2: Sorry. Löfgren group of smokers.

Comment 3: Typo in line 287: CD4CD44+-

Reply 3: thank you, corrected.

Comment 4. Line 291: which compartment comparison are the authors using? Blood or BALF?

Reply 4: We are sorry. BALF. Corrected.

Comment 5: Why were patients with Löfgren syndrome excluded from the lung tissue analysis?

Reply 5: We have no lung biopsy tissue preparations for sarcoid patients with Löfgren syndrome (lung biopsy was performed only for patients without Löfgren syndrome).

Comment 6: Is there an association with granuloma formation and CD31 marker in paired BALF and biopsy?

Reply 6: We have no data about CD31 marker in lung biopsy tissue preparations. We did not have enough financial resources to evaluate all markers.

Comment 7: Since smoking is plays an important role in sarcoidosis and in the lung per se. Have the authors opportunity to look at BALF lymphocyte markers in patients that have COPD?

Reply 7: Yes, we have. However, not in this project.

Comment 8: Tables and graphs: Use the same scale as possible when comparing 2 or more graphs.

Reply 8: We are sorry, but it would be difficult while percentage of CD31+ is much lower comparing to CD44+.

Comment 9: Use consistently the ö/o in Löfgren.

Reply 9: Thank you for the comment. We corrected the Figures.