nature portfolio

Peer Review File

Living sensor display implanted on skin for long-term biomarker monitoring

Corresponding Author: Professor Hiroyuki Fujita

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The researchers have developed a system of genetically modified cells which can act as living sensors platform and detect the presence of a limited number of biomarkers. the noteworthy part is that manages to grow a tissue part from these cell and implanted it into the mice skin to monitor the levels of target biomarkers. However there are some serious concerns here

-authors does not provide the full performance of the living tissue sensor, cross selectivity and sensitivity must be optimized

- the current strategy to implant the living tissue sensors is not applicable for long term applications as the long term applications may pose some risk of reversing the genetic modifications.

-skin implant is not the best was to keep the engineered cell alive, active and highly sensitive over long time periods and physiological conditions.

Reviewer #2

(Remarks to the Author)
Sawayama et al. have developed a tissue-engineered human skin equivalent that, upon transplantation to immunocompromised mice, retains the ability to detect proinflammatory stimuli when injected subcutaneously in the surrounding areas. To this end, they engineered keratinocyte cultures (containing stem cell subsets) with NFkB responsive element-controlled EGFP expression, generated tissue-engineered dermoepidermal constructs that were transplanted onto SCID hairless outbred mice, and noninvasively detected in vivo fluorescence emissions by the skin patches for >200 days.

I found the presented data overall convincing and of high interest for scientists in diverse fields, since they explore quite a novel idea, but the manuscript needs major significant amendments to provide a better picture of how feasible this approach would be in real life.

Major comments:

1. While the authors claim that “(…) Living Sensor Display is an effective technique for measuring the biomarkers within the body”, in truth, they only show convincing detection of repeated proinflammatory stimuli when injected in the vicinity of the skin transplant. This is key for the proposed approach to be of any practical use. Therefore, the detection of i.v. or i.p. delivered stimuli must be further explored to obtain a more convincing response in the absence of s.c. injection in the transplant area. Otherwise transplantation of skin patch into models of disease such as ob/ob mice might be of use to test applicability of the technology.

2. Further on this line, the patches are quite big in terms of the relative proportion of total skin that has to be transduced (one has to imagine) to obtain a significant response. It would be necessary to establish the upper and lower limits of detection of a given metabolite with differently sized skin patches. Only with those data in hand, it would be possible to infer if the proposed technology would be transferable to humans or pets as they imply. Anything larger that a few square millimeters or centimeters would be difficult/impossible to put into practice in human or pet skin.

3. NHEK cells were used in immunocompromised animals. However, to be of any future use in humans or other animals, the technology should be further validated. Specifically, the authors need to provide evidence that:
a) primary human epidermal keratinocyte cultures can be similarly transduced and the resulting skin patches crucially retain responsivity to proinflammatory stimuli long-term, indicating that the stem cell compartment can be targeted also in primary human keratinocyte cultures; and
b)transduced mouse keratinocyte cultures can be grafted in immune competent mice and the system responsivity can still be demonstrated in an in vivo setting where the immune system, with all its checks and balances, is fully competent. Both experiments are key because non-autologous skin patches would quickly be rejected in immune-competent individuals of any species.

4. On a similar vein, EGFP+ cells are immunogenic (See: Ansari et al. https://doi.org/10.1007/s12015-016-9670-8). It is to be expected that such skin patches would be immunogenic upon activation of EGFP expression in immune-competent animals. Therefore, it is key to demonstrate long term feasibility of this technology in animals that are naïve for GFP expression but with a fully competent immune system.

5. All statements on immune stain data are qualitative because no quantitation of positive cells is provided. This must be corrected.

6. Page 8: “These findings confirm that (…) the inflammatory response caused by transplantation subsides within approximately four weeks”. This can only be stated if cellular infiltration was assessed at earlier time points, which is not the case.

7. The discussion needs some relevant amendments. Specifically:
a) “we unexpectedly observed maturation of tissue-engineered skin after transplantation in SHO mice”. I would argue that, on the contrary, this is to be expected as can be ascertained by checking the abundant literature on human dermoepidermal skin equivalents grafted onto SCID mice.
b) Since the authors do not show any data that the technology is transferable outside mammals, any reference to plant leaves and other potential applications of this technology in the plant kingdom should be avoided.
c) “Strategies to limit the spread of GMOs may be necessary”. This is the least concern in skin patch transplants, where migration of EGFP positive cells to germline is certainly not expected.

Other comments:
• Figure 1 would be more appropriate as graphical abstract for the manuscript instead of a main figure.
• Introduction. Skin is not “the organ with the largest area of the body”. See: Sontheimer, https://doi.org/10.1038/jid.2013.335
• Figure 2 panel b lacks scales and unit definitions. The statistical significance of the potential differences found among panels d, e, f are lacking.
• I am concerned that the time allowed for in vitro stratification of the skin equivalents (four days) is too short. Accordingly, the pictures on Figure 3b-e show unconvincing epidermal stratification. 10d would be more appropriate. And more convincing staining of markers for basal and suprabasal keratinocytes should be shown.
• The authors should provide an explanation (at least speculative) of why fluorescence intensity decreases over time in Figure 3g. Once more, is the difference between panels f and g statistically significant?
• Reporting of construct-related data needs to be more detailed and transparent to enable reproducing the results
• Appropriate statistical tests and accurate description of error bars and probability values are lacking

Reviewer #3

(Remarks to the Author)
In this manuscript, the authors developed the Living Sensor Display technique for long-term, in situ monitoring of biomarkers such as TNF-α and lipopolysaccharides. They engineered keratinocyte stem cells into living sensors, which were then transplanted into a tissue-engineered skin and engrafted onto mice. Thus, the tissue-engineered skin could continuously monitor the inflammatory situations in vivo. The concept is interesting, but major revisions, especially on the figures, must be made before considering publication in Nature Communications. The authors should carefully improve their figures, which are not clear, poorly aligned, and full of unprofessional mistakes.

The following detailed comments are given to strengthen their manuscript.
1) Page 1, line 22. " …for the real-time, non-invasive observation of biomarker…". How do the authors define "non-invasive"? It seems that engrafting tissue-engineered skin will cause damage to the original skin surface.
2) Page 4, line 14. The concentration-dependent responsiveness is not well supported as only three concentrations were evaluated (0, 0.2, 20 ng/ml). The authors need to provide a dose-response curve of their TNF-α sensors; that is, how the fluorescence levels change with various TNF-α concentrations. The dose-response curve will provide more information on the sensors' dynamic range, sensitivity, and detection limit.
3) Page 26, Figure 2. A fluorescence histogram with error bars should be provided to show the fluorescence of three biological repeats. The current panel is hard to understand and only shows one biological repeat.
4) Page 27, Figure 3. The fluorescence in Figures 3f and 3g peaks in 16h and then decreases. What causes this phenomenon, especially in the tissues without the TNF-α treatment (Figure 3f)? The tissues also exhibit a significantly higher background fluorescence (Figure 3f) and lower dynamic range (Figure 3g) than the NHEK cell-based sensors in Figure 2f. Discussion on this phenomenon is suggested.
5) Page 28, Figure 4. It's weird that Figure 4c is not included in the result section, but in the discussion section. Please either

describe it in the results or move it to supplementary figures. Similarly, the Figures 6a-f are not properly referred to in the texts.
6) Some panels could use legends or captions to make them clearer. For instance, what do the lines and the symbols represent in Figures 5f, 6f, S3, and S8?
7) Page 10, line 13. " some mice …, while others…". Precise mice numbers are needed here.
8) Brief summaries of the results in the fluorescence images (e.g., Figures 4b, S2, S4, S5, and S7) could be added to the figure captions to help readers understand these figures.
9) Page 32, Figure S1. The skin in Figure S1b should not have been treated with TNF-α; why does it show such obvious fluorescence?

There are also many minor mistakes in the texts and figures. The authors should carefully examine the manuscript and correct them. I listed some, but not all, below:
1) Page 6, line 21. "vimentin" should be "Vimentin".
2) Page 26, Figure 2b. Please separate "Uninfectedunstimulated".
3) Page 26, Figure 2b. The unit, value, and title of the y-axis, and the unit and values of the x-axis are missing.
4) Page 26, Figure 2b, right. The EGFP elements are not aligned correctly in the No.4 construct.
5) Page 28, Figure 4. Three pictures in panel a are not properly aligned.
6) The supplementary figures should be rearranged according to their order in the text. For example, Figure S9 should be changed to S4, as it appeared before the current S4.
7) The scale bars in Figures 6i, 6j, and S4 need to be corrected.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I do not have additional comments

Reviewer #2

(Remarks to the Author)
Samayama et al. have now satisfactorily responded to most of my original concerns, but some aspects need further clarification. To facilitate the conversation, I will order the still unresolved needs according to my original list:

1- I am satisfied by the new results shown on Fig S12, where they now demonstrate significant detection of i.v. stimuli 2 days post-injection. This is a very relevant result.However, the fact that there is a 2-day delay from i.v. stimulus to patch detection should be acknowledged in the Discussion section, because it has implications on the use of this technology that cannot be overseen.

2- I am also satisfied with the response that both patches of 9 and 200 sq mm worked similarly well. While I am not asking for further experimentation here, I believe that the discussion paragraph where they speak of detection limits (“Through this work, we demonstrated the effectiveness of detecting substances…”) should also include some words on the patch size and how it influences detection limits (I am assuming that, in equal conditions, larger patches will have a lower detection limit).

3a- Thanks for the clarification, my mistake.

3b- The PBMC co-culture surrogate assay is acceptable.

4- Limitation of non-testing in immune-competent animals is now correctly acknowledged.

5- Quantitative statements cannot be "approximately 90%" or "approximately 75%". I am afraid we need to see a proper quantitative study where they show graphs with means and standard deviations in an appropriate number of replicates as per journal guidelines.

6- Same here as comment 5.

7a- The authors are free to maintain this interpretation if they wish, but I insist that maturation of monolayer human epidermal equivalents upon transplantation onto mice is a well-known fact and not to be unexpected. On the contrary, the epidermis would be unexpected to remain immature if skin correctly engrafted.

7b & 7c- OK

Other (minor) comments. OK.

Reviewer #3

(Remarks to the Author)
Comments to authors
The authors have made some revisions, mainly to the texts. However, I am not convinced by their explanations in the response letter, and I insist that their figures are far from the publishable level, with detailed comments given below.
1） (Fig. 1) The authors only added one concentration (2 ng/ml) to their previous results, and in total, they only tested four concentrations (0, 0.2, 2, and 20 ng/ml). For biosensor characterization, these concentrations are too few to reflect the crucial properties of biosensors, e.g., the limit of detection and sensitivity. The authors should add fluorescence data at more concentrations below 0.2 ng/ml.
2） (Fig. 2) The authors claim that their error bars are too small to see clearly. However, it is wired that obvious error bars are in Fig. S1a-b, while no deviation can be seen in Figs. 2d-f, 3f-g, S1d, S2, and S3.
3） (Fig. 2b) The authors should change the flow cytometry histograms in Figure 2b into bar graphs. The current version is messy and not clear enough to display the relative fluorescence intensities and dynamic ranges of each construct, both unstimulated and stimulated.
4） (Fig. 5) I noticed that the error bars in Fig. 5g-i were not centered. The lines showing statistical significance in Fig. 5g, 5i, and 6g seem drawn manually and not aligned well. Were these error bars and statistical significance directly generated by the plotting software or manually marked by the author?
5） The quality of the figures is still poor and far from the publishable standards. For instance:
a) (All figures) The authors use "ng/ml" in some figures (Fig. S1c) but "ng/mL" in others (Fig. S1a). Please make them consistent.
b) (Fig. S1a-c) Panels are not properly aligned.
c) (Fig. 3b) An extra line (blue) and arrow (black) are inserted without any annotations. Remove them.
d) (Fig. S1c) The Y-axis title is missing.
e) (Fig. 5k) The text and scale bars overlap.
f) (Fig. S4b) There is no scale on the X-axis. Remove ")" in the legend.
g) (Fig. S14) It's weird that Fig.S14 is not included in the result section.
h) (Fig. S5) Delete ";" in the legend.
i) (Fig. S8) Delete "." in the legend.
j) (Fig. S6, 7 & 10). The scale bars in HE panels are not consistent. Some are labeled with 100 μm, but the others were not.
k) (Figs. 1-6) Remove the "." following "Figure".

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The researchers have developed a system of genetically modified cells which can act as living sensors platform and detect the presence of a limited number of biomarkers. the noteworthy part is that manages to grow a tissue part from these cells and implanted it into the mice skin to monitor the levels of target biomarkers. However there are some serious concerns here

-authors does not provide the full performance of the living tissue sensor, cross selectivity and sensitivity must be optimized

We appreciate the reviewer's valuable feedback on the performance of the Living Sensor Display. In this study, we have demonstrated that the system exhibits:

1. High sensitivity to TNF-α at the ng/mL level, which is difficult to measure in vivo.
2. Reproducible responses in a 3D tissue-engineered skin model.
3. Long-term stability after transplantation, maintaining responsiveness for over 200 days.
4. Real-time visualization of inflammatory states.

Regarding cross-selectivity, as the reviewer pointed out, we have addressed this aspect in our study. Specifically, we examined the fluorescence response of NF-κB-EGFP modified NHEKs to various stimuli. Our findings (Fig. S2) demonstrate that:

・TNF-α and IL-1β, both involved in NF-κB signaling, induced a fluorescence increase.

・IL-2 and LPS, which lack direct receptors in the NF-κB pathway, did not induce a fluorescence response.

These results confirm that the Living Sensor Display exhibits high specificity for NF-κB pathway-related signals, reinforcing its reliability for biomarker monitoring.

For further clarity, we have highlighted this discussion in the revised manuscript (Fig.S2) : "To investigate the cross-selectivity of NHEKs genetically modified with NF-κB-EGFP, we examined the response not only to TNF-α but also to IL-1β, which is involved in the NF-κB signaling pathway, and IL-2 and LPS, which are not directly involved in the NF-κB signaling pathway and lack membrane receptors. These factors were added to the culture medium at a concentration of 20 ng/mL, and changes in fluorescence intensity were observed (Fig. S2a, b, c and d). As a result, an increase in fluorescence intensity similar to that seen with TNF-α was confirmed for IL-1β. However, no increase in fluorescence intensity was observed for IL-2 or LPS. These results demonstrate that NHEKs genetically modified with NF-κB-EGFP exhibit high selectivity, responding specifically to target molecules associated with receptors involved in the NF-κB signaling pathway, while showing no cross-reactivity to IL-2 or LPS."

- the current strategy to implant the living tissue sensors is not applicable for long term applications as the long term applications may pose some risk of reversing the genetic modifications.

We appreciate the reviewer's concern regarding the stability of the genetic modifications in long-term applications. However, our results indicate that the Living Sensor Display remains stable and functional for over 200 days post-transplantation in SHO mice. Given that the median survival time of immunodeficient mice is approximately 37 weeks (6–12 months), the ability of the Living Sensor Display to maintain fluorescence response, turnover, and structural integrity without any maintenance suggests its feasibility for long-term applications.

Furthermore, we observed consistent fluorescence intensity and stable tissue size throughout the experiment, indicating that the engineered skin undergoes natural turnover while preserving its function. This is explicitly stated in the manuscript:
"Moreover, the area of the skin showing a fluorescent response hardly changed and was maintained for over 200 days after engraftment."

These findings suggest that the risk of genetic modification reversal is minimal within the observed timeframe. While further studies in different models would be valuable for future clinical applications, our results support the Living Sensor Display as a long-term, stable biomarker monitoring system.

-skin implant is not the best was to keep the engineered cell alive, active and highly sensitive over long time periods and physiological conditions.

We appreciate the reviewer's comment on the long-term viability of engineered cells in a skin implant. However, we would like to clarify that this study focuses on demonstrating the proof-of-concept (POC) of the Living Sensor Display rather than optimizing the best method for maintaining engineered cells under all physiological conditions.

As detailed in the manuscript, we systematically evaluated multiple approaches to tissue-engineered skin transplantation, ultimately selecting a method that enabled stable engraftment and long-term responsiveness. Our results show that the Living Sensor Display remained functional for over 200 days post-transplantation, maintaining stable fluorescence intensity and cellular turnover, which supports its feasibility for long-term applications.

These findings are explicitly discussed in the Discussion section, particularly regarding the stability and performance of the Living Sensor Display. Given this, we believe the current methodology is appropriate for the scope of this study and no modifications to the manuscript are necessary.

Reviewer #2 (Remarks to the Author):
Sawayama et al. have developed a tissue-engineered human skin equivalent that, upon transplantation to immunocompromised mice, retains the ability to detect proinflammatory stimuli when injected

subcutaneously in the surrounding areas. To this end, they engineered keratinocyte cultures (containing stem cell subsets) with NFkB responsive element-controlled EGFP expression, generated tissue-engineered dermoepidermal constructs that were transplanted onto SCID hairless outbred mice, and noninvasively detected in vivo fluorescence emissions by the skin patches for >200 days.

I found the presented data overall convincing and of high interest for scientists in diverse fields, since they explore quite a novel idea, but the manuscript needs major significant amendments to provide a better picture of how feasible this approach would be in real life.

Major comments:

1. While the authors claim that "(…) Living Sensor Display is an effective technique for measuring the biomarkers within the body", in truth, they only show convincing detection of repeated proinflammatory stimuli when injected in the vicinity of the skin transplant. This is key for the proposed approach to be of any practical use. Therefore, the detection of i.v. or i.p. delivered stimuli must be further explored to obtain a more convincing response in the absence of s.c. injection in the transplant area. Otherwise transplantation of skin patch into models of disease such as ob/ob mice might be of use to test applicability of the technology.

We appreciate your valuable feedback. As you noted, our previously published data only included subcutaneous (s.c.) administration of TNF-α and LPS. In response to your suggestion, we conducted additional experiments using intravenous (i.v.) LPS administration. Although the number of animals was limited, the results showed that LPS administered via the tail vein in mice carrying the Living Sensor Display led to an increase in fluorescence intensity. This suggests that target molecules secreted during in vivo inflammation leak from the vasculature and can be detected by the Living Sensor Display—a significant step toward practical application.

We have added these results in Fig. S12. In the Results section, we have inserted the following text:
"Furthermore, LPS solution (0.1 mg/mL) was administered via the tail vein at a dose of 0.1 mg/kg on Day 0 and Day 7. As a result, fluorescence intensity increased (Fig. S12a–e) from the day following administration, with a significant increase on the second day post-administration compared to baseline (Fig. S12f). By the seventh day post-administration, fluorescence had subsided; however, upon re-administration, fluorescence increased again two days later. On Day 9, fluorescence intensity was significantly higher than before the Day 7 administration (Fig. S12g)."

Additionally, we have updated the Discussion section by adding "We demonstrated the ability to continuously monitor inflammatory cytokines in vivo by transplanting tissue-engineered skin into mice. This skin was constructed from NHEK cells that produce EGFP upon stimulation by inflammatory cytokines, such as TNF-α and IL-1β, secreted from human immune cells."

2. Further on this line, the patches are quite big in terms of the relative proportion of total skin that has to be transduced (one has to imagine) to obtain a significant response. It would be necessary to establish the upper and lower limits of detection of a given metabolite with differently sized skin patches. Only with those data in hand, it would be possible to infer if the proposed technology would be transferable to humans or pets as they imply. Anything larger that a few square millimeters or centimeters would be difficult/impossible to put into practice in human or pet skin.

We appreciate the reviewer's comment regarding the size of the transplanted Living Sensor Display and its implications for practical applications. In the manuscript, we primarily present images of relatively large engrafted patches (10 mm × 20 mm) in Fig. S5a,b, as these cases provided the clearest visualization of fluorescence responses. However, we would like to clarify that successful fluorescence detection and biomarker response were also observed in much smaller patches (~3 mm × 3 mm) as shown in Fig. S5c,d, demonstrating that the system functions effectively even at reduced sizes.

Given this, we believe that the Living Sensor Display could be adapted for practical use in animals, including pets, without significant size constraints. While a precise determination of the upper and lower detection limits relative to patch size would be an important aspect for future clinical translation, our current results indicate that effective responses are achievable in small patches, which supports the feasibility of this approach.

3. NHEK cells were used in immunocompromised animals. However, to be of any future use in humans or other animals, the technology should be further validated. Specifically, the authors need to provide evidence that:
a) primary human epidermal keratinocyte cultures can be similarly transduced and the resulting skin patches crucially retain responsivity to proinflammatory stimuli long-term, indicating that the stem cell compartment can be targeted also in primary human keratinocyte cultures; and

We appreciate the reviewer's concern regarding the validation of the Living Sensor Display in primary human epidermal keratinocytes. However, we would like to clarify that the NHEKs used in this study are, in fact, primary human neonatal epidermal keratinocytes, which we obtained from KURABO Industries LTD. These cells were carefully screened prior to purchase and selected based on their ability to form skin equivalents, undergo natural turnover, and retain functional properties necessary for tissue engineering applications.

Furthermore, it has been reported that NHEKs contain a certain proportion of stem cells[1], and our results indicate that the engineered skin retains turnover ability after transplantation, supporting the conclusion that stem cells within the NHEK population contribute to long-term function. Since our study already

demonstrates stable engraftment, biomarker responsiveness, and turnover in vivo, we believe that the evidence requested by the reviewer has already been provided in the manuscript.

1. Wong, Y. L., Okubo, T., Uno, E., Suda, K. & Ishii, T. Role of CD99 in regulating homeostasis and differentiation in normal human epidermal keratinocytes. *Biochem. Biophys. Res. Commun.* **606**, 108–113 (2022).

b)transduced mouse keratinocyte cultures can be grafted in immune competent mice and the system responsivity can still be demonstrated in an in vivo setting where the immune system, with all its checks and balances, is fully competent. Both experiments are key because non-autologous skin patches would quickly be rejected in immune-competent individuals of any species.

We appreciate the reviewer for the valuable comment regarding the immune system, which is critical for the practical implementation of the Living Sensor Display. The purpose of this study is to obtain proof-of-concept data demonstrating the utility of the Living Sensor Display with human applications in mind. Therefore, we developed human-derived tissue-engineered skin and conducted experiments in immune-competent mice. We acknowledge that this concern is understandable—for instance, when human-derived tissue-engineered skin was transplanted onto C57BL/6 mice, it was rejected and failed to engraft. On the other hand, transplantation into humans would obviate this issue; however, that would require a highly advanced clinical study, which is beyond the scope of our current work.

To address whether the system can respond in the presence of a human immune environment, we performed proof-of-concept experiments using a co-culture system with genetically modified NHEK cells and human-derived Peripheral Blood Mononuclear Cells (PBMCs) in a culture insert. In these experiments, we placed a culture insert containing a solution of 1×10^6 cells of PBMCs over the NHEK cell layer, and in another condition, we added LPS (to a final concentration of 10 μg/mL in the well) to activate 1×10^6 cells of PBMCs within the insert. We then monitored changes in fluorescence intensity. No change was observed with the addition of PBMCs alone, but in the LPS-stimulated PBMC condition, an increase in fluorescence intensity was detected after 60 hours.

These results indicate that human tissue can function in a PBMC environment. Although this co-culture system does not fully recapitulate the complete human immune system, it is considered to mimic the early stages of acute inflammatory responses in humans. This suggests that when the Living Sensor Display is applied in humans, fluorescence responses induced by inflammatory cytokines such as TNF-α and IL-1β can also be observed.

We have added these findings in Fig. S3. In the Results section, the following text has been included:

“To verify whether NHEKs genetically modified with NF-κB-EGFP respond in a human immune environment, we performed co-culture experiments using a culture insert system with NHEK cells and human-derived Peripheral Blood Mononuclear Cells (PBMCs). A culture insert containing 1×10^6 PBMCs was placed over the NHEK cell layer. In another condition, a culture insert containing 1×10^6 PBMCs activated by LPS (final concentration: 10 μg/mL in the well) was used. Changes in fluorescence intensity were then monitored (Fig. S3a, b). As a result, fluorescence intensity remained unchanged with PBMCs alone; however, in the LPS-stimulated condition, it increased after 60 hours. These results suggest that NHEKs exhibit fluorescence responses in the presence of PBMCs, particularly under inflammatory conditions induced by LPS stimulation. Although this system does not fully represent a complete human immune system, it is considered to mimic the early stages of acute inflammatory responses, implying that in practical applications, fluorescence responses to cytokines can be observed in humans.”

Furthermore, in the Discussion section, we have inserted the following text: “We demonstrated the ability to continuously monitor inflammatory cytokines in vivo by transplanting tissue-engineered skin into mice. This skin was constructed from NHEK cells that produce EGFP upon stimulation by inflammatory cytokines, such as TNF-α and IL-1β, secreted from human immune cells into mice.”

4. On a similar vein, EGFP+ cells are immunogenic (See: Ansari et al. https://doi.org/10.1007/s12015-016-9670-8). It is to be expected that such skin patches would be immunogenic upon activation of EGFP expression in immune-competent animals. Therefore, it is key to demonstrate long term feasibility of this technology in animals that are naïve for GFP expression but with a fully competent immune system.

We appreciate the reviewer’s insightful comment regarding the potential immunogenicity of EGFP-expressing cells in immune-competent animals. While EGFP is expressed intracellularly and is not typically secreted into the extracellular space, minimizing the likelihood of directly triggering a local immune response, we acknowledge that EGFP peptides could still be recognized by the immune system if released due to cell death or damage. This could lead to antigen presentation by immune cells, as noted in previous studies (Ansari et al., 2016).

However, as part of our preliminary investigations, we attempted transplantation into fully immune-competent C57BL/6 (B6) mice, but the tissue-engineered skin did not engraft, likely due to immune rejection. This suggests that further optimization, such as the use of autologous cells or immunomodulation, would be necessary for translation to fully immune-competent models. Although evaluating this technology in humans would be the ultimate goal, such experiments require advanced clinical trials and are beyond the scope of this study.

To address this point, we have added the following statement to the Discussion section：

" Unlike conventional wearable sensor technologies, this method enables the long-term, real-time acquisition of minute biochemical signals, with potential applications ranging from health management and early disease detection to veterinary medicine and livestock monitoring. The target biological reactions include inflammatory markers, hypoxic and oxidative stress, and hormones, among others. This system is expected to have broad applications beyond human healthcare, such as improving the health management of pets that cannot communicate their medical conditions. **However, several challenges remain before widespread implementation. While SHO mice lack mature B and T lymphocytes, the observed innate immune responses (Fig. 6j) suggest that the transplanted tissue remains biologically active. Given reports of EGFP immunogenicity25, future studies using fully immune-competent models will be essential. Potential strategies include using autologous cells or alternative reporter proteins with lower immunogenicity to ensure broader applicability. In addition to biological limitations, societal acceptance may also pose a challenge. Strategies to control the spread of genetically modified organisms may be necessary; however, applications targeting non-human subjects, such as veterinary medicine and livestock health, may be more readily accepted."**

5. All statements on immune stain data are qualitative because no quantitation of positive cells is provided. This must be corrected.

We appreciate the reviewer's suggestion and have now quantified the number of positively stained cells in the immunohistochemical images. Specifically, we counted the number of macrophages (F4/80+) and neutrophils (Ly6g+) in the transplanted tissue at different time points.

Our results indicate that macrophage count decreased by approximately 90% from week 3 to week 4 post-transplantation. Neutrophil count decreased by approximately 75% in the same period. These quantitative findings have been added to the caption of Fig.S7 to provide a more precise interpretation of the immune response over time.

6. Page 8: "These findings confirm that (…) the inflammatory response caused by transplantation subsides within approximately four weeks". This can only be stated if cellular infiltration was assessed at earlier time points, which is not the case.

We appreciate the reviewer's suggestion and have now included data from an earlier time point (week 3 post-transplantation) to further support our conclusion.

1. At week 3, inflammatory cells were still observed near the transplant site, confirming that an acute inflammatory response was occurring, even in SHO mice.
2. By week 4, the infiltration of inflammatory cells had significantly decreased, suggesting that the immune response had subsided.

To provide a clearer comparison, we have added these findings to Fig. S6. In the Results section, we have

inserted the following text:
" The infiltration of inflammatory cell groups was observed to have subsided at 3 weeks after transplantation (Fig. S6) compared to 4 weeks after transplantation (Fig. S7), indicating that at least 4 weeks should be waited before fluorescence observation can be performed."
We believe this additional data strengthens our conclusion that the inflammatory response caused by transplantation resolves within approximately four weeks.

7. The discussion needs some relevant amendments. Specifically: a) "we unexpectedly observed maturation of tissue-engineered skin after transplantation in SHO mice". I would argue that, on the contrary, this is to be expected as can be ascertained by checking the abundant literature on human dermoepidermal skin equivalents grafted onto SCID mice.

We appreciate the reviewer's comment and acknowledge that the maturation of tissue-engineered skin after transplantation into immunodeficient mice has been previously reported in the literature. However, the key novel finding of this study is that by transplanting an immature, less-differentiated skin equivalent, we observed an increase in engraftment success, which has not been previously emphasized in the context of skin grafting.

Importantly, our study does not aim to induce maturation but rather demonstrates that a less mature construct enhances engraftment efficiency, which may have implications for optimizing tissue-engineered skin transplantation. This aspect is explicitly discussed in the Discussion section, where we highlight that maturation is not a necessary requirement for successful transplantation in this model.

b) Since the authors do not show any data that the technology is transferable outside mammals, any reference to plant leaves and other potential applications of this technology in the plant kingdom should be avoided.

We appreciate the reviewer's suggestion and agree that the discussion on plant applications extends beyond the scope of this study. To maintain the focus on the core findings, we have removed references to plant-related applications from the manuscript.
We believe this revision enhances the clarity of the discussion and ensures that the manuscript remains aligned with the primary objectives of our study.

c) "Strategies to limit the spread of GMOs may be necessary". This is the least concern in skin patch transplants, where migration of EGFP positive cells to germline is certainly not expected.

We appreciate the reviewer's insightful comment. We agree that the risk of EGFP-positive cells migrating to the germline in skin patch transplants is extremely low, making GMO containment a less critical concern

in this context. Nonetheless, we remain committed to following current safety and regulatory guidelines, ensuring that all necessary precautions are taken as the technology advances.

Other comments:

• Figure 1 would be more appropriate as graphical abstract for the manuscript instead of a main figure.

We appreciate the reviewer's perspective. Figure 1 was intentionally designed to convey the concept of the Living Sensor Display clearly at a glance, helping readers grasp the key idea efficiently. We believe that maintaining Figure 1 as a main figure enhances the manuscript's readability and does not compromise its scientific rigor. Given its role in setting the foundation for the study, we would prefer to keep it in its current placement.

• Introduction. Skin is not "the organ with the largest area of the body". See: Sontheimer, https://doi.org/10.1038/jid.2013.335

We appreciate the reviewer's correction. As suggested, we have revised the text to replace "the organ with the largest area" with "an organ with a large surface area" to ensure accuracy. This change has been implemented in the Introduction section. Thank you for bringing this to our attention.

• Figure 2 panel b lacks scales and unit definitions. The statistical significance of the potential differences found among panels d, e, f are lacking.

We appreciate the reviewer's suggestion and have addressed these concerns as follows:

For Figure 2 panel b, we have added appropriate scales and unit definitions to improve clarity.
For panels d, e, and f, we have included a new graph displaying fluorescence intensity at 24 hours, along with statistical significance markers to clearly indicate the differences among conditions in Fig.S1a.
Additionally, we have revised the manuscript in P.6, line 4, by adding:
"…significantly from 0 ng/mL to 20 ng/mL and from 0.2 ng/mL to 20 ng/mL at 24h after TNF-α stimulation (Fig.S1a,b),"
These revisions have been incorporated into Figure 2 and the corresponding figure legend, ensuring a more precise and statistically supported presentation of our results.

• I am concerned that the time allowed for in vitro stratification of the skin equivalents (four days) is too short. Accordingly, the pictures on Figure 3b-e show unconvincing epidermal stratification. 10d would be more appropriate. And more convincing staining of markers for basal and suprabasal keratinocytes should be shown.

We appreciate the reviewer's concern regarding the in vitro stratification period. However, based on our preliminary experiments, we found that over-maturation of the tissue-engineered skin negatively affects engraftment success. Specifically, prolonging the in vitro culture period before transplantation led to a significant reduction in engraftment efficiency, which is why we selected a four-day stratification period.

This aspect is discussed in the Discussion section, where we emphasize that transplanting an immature skin equivalent enhances engraftment efficiency. Given that successful engraftment, turnover, and long-term function were observed in vivo, we believe our approach is appropriate for this study.

While we acknowledge that longer in vitro maturation may improve stratification in vitro, our focus is on achieving optimal transplantation outcomes rather than maximizing pre-transplant stratification.

• The authors should provide an explanation (at least speculative) of why fluorescence intensity decreases over time in Figure 3g. Once more, is the difference between panels f and g statistically significant?

We appreciate the reviewer's suggestion. The decrease in fluorescence intensity over time in Figure 3g is likely due to EGFP photobleaching, which occurs progressively during prolonged fluorescence measurements. This is a common phenomenon in live-cell imaging and fluorescence-based assays.

To address the reviewer's concern regarding statistical significance, we have now added a new graph in Figure S4a, showing the fluorescence intensity at 17 hours post-peak, along with statistical significance markers. These data confirm that the differences observed in Figure 3f and 3g are statistically significant.
In the Results section, we have inserted the following text: "The fluorescence intensity increased significantly from 0 ng/mL to 20 ng/mL at 17 h, peaking at approximately 17 h before gradually decreasing (Fig.S4a)."

We believe these additions provide a clearer understanding of the fluorescence intensity changes over time and further support the robustness of our findings.

• Reporting of construct-related data needs to be more detailed and transparent to enable reproducing the results

We appreciate the reviewer's suggestion and have added detailed information on the cell constructs in the Methods section (Page 17, lines 4–5) to enhance reproducibility. These additions clarify the specific characteristics of the engineered cells used in our study.

• Appropriate statistical tests and accurate description of error bars and probability values are lacking

We appreciate the reviewer's feedback and have now included statistical analyses for both TNF-α-stimulated cultured cells and tissue-engineered constructs. These results, along with appropriate statistical tests, error bar descriptions, and p-values, have been added to Fig.S1a and Fig.S4a. The new data confirm that the observed differences are statistically significant.
We believe these revisions provide greater clarity and ensure the rigor of our analyses.

Reviewer #3 (Remarks to the Author):
In this manuscript, the authors developed the Living Sensor Display technique for long-term, in situ monitoring of biomarkers such as TNF-α and lipopolysaccharides. They engineered keratinocyte stem cells into living sensors, which were then transplanted into a tissue-engineered skin and engrafted onto mice. Thus, the tissue-engineered skin could continuously monitor the inflammatory situations in vivo. The concept is interesting, but major revisions, especially on the figures, must be made before considering publication in Nature Communications. The authors should carefully improve their figures, which are not clear, poorly aligned, and full of unprofessional mistakes.
The following detailed comments are given to strengthen their manuscript.

1) Page 1, line 22. " …for the real-time, non-invasive observation of biomarker…". How do the authors define "non-invasive"? It seems that engrafting tissue-engineered skin will cause damage to the original skin surface.

We appreciate the reviewer's comment and acknowledge that the initial implantation step involves an invasive procedure. To clarify this, we have revised the sentence in Page 1, line 22, by adding:
"…with the only invasive step being the initial implantation, "
This ensures a more precise definition of "non-invasive" in the context of long-term biomarker monitoring.

2) Page 4, line 14. The concentration-dependent responsiveness is not well supported as only three concentrations were evaluated (0, 0.2, 20 ng/ml). The authors need to provide a dose-response curve of their TNF-α sensors; that is, how the fluorescence levels change with various TNF-α concentrations. The dose-response curve will provide more information on the sensors' dynamic range, sensitivity, and detection limit.

We appreciate the reviewer's suggestion and have now included a dose-response curve showing the concentration-dependent changes in fluorescence intensity for both cultured cells and tissue-engineered skin. These data have been added to Fig.S1b and Fig.S4b.
In the Results section, we have inserted the following text:
"The fluorescence intensity changes with various TNF-α concentrations were observed in both cultured cells and tissue-engineered skin, indicating that both systems exhibit a dynamic range from 0 to 20 ng/mL (Fig.S1b and S4b)"

Furthermore, we note that in vivo, TNF-α concentrations rarely exceed 20 ng/mL, meaning that the Living Sensor Display has a sufficient dynamic range for physiological conditions. This point is already discussed in the Discussion section.

3) Page 26, Figure 2. A fluorescence histogram with error bars should be provided to show the fluorescence of three biological repeats. The current panel is hard to understand and only shows one biological repeat.

We appreciate the reviewer's suggestion and have revised Figure 2d and 2e to improve clarity. Since the standard deviation is very small, the error bars overlapped in the original figure, making them difficult to visualize. To address this, we have added a new graph at the 24-hour time point in Figure S1a, where the fluorescence intensity for different concentrations is displayed with error bars.
Additionally, we have revised the manuscript in P.6, line 4, by adding:
"…significantly from 0.2ng/mL to 20 ng/mL at 24h after TNF-α stimulation (Fig.S1a)"
These modifications ensure that the statistical significance of fluorescence intensity differences is clearly conveyed.

4) Page 27, Figure 3. The fluorescence in Figures 3f and 3g peaks in 16h and then decreases. What causes this phenomenon, especially in the tissues without the TNF-α treatment (Figure 3f)? The tissues also exhibit a significantly higher background fluorescence (Figure 3f) and lower dynamic range (Figure 3g) than the NHEK cell-based sensors in Figure 2f. Discussion on this phenomenon is suggested.

We appreciate your valuable feedback. In particular, the slight increase in fluorescence intensity observed in tissues not treated with TNF-α is likely due to the inevitable peak seen in all cases when using tissue-engineered skin. This suggests that the genetically modified cells may respond to various stimuli, and medium changes might also introduce some stimulation.

Furthermore, regarding the high background and low dynamic range: although fluorescence measurements were performed using the same apparatus as that for cultured cells, the tissue-engineered skin tends to shrink due to its own contractile force when removed from the culture insert, making quantitative measurement difficult. Therefore, we conducted fluorescence observations directly on the tissue-engineered skin fabricated within the culture insert. Because the transparency of both the culture insert and the tissue-engineered skin is lower than that of cultured cells, scattering is increased, which results in a higher background and a smaller dynamic range.
In the Results section, we have inserted the following text:
"Compared to the data obtained from cultured cells, the dynamic range relative to the background in tissue-engineered skin is smaller. This is likely because the fluorescence measurements were conducted through a culture insert and scattering from the culture insert and the tissue-engineered skin may have influenced

the results."

5) Page 28, Figure 4. It's weird that Figure 4c is not included in the result section, but in the discussion section. Please either describe it in the results or move it to supplementary figures. Similarly, the Figures 6a-f are not properly referred to in the texts.

Thank you for your valuable feedback. We have added Figure 4c and Figures 6a–f into the result.

6) Some panels could use legends or captions to make them clearer. For instance, what do the lines and the symbols represent in Figures 5f, 6f, S3, and S8?

We appreciate your valuable feedback and have made the necessary revisions. The dotted lines indicate the timing of fluorescence observation, while the triangles denote the timing of administration. Additional details have been included in the figure caption.

7) Page 10, line 13. " some mice …, while others…". Precise mice numbers are needed here.

Thank you for your valuable feedback. We have included the number of mice and made the necessary revisions.

8) Brief summaries of the results in the fluorescence images (e.g., Figures 4b, S2, S4, S5, and S7) could be added to the figure captions to help readers understand these figures.

We appreciate your suggestion. We have added brief summaries of the fluorescence image results to the figure captions.

9) Page 32, Figure S1. The skin in Figure S1b should not have been treated with TNF- α ; why does it show such obvious fluorescence?

This figure shows a mouse at 4 weeks post-transplantation, when the fluorescence intensity is elevated due to inflammation from the transplant. Subsequently, after waiting 2–3 weeks for the fluorescence intensity to decrease, administration experiments were performed using mice. Further details are provided in the caption in Fig. S5. We have revised the sentence in Page 9, line 26, by adding:
"…suggesting that fluorescence observation should not be initiated until at least four weeks have passed. "

There are also many minor mistakes in the texts and figures. The authors should carefully examine the manuscript and correct them. I listed some, but not all, below:

1) Page 6, line 21. "vimentin" should be "Vimentin".

We appreciate your valuable feedback and have made the necessary revisions.

2) Page 26, Figure 2b. Please separate "Uninfectedunstimulated".

We appreciate your valuable feedback and have made the necessary revisions

3) Page 26, Figure 2b. The unit, value, and title of the y-axis, and the unit and values of the x-axis are missing.

We appreciate your valuable feedback and have made the necessary revisions

4) Page 26, Figure 2b, right. The EGFP elements are not aligned correctly in the No.4 construct.

We appreciate your valuable feedback and have made the necessary revisions

5) Page 28, Figure 4. Three pictures in panel a are not properly aligned.

We appreciate your valuable feedback and have made the necessary revisions

6) The supplementary figures should be rearranged according to their order in the text. For example, Figure S9 should be changed to S4, as it appeared before the current S4.

We appreciate your valuable feedback and have made the necessary revisions

7) The scale bars in Figures 6i, 6j, and S4 need to be corrected.

We appreciate your valuable feedback and have made the necessary revisions

REVIEWER COMMENTS

Reviewer #2

1- I am satisfied by the new results shown on Fig S12, where they now demonstrate significant detection of i.v. stimuli 2 days post-injection. This is a very relevant result. However, the fact that there is a 2-day delay from i.v. stimulus to patch detection should be acknowledged in the Discussion section, because it has implications on the use of this technology that cannot be overseen.

Thank you for your comment. The graph in question was originally plotted at the 2-day time point to match the timing used for i.p. and s.c. stimulations. As a result, it may appear that there is a 2-day delay for i.v. detection. However, as shown in the time-course profile in Fig. S12e, a clear increase in signal following i.v. injection is already evident at 1-day post-stimulation. Therefore, we believe that the patch can detect i.v. stimuli as early as 1 day after injection.
To avoid confusion, we have revised panels (f) and (g) to present the 1-day data instead of the 2-day time point.

2- I am also satisfied with the response that both patches of 9 and 200 sq mm worked similarly well. While I am not asking for further experimentation here, I believe that the discussion paragraph where they speak of detection limits ("Through this work, we demonstrated the effectiveness of detecting substances…") should also include some words on the patch size and how it influences detection limits (I am assuming that, in equal conditions, larger patches will have a lower detection limit).

As suggested, we have revised the Discussion section. Following the sentence "Through this work, we demonstrated the effectiveness of detecting substances…", we added the sentence as :

"We also found that even small-sized patches can exhibit responsiveness; however, as the transplantation area decreases, the total fluorescence signal becomes weaker, leading to a reduced detection limit."

This addition, which addresses the influence of patch size on detection sensitivity as pointed out by the reviewer, has been inserted on page 15, line 16 of the revised manuscript.

3a- Thanks for the clarification, my mistake.
3b- The PBMC co-culture surrogate assay is acceptable.
4- Limitation of non-testing in immune-competent animals is now correctly acknowledged.

Thank you for your understanding.

5- Quantitative statements cannot be "approximately 90%" or "approximately 75%". I am afraid we need to see a proper quantitative study where they show graphs with means and standard deviations in an appropriate number of replicates as per journal guidelines.
6- Same here as comment 5.

We appreciate the reviewer's concern regarding the use of approximate percentages. In this case, our intention was simply to convey whether red fluorescent signals were present or absent in the immunostained tissue images. In response to the previous revision request, we performed cell counting to provide numerical data; however, given that nonspecific staining can occasionally occur in in vivo immunohistochemistry, we acknowledge that the resulting numbers may have limited quantitative precision. Therefore, we would like to respectfully request that the use of approximate percentages (e.g., "approximately 90%") be allowed in this context, as we do not intend to make strong claims based on these figures.

7a- The authors are free to maintain this interpretation if they wish, but I insist that maturation of monolayer human epidermal equivalents upon transplantation onto mice is a well-known fact and not to be unexpected. On the contrary, the epidermis would be unexpected to remain immature if skin correctly engrafted.

We respectfully acknowledge the reviewer's perspective. However, to the best of our knowledge, there have been no previous reports in which tissue-engineered skin transplanted onto mice exhibited both well-developed rete ridge-like structures and sustained epidermal turnover over an extended period. In our experiments, tissue-engineered constructs that included pre-formed dermal and epidermal layers did not engraft successfully. Likewise, subcutaneous co-culture of keratinocytes and fibroblasts resulted in minimal engraftment. Therefore, from our experimental experience, the observed epidermal maturation and long-term turnover were truly unexpected, and we used the term "unexpected" in that context.

7b & 7c- OK

Other (minor) comments. OK.

Thank you for your understanding.

Reviewer #3

1) (Fig. 1) The authors only added one concentration (2 ng/ml) to their previous results, and in total, they only tested four concentrations (0, 0.2, 2, and 20 ng/ml). For biosensor characterization, these concentrations are too few to reflect the crucial properties of biosensors, e.g., the limit of detection and sensitivity. The authors should add fluorescence data at more concentrations below 0.2 ng/ml.

As suggested, we have added fluorescence data for two additional lower concentrations (0.002 ng/ml and 0.02 ng/ml) in Fig. S1d. Based on these results, we have noted in the manuscript that "but detection below 0.02 ng/ml falls outside the dynamic range."

2) (Fig. 2) The authors claim that their error bars are too small to see clearly. However, it is wired that obvious error bars are in Fig. S1a-b, while no deviation can be seen in Figs. 2d-f, 3f-g, S1d, S2, and S3.

All graphs have been recreated using GraphPad Prism to ensure consistent formatting, and the error bars have been adjusted to be clearly visible in all relevant figures.

3) (Fig. 2b) The authors should change the flow cytometry histograms in Figure 2b into bar graphs. The current version is messy and not clear enough to display the relative fluorescence intensities and dynamic ranges of each construct, both unstimulated and stimulated.

As suggested, we have replaced the original flow cytometry histograms in formerly Figure 2b with bar graphs showing the median fluorescence intensity (MFI) for each construct under both unstimulated and stimulated conditions. The original histograms have been moved to Supplementary Figure S1 for reference.
Additionally, we have updated the Lentivirus vector construction section in the Materials and Methods (page 17, line 6) to include the following sentence:

"Flow cytometry data were analyzed using FlowJo software (v10.10.0, FlowJo LLC). After doublet exclusion, SYTOX Blue-negative cells were defined as live cells. GFP fluorescence

was quantified as median fluorescence intensity (MFI). For visual comparison across conditions, fluorescence histograms were normalized to mode, with the highest frequency set to 100% on the y-axis."

4) (Fig. 5) I noticed that the error bars in Fig. 5g-i were not centered. The lines showing statistical significance in Fig. 5g, 5i, and 6g seem drawn manually and not aligned well. Were these error bars and statistical significance directly generated by the plotting software or manually marked by the author?

As the reviewer pointed out, the error bars in Figure 4g-i(formerly Fig. 5g–i) were slightly misaligned. The previous figures were created in Excel and exported as images, and we assure you that no parts were hand-drawn or manipulated. The statistical significance markers were manually added using shapes, as Excel lacks built-in tools for this purpose.
To avoid any misunderstanding or doubt, we have now recreated all relevant graphs using GraphPad Prism, which ensures accurate placement of error bars and significance markers.

5) The quality of the figures is still poor and far from the publishable standards. For instance: a) (All figures) The authors use "ng/ml" in some figures (Fig. S1c) but "ng/mL" in others (Fig. S1a). Please make them consistent. b) (Fig. S1a-c) Panels are not properly aligned. c) (Fig. 3b) An extra line (blue) and arrow (black) are inserted without any annotations. Remove them. d) (Fig. S1c) The Y-axis title is missing. e) (Fig. 5k) The text and scale bars overlap. f) (Fig. S4b) There is no scale on the X-axis. Remove ")" in the legend. g) (Fig. S14) It's weird that Fig.S14 is not included in the result section. h) (Fig. S5) Delete ";" in the legend. i) (Fig. S8) Delete "." in the legend. j) (Fig. S6, 7 & 10). The scale bars in HE panels are not consistent. Some are labeled with 100 μm, but the others were not. k) (Figs. 1-6) Remove the "." following "Figure".

We sincerely thank the reviewer for the careful and detailed feedback. All the points mentioned have been addressed.