THE EMBO JOURNAL

Binary organization of epidermal basal domains highlights robustness to environmental exposure

Sangeeta Ghuwalewala, Seon Lee, Kevin Jiang, Joydeep Baidya, Gopal Chovatiya, Pritinder Kaur, David Shalloway, and Tudorita Tumbar
DOI: 10.15252/embj.2021110488

Corresponding author: Tudorita Tumbar (tt252@cornell.edu)

Review Timeline:

Submission Date:	19th Dec 21
Editorial Decision:	17th Feb 22
Revision Received:	15th May 22
Editorial Decision:	1st Jul 22
Revision Received:	7th Jul 22
Accepted:	8th Jul 22

Editor: Daniel Klimmeck

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision **17th Feb 2022**

Dear Doina,

Thank you again for the submission of your manuscript (EMBOJ-2021-110488) to The EMBO Journal. Please accept my sincere apologies for the unusual delay with the peer-review of your manuscript. It was particularly difficult to get referees assigned at this time of the year. Your manuscript has been initially sent to three reviewers, however one expert got much delayed and in the end did not send us his-her report after repeated chasers. We have received reports from the other two referees, which I enclose below, and decided to proceed with our decision based on these reports.

As you will see, the referees acknowledge the potential interest and novelty of your results, although they also express a number of important concerns that will have to be conclusively addressed before they can be supportive of publication of your manuscript in The EMBO Journal. In more detail, the referees' consistently point to the need to better detail the dynamics of Sox6 acquisition in expanding cell populations, the role of Sox6 in acute UV response versus LRC function, and distinct features of damage response in the two IFE compartments. In addition, the reviewers raise a number of points related to additional controls required to corroborate the findings, overall data and method representation, that would need to be conclusively addressed to achieve the level of robustness and clarity needed for The EMBO Journal.

I judge the comments of the referees to be generally reasonable and given their overall interest, we are in principle happy to invite you to revise your manuscript experimentally to address the referees' comments.

I will let you know in case we hear back from the protracted referee #1.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: https://www.embopress.org/page/journal/14602075/authorguide#transparentprocess

As you know, we generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

When submitting your revised manuscript, please carefully review the instructions below.

Thank you for the opportunity to consider your work for publication.
I look forward to your revision.

with
Best regards,

Daniel

Daniel Klimmeck, PhD
Senior Editor
The EMBO Journal

Instruction for the preparation of your revised manuscript:

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).

3) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point response to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

4) a complete author checklist, which you can download from our author guidelines (https://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/Author Checklist%20-%20EMBO%20J-1561436015657.xlsx). Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

5) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript.

6) It is mandatory to include a 'Data Availability' section after the Materials and Methods. Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database, and the accession numbers and database listed under 'Data Availability'. Please remember to provide a reviewer password if the datasets are not yet public (see https://www.embopress.org/page/journal/14602075/authorguide#datadeposition).
In case you have no data that requires deposition in a public database, please state so in this section. Note that the Data Availability Section is restricted to new primary data that are part of this study.
*** Note - All links should resolve to a page where the data can be accessed. ***

7) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at .

8) We would also encourage you to include the source data for figure panels that show essential data. Numerical data can be provided as individual .xls or .csv files (including a tab describing the data). For 'blots' or microscopy, uncropped images should be submitted (using a zip archive or a single pdf per main figure if multiple images need to be supplied for one panel). Additional information on source data and instruction on how to label the files are available at .

9) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online (see examples in https://www.embopress.org/doi/10.15252/embj.201695874). A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc. in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc. See detailed instructions regarding expanded view here: .

- Additional Tables/Datasets should be labelled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

10) When assembling figures, please refer to our figure preparation guideline in order to ensure proper formatting and readability in print as well as on screen:
http://bit.ly/EMBOPressFigurePreparationGuideline

Please remember: Digital image enhancement is acceptable practice, as long as it accurately represents the original data and conforms to community standards. If a figure has been subjected to significant electronic manipulation, this must be noted in the figure legend or in the 'Materials and Methods' section. The editors reserve the right to request original versions of figures and the original images that were used to assemble the figure.

11) For data quantification: please specify the name of the statistical test used to generate error bars and P values, the number (n) of independent experiments (specify technical or biological replicates) underlying each data point and the test used to calculate p-values in each figure legend. The figure legends should contain a basic description of n, P and the test applied. Graphs must include a description of the bars and the error bars (s.d., s.e.m.).

Further information is available in our Guide For Authors: https://www.embopress.org/page/journal/14602075/authorguide

We realize that it is difficult to revise to a specific deadline. In the interest of protecting the conceptual advance provided by the work, we recommend a revision within 3 months (18th May 2022). Please discuss the revision progress ahead of this time with the editor if you require more time to complete the revisions. Use the link below to submit your revision:

Link Not Available

Referee #2:

Mounting evidence suggests that epidermis is maintained by molecularly and functionally distinct basal cell states. The current manuscript combines label retention, lineage tracing and single-cell transcriptomic approaches to provide novel insights into how cells in these cell states expand and differentiate. The authors identify interesting similarities between mouse tail and human skin and cell state/epidermal domain and/or SOX6 dependent UV exposure responses.
The manuscript provides a direct continuation of prior studies by the group along with a discussion of long-standing controversies and questions in the epidermal stem cell field. The manuscript and data are overall well presented, experiments are well controlled and results are correctly interpreted.

Major comments:

What are the longer time consequences of Sox6 loss in skin epithelium in homeostasis conditions?

Correlation between SOX6 expression and UV response is intriguing. The rapid expansion of Sox6+ cells suggests one of two possibilities: 1.) increase in SOX6 expression levels increase number of cells that are above detection limit. 2.) SOX6 negative cells (in the non-LRC domain) become SOX6 positive. The authors should clarify this point. In addition, it would be critical to directly compare proliferation, apoptosis, gH2AX expression between SOX6+ (LRC) and SOX6- (non-LRC) cells.

Does the expression of other markers of LRCs also increase upon UV exposure? The paper would improve if Sox6 and LRC functions to UV response could be uncoupled.

The role of SOX6 in Acute UV Response is unclear. Proliferation markers are down and apoptosis markers are up in Sox6 iko skin epithelial cells without UV challenge (0hrs) suggesting that these responses are independent of UV? The UV damage appears to be higher in control compared to Sox6 iko cells, yet more of the cells die. It is difficult to understand cause and consequence of the data presented in Fig. 6. Ideally, data before UV treatment should also be presented.

Sox6 marks label retaining (slow cycling cells), yet Sox6 ko results in reduced proliferation and increased apoptosis. How can this be explained? Are the effects autonomous or non-autonomous?

Minor comments:

Some areas in the beginning of the results section of the paper seem focussed on controversies in the field. Although this is interesting and informative on background, it is also somewhat distracting.

Labelling of Fig. 5A is unclear. Which images correspond to which genotype/labelled lineage?

Line 296: We turned to bulk RNA-seq analyses using 10x Illumina sequencing. What does this mean? Are all single cells pooled into one cohort?

In Fig3I it would be helpful to chose sort = TRUE to order FeaturePlots from cells with highest to lowest expression. This would be helpful for Slc1a3, Sox6, Igfbp3 which appear to be expressed in a very small number of cells. It also seems as if only few LRCs express Sox6 in basal epithelial cells and even LRCs. Yet, the staining in Fig. 6G,H suggests ~50% of cells j- presumably in skin are positive. It should be clarified what Percent Sox6+ cells refers to. Skin? Basal Cells?, LRCs?

Fig. 5E, F panels are difficult to understand. How is a z-score of signaling pathways across cell populations generated?

Although this seems outside the scope of this study, one can not help but wonder how Sox6-loss affects LRCs and UV response on s cellular and molecular level. Knocking Sox6 out and analyzing the phenotype superficially doesn't seem a good way of ending a story. It seems it opens more questions at the end of the paper than there were at the beginning or that it answers...

Referee #3:

Summary:
In this manuscript, Ghuwalewala et al. compare basal epithelial domains in the adult interfollicular epidermis (IFE) of mouse and human skin. Work in the Tumbar laboratory has previously identified label (LRC) and non-label retaining cell (non-LRC) populations in the mouse IFE, characterized by distinct marker gene expression and proliferative index. In this follow-up work, the authors provide a more detailled analysis of the associated scale (non-LRC) and interscale (LRC) compartments in mouse

skin, and ask if human skin features a similar organization of IFE domains. Although no differences regarding the proliferative status were noted, the authors report that rete ridges, known regular indentations of human epidermis, share several markers of scale /non-LRC regions in mouse skin. Similarly, the inter-ridges (tissue in between rete ridges) were found enriched in markers expressed by murine interscale/LRC domains. In a remarkable effort the authors utilized single-cell as well as bulk RNA seq methodology and included available published data sets from human skin to report two different basal cell differentiation paths of IFE populations partly shared between mouse and human skin. GSEA analysis revealed various pathways with differential expression between the LRC and non-LRC populations, among others UV response genes. One of the UV responsive genes, Sox6, also showed enrichment in the interscale IFE. The authors went on to demonstrate a physiological function of Sox6 in epidermal homeostasis and in UVB-mediated responses in vivo, which adds to the strength of the manuscript.
Overall this is a very interesting and highly informative manuscript that contains various relevant findings of potential relevance for the broad readership of EMBO J. The significance of the manuscript could be further improved by substantiating the authors' hypothesis on differential UV responses in the specific IFE domains (see below).

Major comments:
• Figures 5/6: The hypothesis that the differential UV response of epidermal basal domains (by means of transcriptome analysis of the different cell populations) is related to a higher UV exposure of inter-ridge and inter-scale cells is very tempting and opens many new interesting questions. To further test that hypothesis it would be very helpful to
a) demonstrate that there is indeed a higher DNA damage (i.e. gamma-H2Ax staining) in the interscale region (in essence as done for the Sox6 loss-of-function analysis but simply in wild-type mice, costaining with LRC vs non-LRC markers),
b) and -if so- also relate this to the skin architecture in more detail. Superficially seen, at least in tail skin scales appear dome-shaped to the exterior, hence supposedly extending the interscale area. I agree with the authors though that this might be very different when only focusing on the basal layer. To clarify this point it would be great to see this addressed in the tissue (e.g. by H&E stainings of intact cross-sections followed by simple measurement of the relative position of the basal layers in scale vs. interscale, as well as the height of overlaying differentiated epidermal layers that may provide shielding functions).
Above analyses in wild-type tissue would help to understand if the populations of interest differ specifically in their transcriptional responses to UV (despite similar extent of exposure or damage), or if the initial damage is indeed different (which would support the idea of more vs. less UV protected regions within the basal epidermis).
c) What do the asterisks in Figure 5A and F refer to? In A: do we see tail or back?
d) Line 99-101: The authors state that Aspm expression is enriched in non-LRCs (as shown previously by them) but the genetic approach to mark this population by means of Aspm-CreER did not reveal a selective labeling of non-LRCs; instead the marked cells were found in both scale and interscale IFE, and line (Fig. 1B, C). The RNAseq data in Figure 5 utilized -I think- that very same Cre-ER model. Is it not surprising that the Aspm-CreER-derived cell population has the strongest 'UV response up' signature (judged by the heat map result in 5E), whereas the samples derived from Dlx1-CreER (higher specificity for interscale) and the Slc1A3-CreER (higher specificity for scale) are more alike regarding this class of regulated genes?
e) What was the reason not to compare the Slc1A3-CreER to the Dlx1-CreER population in back skin, as done in tail skin? If I am not mistaken, from the data sets presented in 5F the Slc1-CreER population represents the most 'selective' population (Fig. 1C), whereas the Aspm-CreER- and even more so the K14-CreER-derived samples represent broader basal IFE identities. This in mind, is the heat-map for UV responsive genes in back skin samples not surprising? Would be nice to at least discuss these points in more detail.
• Figure S2E:
a) The reported clustering of the analysed markers in non-tail skin is interesting. The authors state in the legend one should note the overlap of H2B-GFP-LRC/non-LRC with the presented markers in paw, ear, back. In several examples this is somewhat difficult to see. Please present a quantification of such overlap at least for the most prominent marker (i.e.: Slc1A3+/H2B-GFP- vs. Slc1A3+/H2B-GFP+ vs. Slc1A3-/H2B-GFP+ vs. Slc1A3-/H2B-GFP- cells) to provide a quantitative idea how pronounced such clustering is.
b) How about K31? Does it also cluster in these skin regions outside the tail?
c) Please indicate the genetic model used here, and the time point after chase, if chased.

Minor comments:
• Based on which criteria did the authors identify the indicated scale and interscale regions as outlined in e.g. Figure S1? Please mention this in the figure legends.
• Lines 121, 122: While the differential levels of K14 and K15 expression in scale and interscale is noted, calling them 'markers' for each compartment does not seem appropriate. They are both expressed in the respective other region, too.
• Line 124: Based on which observation is this last sentence? Please include a figure reference or remove the statement.

Non-essential suggestions:
• Regarding the terminology interscale 'line': Schweizer and Marks previously described the 'interscale fold' in mouse tail skin (PMID: 409492; Cancer Research 37,4195-4201, November 1977). Could the authors comment on whether this is the same compartment the authors refer to here? If so, renaming the 'line' into 'fold' could perhaps help the community for future reference.
• Figure 1E, 2G: Bar diagrams are rather difficult to interpret due to small size and thick boundaries.

We thank the reviewers for their comments and positive feedback that substantially improved our paper. We have performed new experiments and edited our manuscript to include the new requested data. Although not required by our reviewers, in the meantime we obtained several more thousand cells from scRNA-seq of tail skin sorted basal layer, which we add here to the previous data. The results and interpretation have not changed, but the panels appear somewhat different from the previous version. We address the specific critiques here point by point in this response to reviewers.

Referee #2:

Mounting evidence suggests that epidermis is maintained by molecularly and functionally distinct basal cell states. The current manuscript combines label retention, lineage tracing and single-cell transcriptomic approaches to provide novel insights into how cells in these cell states expand and differentiate. The authors identify interesting similarities between mouse tail and human skin and cell state/epidermal domain and/or SOX6 dependent UV exposure responses.
The manuscript provides a direct continuation of prior studies by the group along with a discussion of long-standing controversies and questions in the epidermal stem cell field. The manuscript and data are overall well presented, experiments are well controlled and results are correctly interpreted.

Major comments:

(1) What are the longer time consequences of Sox6 loss in skin epithelium in homeostasis conditions?

Previously the latest time point was 7-days post tamoxifen (TM) induction. We now add 30-days post induction, with n=3 Sox6 iKO and 3 WT mice sacrificed at PD49 (TM induction at PD19). To analyze the skin phenotype, we stained with Ki67 (proliferation), Caspase3 (cell death), and gammaH2AX (DNA damage) and K10 (differentiation). Interestingly, while by 7 days Sox6 loss caused decreased proliferation and increased cell death, by 30 days cell death was resolved and proliferation was increased relative to control. This is likely a compensatory effect of the epidermis, which results in no grossly notable skin phenotype in skin. We are including these new longer-term data to supplementary Figure S10R-U.

(2) Correlation between SOX6 expression and UV response is intriguing. The rapid expansion of Sox6+ cells suggests one of two possibilities:

1.) increase in SOX6 expression levels increase number of cells that are above detection limit.

2.) SOX6 negative cells (in the non-LRC domain) become SOX6 positive.

The authors should clarify this point.

We apologize for this confusion. Our so-called Sox6+ cells are more precisely Sox6high cells, with clear levels above the staining background, whereas the Sox6- cells are in fact Sox6low. We now clarify this point in Figure 5H and now split the basal cells in Sox6high and Sox6low cells instead of Sox6+ and Sox6-. We also point the reviewer to our previous quantifications that showed (1) similar expression levels per Sox6+ cell (now named Sox6high cell) in normal vs UVB irradiated epidermis (Fig. S8E) and (2) an increase in % Sox6high cells in basal layer upon UVB irradiation (Fig. 5H). In conclusion, in response to UVB irradiation of the mouse back skin, Sox6low cells acquire Sox6high levels. We now show more illustrative images in new Fig. S8F.

(3) In addition, it would be critical to directly compare proliferation, apoptosis, gH2AX expression between SOX6+ (LRC) and SOX6- (non-LRC) cells.

As requested, we have performed these quantifications in wild type scales (non-LRC domains), which have low Sox6 levels, and inter-scales (LRCs domains), which have high Sox6 levels. These data are detailed in Figure S10N-Q and discussed in text. We find no detectable differences between scales and inter-scales by these assays. This is not surprising to us, as the differences in proliferation between scales and inter-scales are subtle (~2-3x differences detectable only after 2-3 weeks of H2B-GFP pulse-chase experiments or more). Furthermore, although the basal layer cells are less exposed in scales and rete-ridges than in interscales and inter-ridges, we might not expect the DNA damage or cell death to be different. In fact, our data suggest differential regulation of the UV-response transcriptome, which we believe provide adaptation or protection of cells in the different regions. This would provide added assistance to genomes in the more exposed epidermal regions, resulting in no differences in DNA damage and cell death.

Does the expression of other markers of LRCs also increase upon UV exposure?

We had already tested Vamp1, Sox6, Slc1a3, and Aspm by IF staining in the previous version. In response to this request, we attempted a more systematic screen using UVB induction in mouse back skin followed by sorting basal cells at 6-hr induction and bulk RNA-seq. Unfortunately, transcriptome changes of irradiated and un-irradiated samples were un-reliable with this approach and will require more optimization. However, to respond to this point we used a list of UV-regulated genes from mouse epidermis available from the literature (Li et al 2021, Shen et al 2019). We used these and the UV-UP and UV-DOWN genes defined by Hallmark Analysis and computed an UV-response gene scores. We found that the UV-response gene score was indeed enriched in the single cell clusters with LRCs vs non-LRCs enriched gene scores, and we provide specific examples of some of these genes. Please see Figure S7 and S9 for summary of the new data.

The paper would improve if Sox6 and LRC functions to UV response could be uncoupled.
The role of SOX6 in Acute UV Response is unclear. Proliferation markers are down and apoptosis markers are up in Sox6 iko skin epithelial cells without UV challenge (0hrs) suggesting that these responses are independent of UV? The UV damage appears to

be higher in control compared to Sox6 iko cells, yet more of the cells die. It is difficult to understand cause and consequence of the data presented in Fig. 6. Ideally, data before UV treatment should also be presented.

Our time point 0 in Figure 6H-N is indeed the no-UVB data of Sox6 loss in normal homeostasis. Also Sox6 loss data in the absence of UVB is detailed in Figure S10. We also have the BrdU pulse chase followed by stainings for Ki67 and caspase in Figure 6A-G. All these data indeed show increased cell death and decreased proliferation during homeostasis, as summarized in our cartoon Figure 6O, left. There was no effect of Sox6 loss on DNA damage during homeostasis. Upon UVB exposure, proliferation goes down at early time points but is quickly compensated for, DNA damage goes down, and cell death goes further up when compared to both homeostasis or UVB response alone. Our interpretation is that in acute UV response Sox6KO cells fail to repair the induced DNA damage and die instead, causing a compound effect on the cell death phenotype. We believe the Sox6 role in homeostasis is to counteract low and chronic UV exposure and cannot be easily uncoupled from its role in acute UV response. Higher expression of Sox6 in the more UV-exposed regions during homeostasis and its rapid upregulation upon direct acute UVB exposure support our interpretation.

Sox6 marks label retaining (slow cycling cells), yet Sox6 ko results in reduced proliferation and increased apoptosis. How can this be explained?

We were also puzzled about the short-term effect of Sox6 loss on proliferation. This suggests that Sox6 normally promotes proliferation in the epidermis, although it is a gene upregulated in the more quiescent regions. One possibility is that the LRC domain cells have an increased set of quiescence signals (separate from Sox6). Sox6 is then necessary to counteract these signals to ensure healthy rates of renewal in these regions. In long-term (30-days) Sox6 loss has the opposite effect and it upregulates proliferation, likely due to skin compensation mechanisms.

Are the effects autonomous or non-autonomous?

In this study we focus on an in-depth characterization of spatial epidermal domains, compare mouse and human skin, and begin to decipher the physiological meaning of this spatial organization. In this context, Sox6 is just one LRC prototype gene that helped connect mouse tail and human skin organization and aided us to uncover a link between epidermal domain organization and UV response. Given the focus of this paper, the extent of our work here, and the time constrain of these revisions, a detailed mechanistic study on Sox6 will be the subject of a future study. With that said, it is likely that the short-term effects we see within days after Sox6 loss are cell autonomous, whereas the long-term effects (by 30 days) are non-cell autonomous.

Minor comments:

Some areas in the beginning of the results section of the paper seem focused on

controversies in the field. Although this is interesting and informative on background, it is also somewhat distracting.

Because detection of LRCs and non-LRCs is so central to our paper it is important for the reader to understand what the assumptions are. Upon feedback from some colleagues, the term 'LRC' may make the reader immediately think 'stem cells', but here LRCs simply reflect proliferation history and regenerative rates in different IFE domains. Also, it can be very confusing to readers in the field as to why some labs have not detected LRCs, so our careful documentation here clarifies this confusion. With that said, we now re-worded the sentence pointing to the controversy, to reduce distraction. We focus the text on our results, confirming the LRC/non-LRC distribution and minimum chase time required to detect them, so the results can be easily repeated by other labs.

Labelling of Fig. 5A is unclear. Which images correspond to which genotype/labelled lineage?

5A shows examples of images from Dlx1-CreER mice, and Figure S7 shows images from all the lineages analyzed. This has now been clarified in figures and legends.

Line 296: We turned to bulk RNA-seq analyses using 10x Illumina sequencing. What does this mean? Are all single cells pooled into one cohort?

This was a mistake in the text. This is just bulk RNA-seq of >10000 sorted cells. This is now corrected in text.

In Fig3I it would be helpful to chose sort = TRUE to order FeaturePlots from cells with highest to lowest expression. This would be helpful for Slc1a3, Sox6, Igfbp3 which appear to be expressed in a very small number of cells. It also seems as if only few LRCs express Sox6 in basal epithelial cells and even LRCs. Yet, the staining in Fig. 6G,H suggests ~50% of cells j- presumably in skin are positive. It should be clarified what Percent Sox6+ cells refers to. Skin? Basal Cells?, LRCs?

We thank the reviewer for this suggestion. We have now used this command for some lowly expressed genes in figure 3 and new Fig S9, and it indeed helps to visualize the low expressing genes better. Our sc-RNA data suggests Sox6 is expressed in 47% of BL cells in back skin which matches our % Sox6-high cells from immunofluorescence staining (Fig. 5H), which we now label as '% Sox6-high cells in BL' to clarify.

Fig. 5E, F panels are difficult to understand. How is a z-score of signaling pathways across cell populations generated?

This was mislabeled and is now changed to NES in revised Fig 5 and S7 (normalized enrichment score) the actual parameter used to plot the heatmap clustering.

Although this seems outside the scope of this study, one cannot help but wonder how Sox6-loss affects LRCs and UV response on s cellular and molecular level. Knocking

Sox6 out and analyzing the phenotype superficially doesn't seem a good way of ending a story. It seems it opens more questions at the end of the paper than there were at the beginning or that it answers...

We agree that the Sox6 downstream mechanisms is an interesting future question outside of the scope of this study. Here we focus on a broad characterization of spatial epidermal domains, their physiological implications, and their similarities in mouse and human skin. Sox6 is an LRC-domain prototype gene that really helped us begin to decipher the physiological meaning of these spatial domains, and without it the link with UV would suffer. On the other hand, given the extent of this study, a future separate study on Sox6 downstream mechanisms seems warranted.

Referee #3:

Summary:
In this manuscript, Ghuwalewala et al. compare basal epithelial domains in the adult interfollicular epidermis (IFE) of mouse and human skin. Work in the Tumbar laboratory has previously identified label (LRC) and non-label retaining cell (non-LRC) populations in the mouse IFE, characterized by distinct marker gene expression and proliferative index. In this follow-up work, the authors provide a more detailled analysis of the associated scale (non-LRC) and interscale (LRC) compartments in mouse skin, and ask if human skin features a similar organization of IFE domains. Although no differences regarding the proliferative status were noted, the authors report that rete ridges, known regular indentations of human epidermis, share several markers of scale /non-LRC regions in mouse skin. Similarly, the inter-ridges (tissue in between rete ridges) were found enriched in markers expressed by murine interscale/LRC domains. In a remarkable effort the authors utilized single-cell as well as bulk RNA seq methodology and included available published data sets from human skin to report two different basal cell differentiation paths of IFE populations partly shared between mouse and human skin. GSEA analysis revealed various pathways with differential expression between the LRC and non-LRC populations, among others UV response genes. One of the UV responsive genes, Sox6, also showed enrichment in the interscale IFE. The authors went on to demonstrate a physiological function of Sox6 in epidermal homeostasis and in UVB-mediated responses in vivo, which adds to the strength of the manuscript. Overall this is a very interesting and highly informative manuscript that contains various relevant findings of potential relevance for the broad readership of EMBO J. The significance of the manuscript could be further improved by substantiating the authors' hypothesis on differential UV responses in the specific IFE domains (see below).

We thank this reviewer for the clear and fair summary of our study and for noting the broad relevance of our findings.

Major comments:
• Figures 5/6: The hypothesis that the differential UV response of epidermal basal domains (by means of transcriptome analysis of the different cell populations) is related to a higher UV exposure of inter-ridge and inter-scale cells is very tempting and opens many new interesting questions. To further test that hypothesis it would be very helpful

to
a) demonstrate that there is indeed a higher DNA damage (i.e. gamma-H2Ax staining) in the interscale region (in essence as done for the Sox6 loss-of-function analysis but simply in wild-type mice, costaining with LRC vs non-LRC markers).

In our understanding of the system, although the basal layer cells are less exposed in scales and rete-ridges than in inter-scales and inter-ridges, we might not expect the DNA damage to be different. Our data indicate differential regulation of the UV-response transcriptome, which might reflect differential adaptation to different levels of UV exposure. This would provide added assistance to genomes in the more exposed epidermal regions, resulting in no differences in DNA damage and cell death. Our quantification in Figure S10P and S10Q are in line with this prediction. We add some text in Discussion touching on this important point.

b) and -if so- also relate this to the skin architecture in more detail. Superficially seen, at least in tail skin scales appear dome-shaped to the exterior, hence supposedly extending the interscale area. I agree with the authors though that this might be very different when only focusing on the basal layer. To clarify this point it would be great to see this addressed in the tissue (e.g. by H&E stainings of intact cross-sections followed by simple measurement of the relative position of the basal layers in scale vs. interscale, as well as the height of overlaying differentiated epidermal layers that may provide shielding functions).

We have now quantified the thickness of epidermis in skin sections stained with anti-K10 antibodies to clearly delineate the inter-scale region and compare it with scale (n=4 mice; Fig. S8A-C). Although we did not bias the measurements to the middle of the scale, which would be the thickest, the results still show the scales are more than double in thickness when compared with inter-scale. We also provide an H&E image for a different visual of this aspect of tail skin morphology.

Above analyses in wild-type tissue would help to understand if the populations of interest differ specifically in their transcriptional responses to UV (despite similar extent of exposure or damage), or if the initial damage is indeed different (which would support the idea of more vs. less UV protected regions within the basal epidermis).

Despite differential exposure as indicated by thickness measurements, we do not see differences in DNA damage or cell death. Again, we believe this is because the basal cells in the different regions already have differentially regulated UV response genes even at steady state, as demonstrated by our RNA seq data in the absence of UV irradiation (Fig. 5E).

c) What do the asterisks in Figure 5A and F refer to? In A: do we see tail or back?

Asterisk refers to back skin data placed in supplement. We have now clarified this missed point in the figure legend.

d) Line 99-101: The authors state that Aspm expression is enriched in non-LRCs (as shown previously by them) but the genetic approach to mark this population by means of Aspm-CreER did not reveal a selective labeling of non-LRCs; instead the marked cells were found in both scale and interscale IFE, and line (Fig. 1B, C). The RNAseq data in Figure 5 utilized -I think- that very same Cre-ER model. Is it not surprising that the Aspm-CreER-derived cell population has the strongest 'UV response up' signature (judged by the heat map result in 5E), whereas the samples derived from Dlx1-CreER (higher specificity for interscale) and the Slc1A3-CreER (higher specificity for scale) are more alike regarding this class of regulated genes?

This is an interesting question. We are currently characterizing in more detail the behavior and characteristics of Aspm. Although Aspm is not enriched in the non-LRC domains, it is in fact strongly expressed in the proliferative clusters from mouse back and tail skin and from all human skin data, as indicated by our violin plots. This is not the case for the other populations. Perhaps this expression in cell cycling cells makes the Aspm-marked population most sensitive to damage and might explain why it has the strongest 'UV response up'. Furthermore, both Dlx1 and Slc1a3 are expressed to some extent in the opposite domain, which might obscure some of the differences between these populations. We are now developing Sox6 genetic marking tools to isolate the inter-scale more specifically and address these questions more systematically in future.

e) What was the reason not to compare the Slc1A3-CreER to the Dlx1-CreER population in back skin, as done in tail skin?

The Dlx1-CreER driver is highly inefficient in back skin and we could not obtain enough cells for these experiments. The K14-CreER population served as control for the Aspm-CreER and Dlx1-CreER marked cells in the back skin. We spell this out clearly now in text.

If I am not mistaken, from the data sets presented in 5F the Slc1-CreER population represents the most 'selective' population (Fig. 1C), whereas the Aspm-CreER- and even more so the K14-CreER-derived samples represent broader basal IFE identities. This in mind, is the heat-map for UV responsive genes in back skin samples not surprising? Would be nice to at least discuss these points in more detail.

We have evidence from clonal analysis (Ghuwalewala et al, in preparation) that these 3 IFE populations differ in their long-term behavior in a manner independent of their spatial localization. We believe it'd be more clear if we reserved these discussions for the new study.

• Figure S2E:
a) The reported clustering of the analysed markers in non-tail skin is interesting. The authors state in the legend one should note the overlap of H2B-GFP-LRC/non-LRC with the presented markers in paw, ear, back. In several examples this is somewhat difficult to see. Please present a quantification of such overlap at least for the most prominent

marker (i.e.: Slc1A3+/H2B-GFP- vs. Slc1A3+/H2B-GFP+ vs. Slc1A3-/H2B-GFP+ vs. Slc1A3-/H2B-GFP- cells) to provide a quantitative idea how pronounced such clustering is.

We thank the reviewer for this suggestion, which led to new and revealing data. We have now quantified Slc1a3, Vamp1 and Sox6 expression in ear, paw and back skin (Figure S2F and G). We demonstrate that in all tissues only a fraction of basal cells expresses these markers. Although visually LRC and non-LRC clusters tend to overlap with the different markers, upon quantification with an arbitrary cut-off for GFP^{low} vs GFP^{high} only few significant differences were obtained. Slc1a3+ cells tend to be found in non-LRCs in both ear and back skin, while Vamp1 was up in LRCs in back skin. We now mention that the overlap of our markers with LRCs/non-LRCs is mild in these non-tail tissues, although their basal layer is still heterogeneous with respect to these markers.

b) How about K31? Does it also cluster in these skin regions outside the tail?

As far as we can tell the K31 expression in the epidermis is specific to the tail skin. We stained several times the other tissues but were unable to detect expression in the IFE supra-basal layers of back, ear and paw.

c) Please indicate the genetic model used here, and the time point after chase, if chased.

We already have a schematic in Figure 2A indicating the system and the time points where we collect the tissue. We have now also marked every panel on the figure for clarity.

Minor comments:
• Based on which criteria did the authors identify the indicated scale and interscale regions as outlined in e.g. Figure S1? Please mention this in the figure legends.

The H2B-GFP or BrdU LRCs clustering in inter-scales has been clearly shown previously by whole mount and sections by co-staining with K10 (Sada et al, 2016, Mascre et al 2012, Gomez et al, 2012). It is also seen clearly in the sections that we stained below those panels. Based on the LRC clustering, the dim DAPI staining in the scales, and the position of the 3-hair follicle groups, we outline the predicted location of scales and inter-scales in these whole mount images to help orient the reader. This is now mentioned in the legends.

• Lines 121, 122: While the differential levels of K14 and K15 expression in scale and interscale is noted, calling them 'markers' for each compartment does not seem appropriate. They are both expressed in the respective other region, too.

We agree that markers may be not appropriate. In fact, not just K14 and K15 but all the ‘markers’ are expressed to some extent in all the regions. We now call them ‘preferred genes’ or ‘enriched genes’ instead of ‘markers’.

• Line 124: Based on which observation is this last sentence? Please include a figure reference or remove the statement. “These molecular differences between basal IFE spatial LRC/non-LRC domains also exist to some extent in other body regions of the mouse skin. “
The text has been revised and please check Fig S2 revised for quantification. Basically, we find that our markers are expressed heterogeneously in the basal layer in all tissue analyzed, but the enrichment in LRCs and non-LRCs is best seen in tail skin.

Non-essential suggestions:
• Regarding the terminology interscale 'line': Schweizer and Marks previously described the 'interscale fold' in mouse tail skin (PMID: 409492; Cancer Research 37,4195-4201, November 1977). Could the authors comment on whether this is the same compartment the authors refer to here? If so, renaming the 'line' into 'fold' could perhaps help the community for future reference.

We checked this paper, and we could only find the term ‘fold’ once in that text and its definition is not discussed in depth. We have already introduced the term ‘line’ in Sada et al 2016, and we noted that the term was picked up in a couple of follow-up review papers. So, we believe that it is perhaps better for the field that we stick with this ‘line’ denomination now.

• Figure 1E, 2G: Bar diagrams are rather difficult to interpret due to small size and thick boundaries.

We have now revised the figures to be more visible

1st Revision - Editorial Decision **1st Jul 2022**

Dear Doina,

Thank you for submitting your revised manuscript (EMBOJ-2021-110488R) to The EMBO Journal. Your amended study was sent back to the two referees for their re-evaluation, and we have received comments from both of them, which I enclose below. As you will see, the reviewers stated that their issues have been comprehensively resolved and they are now in favour of publication.

Thus, we are pleased to inform you that your manuscript has been accepted in principle for publication in The EMBO Journal.

We now need you to take care of a number of minor issues related to formatting and data presentation as detailed below, which should be addressed at re-submission.

Please contact me at any time if you have additional questions related to below points.

Thank you for giving us the chance to consider your manuscript for The EMBO Journal. I look forward to your final revision.

Again, please contact me at any time if you need any help or have further questions.

with
Best regards,

Daniel

Daniel Klimmeck PhD
Senior Editor
The EMBO Journal

Formatting changes required for the revised version of the manuscript:

>> Appendix file: please add a ToC on its first page.

>> Dataset availability section: please add GEO dataset identifiers to the section and release privacy from the datasets. Adjust the Author checklist accordingly.

Further information is available in our Guide For Authors: https://www.embopress.org/page/journal/14602075/authorguide

Please use the link below to submit your revision:

Link Not Available

Referee #2:
The revised manuscript reads well and I find the data and conclusions interesting and well documented. The authors have addressed all my previous comments.

Referee #3:
The authors made substantial efforts to improve the manuscript. I do not have further requests, just like to note that the naming/numbering of supplemental figures requires correction.

The authors performed the requested editorial changes.

2nd Revision - Editorial Decision **8th Jul 2022**

Dear Doina,

Thank you for submitting the revised version of your manuscript. I have now evaluated your amended manuscript and concluded that the remaining minor concerns have been sufficiently addressed.

Thus, I am pleased to inform you that your manuscript has been accepted for publication in the EMBO Journal.

Please note that it is EMBO Journal policy for the transcript of the editorial process (containing referee reports and your response letter) to be published as an online supplement to each paper.

Also, in case you might NOT want the transparent process file published at all, you will also need to inform us via email immediately. More information is available here:
https://www.embopress.org/page/journal/14602075/authorguide#transparentprocess

---

Please note that in order to be able to start the production process, our publisher will need and contact you shortly regarding the page charge authorisation and licence to publish forms.

Authors of accepted peer-reviewed original research articles may choose to pay a fee in order for their published article to be made freely accessible to all online immediately upon publication. The EMBO Open fee is fixed at $6,100 USD / £4,950 GBP / €5,500 EUR (+ VAT where applicable).

We offer two licenses for Open Access papers, CC-BY and CC-BY-NC-ND.
For more information on these licenses, please visit:
http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en_US

Should you be planning a Press Release on your article, please get in contact with embojournal@wiley.com as early as possible, in order to coordinate publication and release dates.

On a different note, I would like to alert you that EMBO Press is currently developing a new format for a video-synopsis of work published with us, which essentially is a short, author-generated film explaining the core findings in hand drawings, and, as we believe, can be very useful to increase visibility of the work. This has proven to offer a nice opportunity for exposure i.p. for the first author(s) of the study. Please see the following link for representative examples and their integration into the article web page:
https://www.embopress.org/video_synopses
https://www.embopress.org/doi/full/10.15252/embj.2019103932

Please let me know, should you be interested to engage in commissioning a similar video synopsis for your work. According operation instructions are available and intuitive.

Finally, we have noted that the submitted version of your article is also posted on the preprint platform bioRxiv. We would appreciate if you could alert bioRxiv on the acceptance of this manuscript at The EMBO Journal in order to allow for an update of the entry status. Thank you in advance!

If you have any questions, please do not hesitate to call or email the Editorial Office.

Thank you for this contribution to The EMBO Journal and congratulations on a successful publication!

Please consider us again in the future for your most exciting work.

Best regards,

Daniel

Daniel Klimmeck, PhD
Senior Editor
The EMBO Journal
EMBO
Postfach 1022-40
Meyerhofstrasse 1
D-69117 Heidelberg
contact@embojournal.org
Submit at: http://emboj.msubmit.net

EMBO Press Author Checklist

Corresponding Author Name: Prof. Tudorita Tumbar	**USEFUL LINKS FOR COMPLETING THIS FORM**
Journal Submitted to: EMBO journal	The EMBO Journal - Author Guidelines
Manuscript Number: EMBOJ-2021-110488R	EMBO Reports - Author Guidelines
	Molecular Systems Biology - Author Guidelines
	EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January 2022)

This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your manuscript.

Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Not Applicable	N/A

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Yes	Materials and Methods

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Yes	Materials and Methods

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Not Applicable	N/A
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	N/A
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	N/A

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Yes	Materials and Methods
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	N/A
Please detail **housing and husbandry conditions**.	Yes	Materials and Methods

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	N/A
Microbes: provide species and strain, unique accession number if available, and source.	Not Applicable	N/A

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Yes	Our samples were de-indentified with respect to ethnicity and gender, but we noted available data on age and sex in Materials and Methods and Figures.

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Yes	Acknowlegments

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Yes	https://doi.org/10.1101/2022.02.23.481662 (BioRxiv)
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	N/A

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	N/A

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Sample size is determined by experimental requirements and not by statistical methods.
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Not Applicable	N/A
Include a statement about **blinding** even if no blinding was done.	Yes	No blinding was done, but experiments are often repeated by independent investigators including undergraduate students naïve to the expected
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Yes	Control and test mice were preferably littermates, age and sex matched. However, our sample sizes were not sufficeint to distinguish between sex differences as a biological variable.
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Materials and Methods/Figures

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Yes	Figure legends
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Figure legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Yes	Cornell Office of Research Integrity and Assurance IRB for Human Paricipants 0908000777
Studies involving **human participants**: Include a statement confirming that **informed consent** was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Yes	
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	N/A
Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Yes	Cornell Univeristy Institute of Animal Care and Use Committee Protocol #2007-0125
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	N/A

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	N/A
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	N/A
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	N/A

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	N/A
For **tumor marker prognostic studies,** we recommend that you follow the **REMARK** reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	N/A
For **phase II and III randomized controlled trials**, please refer to the **CONSORT** flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	N/A

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Datasets are currently being deposited in a public repository and an accesion number will be provided prior to publication.
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	N/A
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Yes	All codes used were publicly available and properly cited.
If publicly available data were reused, provide the respective **data citations in the reference list**.	Yes	Figures/Figure legends and in-text citations