Author’s response to reviews

Title: Characterization of disease-specific cellular abundance profiles of chronic inflammatory skin conditions from deconvolution of biopsy samples

Authors:

Zandra Félix Garza (z.c.felix.garza@tue.nl)

Michael Lenz (mlenz@uni-mainz.de)

Joerg Liebmann (joerg.liebmann@philips.com)

Gökhan Ertaylan (gokhan.ertaylan@maastrichtuniversity.nl)

Matthias Born (matthias.born@philips.com)

Ilja Arts (ilja.arts@maastrichtuniversity.nl)

Peter Hilbers (p.a.j.hilbers@tue.nl)

Natal van Riel (n.a.w.v.riel@tue.nl)

Version: 1 Date: 03 May 2019

Author’s response to reviews:

Manuscript: ‘Characterization of disease-specific cellular abundance profiles of chronic inflammatory skin conditions from deconvolution of biopsy samples’

Zandra C Félix Garza‡, Michael Lenz‡, Joerg Liebmann, Gökhan Ertaylan, Matthias Born, Ilja CW Arts, Peter AJ Hilbers, Natal AW van Riel.

‡ These authors contributed equally

Dear Chief Editor,

We have addressed the remarks and questions raised by the reviewers and modified the manuscript accordingly. Please find all our answers attached to this letter. Please note that all changes are highlighted in yellow in the manuscript itself. Additionally, we have verified that our revised manuscript adheres to the journal style.

Thank you for considering our manuscript for publication. We appreciate your time and look forward to your response.
Reviewer report- Mayte Suárez-Fariñas (Reviewer 1): The authors had generated an skin-specific gene signature to be used with deconvolution algorithms with the objective to derived estimates of cell-specific abundance from gene-array data. They evaluate the performance of their approach using reported summaries of cell abundance obtained by FACS and expression profiles of single cell and by reconstructing known biology of cell composition in different layers of the skin and in AD and psoriasis skin biopsies.

1. Overall the analysis is rigorous and the manuscript well written and an important result for the community. Authors should make the DERM22 signature available as supplementary material for this paper.

DerM22 is now included as supplementary material in the manuscript (TableS2).

2. Some of the nomenclature in the assessment of the performance in Figure 1 needs to be adjusted. The authors are showing overall mean fractions for each tissue/method (deconvolution or Flow). This only indicates if the two methods produce average estimates that are close. As such, the use of agreement/consistency is not appropriate from the statistical point of view. Agreement/consistency refer to how close each measure is across all samples. The average of two measures can be the same even if they are perfectly negative correlated (think for example: x=\((-10, -9, -1, 0, 1, 9, 10\)), y=-x, the mean is the same (0) but they actually disagree in all values but 0.

The nomenclature has been adjusted in the first paragraph of the results section as suggested, lines 209-228, particularly lines 214 and 228.

3. The performance analysis of Fig 1 is focused on showing that on average the cell abundance is similar. If gene array and immune-histo-chemistry exists for all samples in the 3 studies presented in Figure1B, could the authors formally assess the consistency (using ICC) of the results of IHC and deconvolution across all samples instead of just the average across groups? Ideally one would like to evaluate the deconvolution performance at the sample level, to for example, study heterogeneity of the cell population so showing the consistency. Of note it is not clear if the ratios plotted here are the ratios of the average for each subgroup or the average of the ratios taken at the sample level. It should be clarified.
We agree with the reviewer on the added value of including a sample level analysis on the data behind Figure1B. Currently, the ratios plotted in Figure1B correspond to the ratio of the average for each subgroup. This has been clarified in the text too, namely in lines 187, 239, and 251. These ratios were computed based on our deconvolution results and the values reported in the figures and text of the cited studies. Implementing the suggested analysis requires the sample specific immunohistochemistry data, which is only available upon request. In an effort to implement the sample level analysis, we requested the raw IHC data from the corresponding studies to implement the proposed ICC of the IHC results and the deconvolution across the samples. However, the conditions set by the data holder for the sharing of data were beyond what we were willing to accept at this stage of the manuscript and thus we did not include this additional analysis.

4. In Fig 1D the authors tried to assess the variance due to the platform used. I wonder if a principal variation analysis with factors platform and group will help quantify the variance due to the study and tissue/disease.

We performed a principal variation analysis as well as a principal components analysis based on the complete gene expression data, revealing a strong batch effect distinguishing two groups of studies on the first principal component (see Figure R1).

(For details on this figure please see .doc version of this letter)

Figure R1: Principal components analysis and Principal variance component analysis on the complete expression dataset reveal strong batch effects for the expression data.

However, the estimation of cell type composition from expression data is only slightly (if at all) affected by these batch effects. This can be seen in Fig 1D and Supplementary Figure S2, which show only slight differences between studies. Furthermore, the observed pattern of differences in cell type composition between studies cannot be explained by the batch effects seen in the PCA, as the studies clustering together in the PCA do not cluster in a similar way when looking at the cell fraction estimates (Figure R2, left). Instead, the cell fraction estimates cluster largely according to disease status (Figure R2, right).

(For details on this figure please see .doc version of this letter)

Figure R2: Principal components analysis on the cell fraction estimates shows that the batch effects from the gene expression data are not transferred to the cell fraction estimates.

Nevertheless, it is in general advisable to focus on comparisons within studies for the interpretation of the results (e.g. as in Fig. 3C). In our analyses, we observe that such relative comparisons within studies are largely consistent across studies (Fig 3C: rows corresponding to same comparisons are largely colored in the same way).
5. The deconvolution of psoriasis samples largely reflects known biology, although the number of myeloid DCs are lower in lesional biopsies than non-lesional) but some findings in AD do not. Specifically, Keratinocytes and immune cells appear as less abundant in non-lesional skin than in healthy (Fig 2D, Fig3A,B). As the authors acknowledged this does not agree with other reports (see PMID:21388663, PMID: 22951056). In Figure 3B one can see that abundance of all cell types is low for non-lesional AD (Fig 3 A and B) as compared with healthy. As the estimates from Healthy in F3A,B are a pool from healthy samples from different studies, I ponder if batch effects are transferred into cell frequency estimates.

Although we showed above that batch effects are not significantly transferred into cell fraction estimates, we agree with the reviewer that detailed comparisons should ideally be done within studies. We did such a comparison in Figure 3C and additionally, it is possible to extract information on study-specific comparisons from Fig. 1D and Supplementary Fig. S2. These comparisons show that (i) the difference in number of keratinocytes as well as in the number of immune cells is not statistically significant when comparing healthy vs non-lesional AD samples (Fig. 3C) and (ii) there is a tendency towards a lower keratinocyte fraction for non-lesional AD samples compared to healthy in dataset GSE36842, while no trend can be observed for dataset GSE32924 (Fig. 1D). We have adjusted lines 428 to 433 in the discussion section to reflect this issue.

A possible explanation for these unexpected results are the structural changes that occur in the non-lesional epidermis of AD patients, namely the swelling of this skin layer, the weakened epidermal barrier, and reduced epidermal thickness. These result in an increased percentage of dermal skin tissue or even subcutaneous adipose tissue within the biopsy samples, and thus a lower abundance of keratinocytes.

6. Indeed if we look at F3C, where the comparisons to healthy seem to be made within each Geo series (and hence no batch /study effect), the non-lesional AD vs healthy comparisons have more color than the non-lesional psoriasis comparisons. The authors should comment on this and should openly indicate if batch effects are 'transferred' into the cell frequency estimates. In this regard, besides Fig2D, a PCA of all samples from hgu133plus2 color-coded by study/GEO series, and a PVCA of the healthy samples may be of help to indicate how much variance is explained by the study.

Indeed, the non-lesional AD vs. healthy comparisons have more color – hence are more different – than non-lesional PS vs. healthy in Fig. 3C. However, we do not see how this should contradict our observations in Fig. 3A and Fig. 3B, as non-lesional AD is also more different from healthy individuals (e.g. more adipocytes, less stem cells, less immune cells) than non-lesional PS from healthy.

7. It will be important to include all cell types in the last figure comparing the changes in cell abundance with NB-UBV. The original report shows statistically significant decrease in keratinocytes proliferation and many immune markers, including myeloid DC's, but I gather that there was no a significant decrease in keratinocytes here. It will add to the manuscript if the
authors use some of the data for biologics in psoriasis, where a decrease in keratinocytes should be seen.

Regarding figure 4, results on the other cell types can be found in the supplemental figure S4 of the manuscript. As pointed out by the reviewer, we did not observe any significant decrease for the keratinocytes. As for the inclusion of biologics data, we applied our approach to two datasets on biologics, i.e. GSE47751 and GSE117239, which assess the use of Etanercept treatment of psoriasis (Fig. 5 and supplemental figures S5 and S6). These datasets comprise expression data on biopsy samples from lesional skin before, during, and after treatment with Etanercept, as well as samples from non-lesional skin at baseline. As expected, our results (Fig. 5) showed a clear and significant (particularly for GSE117239) reduction of the keratinocytes fraction when treated with Etanercept. The estimated keratinocytes fraction at the end of the treatment was similar to the predicted abundance of this cell type in non-lesional skin. Cell types associated with inflammation also returned to values similar to those estimated for non-lesional skin (Supplemental Figures S5 and S6). (See paragraph on treatments, i.e lines 359 to 382).

Minor points:

- Figure 1. A. This figure is confusing. The bars mean either SE, SD or range. Why isn't this reported as a unique measurement? Especially the SE and SD transformation is just a factor of square root of the number of samples.

The results weren’t reported as a unique measurement because we were showing the measurements reported in the original studies. This has been changed in Fig. 1A and is now reported in terms of SE with the exception of the flow cytometry acquired pDCs value for lesional psoriasis skin since no measure of variation was reported in the original study.

- Line 250: it should be Spp Fig S2 not S3.

This has been corrected (see line 246).

- Line 256: "All data points calculated from our deconvolution" should be "all average fraction per groups from our deconvolution". Authors are not showing frequency of each sample, so not 'all data points'. Additionally why only the ratios are presented. It would be important to see if the estimated frequencies are close for each cell population.

Lines 252 – 253 have been modified accordingly. We decided to present only the ratios due to data availability constraints in the immunohistochemistry data.

- Line 291 and Figure 2: Should be said right at the beginning that those are healthy individuals, as it will aid the interpretation of the 79% estimate
This has been included in lines 291, 294,308, and 309.

- I suggest that the same color scheme for cell types be kept across figures

The color scheme has been changed for figure 3B for the sake of consistency.

- Figure 3A. The stacked plots in Figure 2 seem easier to compare than the doughnut plot.

The doughnut plot in Fig. 3B has been replaced by a stacked bar plot as recommended.

Reviewer report- Raghunath Chatterjee (Reviewer 2): The present manuscript entitled "Characterization of disease-specific cellular abundance profiles of chronic inflammatory skin conditions from deconvolution of biopsy samples" by Garza et. al. developed a signature matrix from 22 reference cell types and used it to quantify the cellular abundance in psoriasis and atopic dermatitis.

Authors should address the following comments:

1. The main contribution of this manuscript is the development of DerM22, however the matrix is not provided as a supplementary material. Authors should provide the matrix for the scientific community.

DerM22 is now included as supplementary material in the manuscript (TableS2).

2. Fig.1 A: Flow cytometry data for immune infiltrates in lesional skin of psoriasis showed 6 data sets. Are these different study samples? Two clusters are formed for these data sets, one comprising 2 datasets and another 4 datasets? One dataset overlap with the HG-U133 and another with the HG-U133 Plus 2.0. Are they different in terms of disease severity?

The points shown in Fig1A represent the average values reported in the original flow cytometry studies. These studies considered chronic plaque psoriasis patients, however little is reported on the severity of their psoriasis. Due to the lack of data on the severity of the disease, however likely, it cannot be concluded whether the range of severity levels considered in the studies is a factor that could explain the discrepancies between the data points. The 4 data points overlapping with HG-U133A (green square) come from two different studies, i.e. two data points from Glade et al. 1996 and two from Raxworthy et al. 1987. The two data points overlapping HG-U133 Plus 2.0 were reported in Glade et a. 1996.
2. Fig 1C: Authors interpreted that the keratinocytes identified using CIBERSORT are correctly identified as they showed good overlap with the Epidermal SCs, TAs and keratinocytes. However, I can see that the CD8+ cells also showed good overlap with different other cell types. Authors should justify this observation.

The purpose of figure 1C was to assess the cross-platform performance of the approach using data from Newman et al. As pointed out by the reviewer, expression data from CD8+ cells also overlapped with other cell types, however more than 80% corresponded to the actual CD8+ cells. From the remaining 20%, the majority of the overlap was with subsets of T cells, particularly T cells. The rest of the overlap was small and could be explained by the differences in microarray platforms. Despite the good performance of DerM22 when used to deconvolve expression data from a different platform, one should be cautious as cross-platform bias remains a potential error factor. The observation on the CD8+ has been included in the third paragraph of the results section. (lines 262-280). Additionally, a sentence has been added to the limitations paragraph were the cross-platform issue is addressed. This sentence warns the reader on the use of cross-platform data with this approach (line 487).

3. In basal and suprabasal epidermal sublayers, 12% adipocytes (10%) and adipose stem cells (2%) is very surprising. What is the proportion of CD8+ T-cells? There are reports that the CD8+ T-cells are abundant in epidermis, while CD4+ T-cells in the dermis. But I do not see such abundance in the observed results. Authors should discuss about this differences with reference to the published literatures.

We agree with the reviewer on how surprising the fraction of adipocytes is in the epidermal sublayers. This could be partially explained by the trans-differentiation of adipose stem cells into keratinocytes, described in the literature. The estimated fraction of CD8+ T cells in the skin was below 1% for both the epidermis and the dermis while the CD4+ were estimated at 2% for the epidermis and 4% for the dermis. In consistency with literature, we observe a higher fraction of CD4+ T cells in the dermis (Mueller et al, Front Immunol, 2014). Despite being predominantly found in the dermis, CD4+ T cells may also be present in the epidermis (Cheuk et al, J Immunol, 2014). Regarding the CD8+ T cells, our results disagree with the literature. Little information is available on the original samples within dataset GSE42114. Even though it is unlikely, there is a slight possibility that these cells were cut out of the epidermis during the sample acquisition given that they appear to be located at the borderline of the epidermis and the dermis in humans in healthy or psoriatic skin (Cheuk et al, J Immunol, 2014). The observations on the T cells are now discussed on lines 462 to 470.

4. Fig 3E is very difficult to read. Authors stated that there is a significant increase/decrease of some cell types with disease severity. They used Wilcoxon rank sum test, and without adjustment for multiple hypothesis testing. They should present the corrected P-values. How many samples for each category were considered? Looking into the variances in Fig 1, I have doubt about the significant differences for low abundant cell types after adjustment.
We considered 13 samples for severe and 14 for mild psoriasis, and 10 samples for severe and 37 for mild atopic dermatitis. Upon correction, severe cases of psoriasis showed a significant increase (corrected p<0.05) in the abundance of adipose stem cells and keratinocytes compared to mild cases. Further, a significant decrease (corrected p<0.05) in the relative fraction of adipocytes and endothelial cells was observed for severe psoriasis. In severe cases of atopic dermatitis, only adipose stem cells showed a significant decrease (corrected p<0.05) in their relative fraction. This is described in lines 349 to 358.

5. The validation part of the observed results is limited, it is just based on some IHC and limited flow cytometry data. Author should mention the limitation of their results in the discussion.

It has been added to the limitations paragraph in the discussion at lines 473 to 475.

6. There are some typos, e.g., "overestimated in health and disease"; page 12, line 318: "the layer-specific composition of the skin described in Fig. ,". Authors should correct these typos.

The typo on line 319 has been corrected. It was not clear to the authors what the typo in line 227 "overestimated in health and disease" was. The text has also been screened for other typos.