nature portfolio

Peer Review File

An unbiased tissue transcriptome analysis identifies potential markers for skin phenotypes and therapeutic responses in atopic dermatitis.

Corresponding Author: Professor Masayuki Amagai

Parts of this Peer Review File have been redacted as indicated to maintain patient confidentiality.

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 authors present a highly innovative approach of skin monitoring, using 1mm skin biopsies of a very large, unprecedented AD patient population. The data are novel and of high interest to the dermatological and immunological community. The authors assess both disease heterogeneity as well as treatment responses, which is of very high importance to better understand AD altogether, and for the development of novel treatment approaches.

Major

The stratification of baseline expression patterns between good and poor dupilumab responders is very interesting, but seems rather semi-quantitative. Are there any statistical measures that can give us a better idea as to this stratification (some sort of score with some measure of standard deviation/variance?)

To substantiate these claims, would it be possible to split their large cohort into training and validation sets to better understand reproducibility of their data?

Can the authors give a more exact idea as to the inflammatory level of dysregulation in healthy, nonlesonal and lesional AD skin? E.g. via PCA plot or other measures? It would be good to understand the order of magnitude of transcriptomic dysregulation particularly in nonlesional AD skin, to also get a better sense and comparability to previous studies who have investigated the level of transcriptomic dysregulation on nonlesional AD skin.

Can the authors rule out that type 17 baseline activation is not just a measure of overall disease severity, but truly a predictor of dupilumab response?

Reviewer #2

(Remarks to the Author)
Fukushima-Nokura and colleagues report a large-scale integration of phenotypic and molecular data in a prospective cohort of Atopic dermatitis patients. In a hypothesis-free approach, 29 gene clusters or metagenes were identified based on lesional transcriptome analysis. These clusters are subsequently correlated with blood biomarkers and investigated for their clinical utility to monitor disease severity as well as therapeutic response to dupilumab.
Great strenghtes of the manuscript are the large patient cohort and with that the valuable data resource, the novel method of minimally invasive 1mm punch biopsy collection that enables to obtain several biopsies in the same patient, and the prospective skin transcriptome analysis of patients under dupilumab treatment. General findings are that AD phenotypes are pretty stable over time, that several metagenes could be identified that were associated with disease severity, and that type 3 immunity associated metagenes are asscoiated with poor response to dupilumab.
Major comments:

1) Is there a correlation with the previously identified modules and endotypes that were published recently in a landmark paper by the same group (Sekita A et al, Nat Comm 2023)?
2) Was the association of IL-22 and IL-18 blood cytokine levels to treatment outcomes corrected for severity? As these are known markers associated with AD severity.
3) The longitudinal analysis of AD patients under dupilumab therapy is unprecedented in its quality and delivers precious insights. Nevertheless, the definition of the clinical responder types is not broadly accepted and should be defined more clearly.
4) The authors have obtained more than one lesional biopsy from many AD patients – it would highly upvalue the study to show some analysis and comment on stability depending on the lesional site and over time within one individual AD patient. Is the distribution of metagenes stable in an individual patient or may it change?
5) As a proof of principle: does it give a higher predictive value to determine the KC/FB1 metagene as compared to CCL20 (or IL-19, IL-17A itself) alone to predict dupilumab response in AD patients?
Minor: line 476: IGE instead of IEG

Reviewer #3

(Remarks to the Author)
Comments on “dissecting the heterogeneity of skin phenotypes and therapeutic responses in atopic dermatitis through a non-biased tissue transcriptome analysis” by Dr. Ayana Fukushima-Nomura. In this study, the authors performed a mixture ofa cross-sectional and longitudinal cohort study collecting 951 1mm skin biopsies for RNA-sequencing along with plasma and serum data. Overall, the data presentation and writing make it very difficult to appreciate the data and the findings reported here. I also have a hard time appreciating what findings here are truly novel and how this is telling us something novel about AD. I have the following comments.

The analyses and data presented here are very hard to follow, and the data is presented in a manner that is hard to interpret and understand. It is hard to appreciate the biological implications of what a KC/FB/MD component/module is, and a concern, as shown in Figure 2b, is the extensive overlap between many of these components/modules, making interpretation difficult.

What is the basis of the definition of “excellent”, “good” and “poor” response groups? What defined as “rapid response to dupilumab” – this is very unclear. This needs to be clarified and if the authors can refer back to an accepted definition of these response it would strengthen this type of stratification, which otherwise seems arbitrary.

The basis of the type 17 immune response seems weak. The authors need to address and show what are the metagenes associated with “IL-17 inflammation-induced responses”. Also, with bulkRNA-seq the data should be sensitive enough to pick up expression of IL17A or IL17F. If those are absent from the bulkRNA-seq data what is the reliability of this type-17 response? Also, does this response differentiate from other IL-17 members such as IL-17C or other non-IL-17 family members that act with IL-17 such as TNF?

Why are the serum levels shown as log10 values(+0.1)? Is this based on the pg /ml measurements? Please clarify.

It is unclear what the data quality was – how many differentially expressed genes between the different groups? Was the data of sufficient quality to differentiate between psoriasis and atopic dermatitis? This would be helpful to have. Please provide information about the quality of the data from the RNA-biopsies including RIN scores.

The abstract could be made clearer. The size of the cohort is impressive, but the findings are not clearly highlighted, and it is not clear what was assessed and how. The abstract would benefit from some rewriting. In the body of the manuscript, there are similar awkward sentences and grammatical structures that could benefit from careful read-over by a native English speaker.

I do not agree with the authors that the 1mm punch biopsy reveals robust expression profiles in AD. It appears that there is a high level of variability in the data that the authors are getting, and this reflects the way the data is analyzed and presented.

The Discussion is very long and should be condensed.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have sufficiently addressed all the reviewers' comments and concerns.

Reviewer #2

(Remarks to the Author)

Reviewer #3

(Remarks to the Author)
The authors have done quite extensive revisions and some of my concerns have been addressed. However, several still remain - my main concern with the data, while impressive in terms of sample size, is the quality of the transcriptomic data, the lack of validation of the Th17 response, and the definition of treatment response to dupilumab, which I feel is overstated. My remaining concerns are listed below:

In the Ferrucci study that the authors refer to, the cutoff for mild disease was EASI<7, which I agree is close, but I am not sure why the authors do not follow that but instead select <6.3 as the cutoff in this study. Therefore, this specific cutoff seems arbitrary and not adequately justified. How much would the findings change if the cutoff was the same as in the Ferrucci study?

Also, a significant limitation of this definition and the data presented here is that it doesn't consider the disease severity before treatment. Therefore, the "excellent" vs. "poor responders" is more or less meaningless, and if this was considered as an EASI response (i.e., EASI50/75/90), I'm not sure it would be defined as an excellent response. It's very clear, based on the clinical data shown in Figure 3 and S13 that their "excellent responders" had milder disease at baseline than the other two groups and the authors acknowledge this in the results section – my concern, therefore, is that this signature doesn't necessarily reflect, a response to dupilumab treatment but rather reflects baseline disease activity differences between these groups – it also seems counterintuitive that patients with an overall "poor" response to dupilumab would still be on the drug for 20 months as shown in Figure S13 - this is a major concern and needs to be addressed as the therapeutic stratification is a major message of this manuscript

The authors have not adequately addressed my comment regarding the IL-17 response and what it means. Again, the authors are using specific modules and not showing expression of IL17A, IL17F, or IL17C as requested and not addressing what it means to have an enriched "Th17 module" if there is no detectable expression of IL17A/F. The authors should note that many of the genes in those IL-17 signatures are induced by multiple other proinflammatory mediators, that is why many of these genes are showing up across multiple modules (i.e., in Figure S9B). The overlap between IL-17A and TNF appears quite significant. Therefore, this needs further validation and the authors should show results for IL17A/F and 17C from their bulk RNA-seq data.

The authors state that they provide RIN numbers, but I cannot find these data. Selecting the RIN cutoff at RIN<2 is concerning. Most groups that do this type of work draw the line at RIN>6. How many of the samples were above that threshold (RIN=6) and how many below?

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
I do not feel that the authors have adequately addressed my concerns and comments and the same issues remains despite some changes and revisions.

Version 3:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
I have no further comments.

Responses to Reviewers

Reviewers' comments are itemized below in bold font.
The authors' responses to each specific comment are in blue font with text line numbers and modified and new figures in red.

Reviewer #1 (Remarks to the Author):
The authors present a highly innovative approach of skin monitoring, using 1mm skin biopsies of a very large, unprecedented AD patient population. The data are novel and of high interest to the dermatological and immunological community. The authors assess both disease heterogeneity as well as treatment responses, which is of very high importance to better understand AD altogether, and for the development of novel treatment approaches.

Response:
We sincerely thank the reviewer for their thorough evaluation of our manuscript and for acknowledging the significance of our study. We are grateful for the insightful suggestions that have enhanced the presentation of our work.

Comment 1: ***The stratification of baseline expression patterns between good and poor dupilumab responders is very interesting, but seems rather semi-quantitative. Are there any statistical measures that can give us a better idea as to this stratification (some sort of score with some measure of standard deviation/variance?). To substantiate these claims, would it be possible to split their large cohort into training and validation sets to better understand reproducibility of their data?)***

Response:
Thank you for your comment. In our initial analysis of dupilumab responders, we stratified patients using unsupervised clustering based on clinical severity scores during treatment. However, following the insightful recommendations from other reviewers regarding the need for a clearer definition of clinical responder types, we have reclassified the patients according to a revised definition. The "Excellent" responder group now includes patients who achieved a mild disease level (defined by mEASI < 6.3) within the first month and maintained that level throughout the six months of dupilumab treatment. "Poor" responders are now defined as those who did not reach the mild level at any point during the six-month study period. Regrettably, involved a limited number of participants (n=24), which proved insufficient for conducting machine learning analyses on training and validation datasets. Consequently, we retrospectively examined the association between skin and blood biomarkers and treatment responses.

Accordingly, we have made modifications in the Results section (Page 9, lines 244-250), the Methods section (Page 30, lines 914-920). We have made new Figures 3A and 3B, and we modified Figures 5A, 5B, S13, S17, and S18 and Supplementary Table 16. In the revised manuscript, tables, and figures, all patients have been assigned new IDs (Pt01, etc.).

Comment 2: ***Can the authors give a more exact idea as to the inflammatory level of dysregulation in healthy, nonlesonal and lesional AD skin? E.g. via PCA plot or other measures? It would be good to understand the order of magnitude of transcriptomic dysregulation particularly in nonlesional AD skin, to also get a better sense and comparability to previous studies who have investigated the level of transcriptomic dysregulation on nonlesional AD skin.***

Response:
We appreciate your constructive comment. Accordingly, we have performed additional analyses to evaluate the degree of immune and barrier dysregulation in healthy, non-lesional and lesional AD skin. For this analysis, we assessed the correlation between the metagene scores (the sum of the expression values of AD-associated metagenes) and clinical severity. The results showed that the metagene expression score was significantly correlated with both local severity and total disease severity, as depicted in new Figures 2A to 2C. Additionally, we have included figures that illustrate the intensity of molecular dysregulation in representative patients (new Figures S10A to S10H) and we have modified the Results section in the revised manuscript (Page 7, lines 177-182).

Comment 3: ***Can the authors rule out that type 17 baseline activation is not just a measure of overall disease severity, but truly a predictor of dupilumab response?***

Thank you for your comment. Based on our relational analysis (as opposed to predictive analysis, which requires larger datasets including training and testing sets), we inferred that baseline activation of type 17 response was associated with a poorer response to dupilumab. Accordingly, we have removed the term “predictive biomarker” from the last paragraph of the discussion part to avoid overstatement (Page 18, line 519). Additionally, we conducted a regression analysis corrected for severity using the model: metagene expression ~ $\alpha + \beta 1$ (responder group) + $\beta 2$ (disease severity) + age + sex + ε . The results were insignificant when adjusted for disease severity, likely due to a strong confounding relationship between KC/FB1 expression and disease severity in our patients.
Further investigation is warranted in patients with different gene expression profiles, especially in patients with predominantly high type 2 and low type 17 inflammation profiles, to investigate whether type 17-related inflammation independently contributes

to reduced treatment efficacy. Nevertheless, as previously suggested in asthma (Hasegawa et al., 2017), multi-inflammatory pathogenesis may contribute to both severity and poor response to treatment.

Accordingly, we have deleted the following sentence from the Discussion section to avoid overstatement: "The limited number of metagenes associated with treatment response implies that patients with specific molecular profiles are at a higher risk of experiencing a poor response" (Page 17, lines 482-489). We have also added relevant description (Page 17, lines 488-490) to discuss the limitations of our analysis.

Reference:

Hasegawa, T., Uga, H., Mori, A. & Kurata, H. Increased serum IL-17A and Th2 cytokine levels in patients with severe uncontrolled asthma. *Eur. Cytokine Netw.* **28**, 8–18 (2017)

Reviewer #2 (Remarks to the Author):

Fukushima-Nokura and colleagues report a large-scale integration of phenotypic and molecular data in a prospective cohort of Atopic dermatitis patients. In a hypothesis-free approach, 29 gene clusters or metagenes were identified based on lesional transcriptome analysis. These clusters are subsequently correlated with blood biomarkers and investigated for their clinical utility to monitor disease severity as well as therapeutic response to dupilumab.

Great strenghtes of the manuscript are the large patient cohort and with that the valuable data resource, the novel method of minimally invasive 1mm punch biopsy collection that enables to obtain several biopsies in the same patient, and the prospective skin transcriptome analysis of patients under dupilumab treatment. General findings are that AD phenotypes are pretty stable over time, that several metagenes could be identified that were associated with disease severity, and that type 3 immunity associated metagenes are asscoiated with poor response to dupilumab.

We are deeply grateful to the reviewer for dedicating their valuable time to evaluate our manuscript. We also extend our sincere thanks for the insightful comments, which have significantly enhanced the presentation of our work.

Comment 1: ***Is there a correlation with the previously identified modules and endotypes that were published recently in a landmark paper by the same group (Sekita A et al, Nat Comm 2023)?***

Response:
We appreciate your insightful comment. Accordingly, we performed additional analysis to verify the similarity between the metagenes found in this study and the skin transcriptomic modules reported earlier (Sekita et al., 2023). We observed correlations in most disease-associated metagenes, with the exception of the barrier-related metagene. We ascribed this to the difference in the methodology; while the previous study used weighted gene co-expression network analysis (WGCNA), which assigns each gene to a single module, our approach using NMF allows a single gene to be clustered into multiple metagenes. Indeed, many of the highly-weighted genes in the barrier-related metagene (KC/MD2) showed high weights in several metagenes in our analysis, suggesting that these genes contribute to multiple pathways in human skin.

Based on these results, the Results section (Page 7, lines 158-160), Discussion section (Page 13, lines 367-370), and Methods section (Page 27, lines 800-811) have been modified, and Figure S8A and S8B has been added.

Reference:
Sekita, A. *et al.* Multifaceted analysis of cross-tissue transcriptomes reveals phenotype-endotype associations in atopic dermatitis. *Nat. Commun.* **14**, 6133 (2023)

Comment 2: ***Was the association of IL-22 and IL-18 blood cytokine levels to treatment outcomes corrected for severity? As these are known markers associated with AD severity.***

Response:
Initially, our correlation analysis of the pretreatment levels of blood biomarkers and treatment response was not corrected for disease severity. To address this, we conducted a regression analysis using the model: blood cytokine level ~ $\alpha + \beta 1$ (responder group) + $\beta 2$ (disease severity) + age + sex + ε. The results were insignificant, presumably due to the strong confounding relationship between IL-22 and IL-18 levels and disease severity in our patients.
Further investigation is warranted in patients with different serum cytokine profiles, such as highly severe patients with low serum IL-22 or IL-18, to investigate whether these cytokines are independent factors contributing to less treatment efficacy. We have also added relevant description (Page 17, lines 487-489) to discuss the limitations of our analysis.

Comment 3: ***The longitudinal analysis of AD patients under dupilumab therapy is unprecedented in its quality and delivers precious insights. Nevertheless, the definition of the clinical responder types is not broadly accepted and should be defined more clearly.***

Response:
We acknowledge your concern regarding the definition of clinical responder types. In response, we have reclassified the patients according to new criteria, as referenced in recent studies (Ferrucci et al., 2023); "Excellent" responders are now defined as patients who achieved a mild disease level (defined by mEASI < 6.3) within the first month and maintained it during six months of dupilumab treatment. "Poor" responders are those who did not reach the mild level at any point during the six-month period, and "Good" responders are those who do not fit the criteria of either "Excellent" or "Poor" responders. Consequently, this reclassification resulted in one patient being moved from the “Excellent” to the “Good” response group, and another from the “Poor” to the “Good” response group.

Accordingly, we have made modifications in the Results section (Page 9, lines 244-250), the Methods section (Page 30, lines 914-920). We have made new Figures 3A and 3B, and we modified Figures 5A, 5B, S13, S17, and S18 and Supplementary Table 16. In the revised manuscript, tables, and figures, all patients have been assigned new IDs (Pt01, etc.).

The two patients reassigned to new response groups were on oral cyclosporine at the start of dupilumab treatment and were thus excluded from the initial analysis of pretreatment biomarkers in the original manuscript. As a result, the findings of skin/blood biomarkers associated with treatment response (Figures 3C, 3D, 3E, S14, and S15) remain unchanged. Additionally, the eight patients included in the analysis of personal monitoring remained within their respective response groups as defined by the new criteria. Consequently, the results for these analyses (Figures 5C, 5D, S19, S20, S21, S22) were also unchanged.

Reference:
Ferrucci, S. *et al.* Predictive Factors of Early Response to Dupilumab in Patients with Moderate-to-Severe Atopic Dermatitis. *J. Clin. Med. Res.* **12**, (2023)

Comment 4: ***The authors have obtained more than one lesional biopsy from many AD patients – it would highly upvalue the study to show some analysis and comment on stability depending on the lesional site and over time within one individual AD patient. Is the distribution of metagenes stable in an individual patient or may it change?***

Response:
We deeply appreciate the reviewer’s suggestions. As observed in our cross-sectional analysis, the transcriptomic patterns in the skin are significantly influenced by the

clinical phenotype of the sampled lesion and overall disease severity. Consequently, we further investigated the transcriptional profiles of samples with consistent local severity, obtained from the same patient at different time points (n = 7 subjects) to assess the stability of expression profiles within individuals. Patients on the same treatment at the two-time points were analyzed to exclude treatment effects on the expression profiles.

Accordingly, we have added Figure S11C descriptions in the Method section (Page 29, lines 864-871) in the revised manuscript. The left column panel illustrates the skin metagene expression sampled on different dates from lesions of the same patient with the same local severity. The right panel shows the distribution of skin metagene expression across all the lesions of AD patients. The metagene expressions of samples obtained at different time points showed no statistical significance.

Comment 5: ***As a proof of principle: does it give a higher predictive value to determine the KC/FB1 metagene as compared to CCL20 (or IL-19, IL-17A itself) alone to predict dupilumab response in AD patients?***

Response:
Thank you for your comment. We conducted additional analyses to examine the correlation between disease severity scores post-treatment and pretreatment levels of skin and blood biomarkers. We found that the average mEASI scores at 5 and 6 months after the initiation of dupilumab treatment significantly correlated with baseline expression of KC/FB1 (type 17 response) in lesional sites, but not with plasma CCL20 levels. These findings, combined with the results from the stratified treatment response groups, suggest that skin gene expression may serve as a more effective biomarker for treatment response. In our analysis, IL-17 showed insignificant results, and we were unable to evaluate IL-19 as it was not measured in our study.

The results and methods of additional analysis are presented in Results section (Page 10, lines 276-281), Methods section (Page 29, lines 864-871). Figure 4A, 4B, and 4C has been additionally presented.

Comment 6: ***Minor: line 476: IGE instead of IEG***

Response:
We apologize for this oversight. We have now replaced "IEG" with "IGE" in the revised manuscript.

Comments from Reviewer 3
Comment 1: ***The analyses and data presented here are very hard to follow, and the data is presented in a manner that is hard to interpret and understand. It is hard to***

appreciate the biological implications of what a KC/FB/MD component/module is, and a concern, as shown in Figure 2b, is the extensive overlap between many of these components/modules, making interpretation difficult.

Response:
We apologize for any lack of clarity and interpretability in our initial presentation and analysis of data. To address these concerns, we have revised the whole manuscript to enhance the clarity and accessibility of our data presentation. Accordingly, we have changed the title, and rewritten the Abstract, Introduction, and the first half of the Results section.
For Figure 2B, we considered that blood cytokine levels represent individuals' overall inflammatory status and that it is biologically natural that multiple cytokines are associated with several skin gene profiles that may overlap in an individual.

Comment 2: ***What is the basis of the definition of "excellent", "good" and "poor" response groups? What defined as "rapid response to dupilumab" – this is very unclear. This needs to be clarified and if the authors can refer back to an accepted definition of these response it would strengthen this type of stratification, which otherwise seems arbitrary.***

Response:
We appreciate your insightful comment.
We acknowledge your concern regarding the definition of clinical responder types. To address this, we have reclassified the patients according to new criteria, as referenced in recent studies (Ferrucci et al., 2023). We have newly defined "Excellent" responders as patients who achieved a mild disease level (defined by mEASI < 6.3) within the first month and maintained it during six months of dupilumab treatment. "Poor" responders are now those who did not reach the mild level at any point during the six-month period, and "Good" responders are those who do not fit the criteria of either "Excellent" or "Poor" responders. Consequently, this reclassification resulted in one patient being moved from the "Excellent" to the" Good" response group, and another from the "Poor" to the "Good" response group.

Accordingly, we have made modifications in the Results section (Page 9, lines 244-250), the Methods section (Page 30, lines 914-920). We have made new Figures 3A and 3B, and we modified Figures 5A, 5B, S13, S17, and S18 and Supplementary Table 16. In the revised manuscript, tables, and figures, all patients have been assigned new IDs (Pt01, etc.).

The two patients reassigned to new response groups were on oral cyclosporine at the start of dupilumab treatment and were thus excluded from the initial analysis of pretreatment biomarkers in the original manuscript. As a result, the findings of skin/blood biomarkers associated with treatment response (Figures 3C, 3D, 3E, S14, and S15) remain unchanged. Additionally, the eight patients included in the analysis of personal monitoring remained within their respective response groups as defined by the new criteria. Consequently, the results for these analyses (Figures 5C, 5D, S19, S20, S21, S22) were also unchanged.

Reference:
Ferrucci, S. *et al.* Predictive Factors of Early Response to Dupilumab in Patients with Moderate-to-Severe Atopic Dermatitis. *J. Clin. Med. Res.* **12**, (2023)

Comment 3: ***The basis of the type 17 immune response seems weak. The authors need to address and show what are the metagenes associated with "IL-17 inflammation-induced responses". Also, with bulkRNA-seq the data should be sensitive enough to pick up expression of IL17A or IL17F. If those are absent from the bulkRNA-seq data what is the reliability of this type-17 response? Also, does this response differentiate from other IL-17 members such as IL-17C or other non-IL-17 family members that act with IL-17 such as TNF?***

Response:
Thank you for highlighting this issue. Our study of 1 mm punch biopsy focuses on transcriptional biomarkers derived from keratinocytes and fibroblasts expressed under various inflammatory signals in the skin. Therefore, our biological interpretation of the biomarkers is based on accumulated knowledge on the characteristic and specific responsive expression profiles of the keratinocytes and fibroblasts.

To further support our annotation of the "type 17 response", we have now performed additional enrichment analyses presented in the new Figures S9A to S9C. Figure S9A displays the results of the KEGG pathway analysis conducted on the highly weighted genes, including the "IL-17 signaling pathway". Figures S9B and S9C illustrate the significance of the number of shared genes included in the reference data and highly weighted or marker genes, tested by Fisher's test. The reference data represent genes induced in keratinocytes by IL-17-related cytokines, as reported by the publications below. Overall, the results were all significant for IL-17A, IL-17F, and TNFa, with less significant results for IL-17C. This suggests that KC/FB1 also represents the transcriptomic response to IL-17 and its co-working cytokines, such as TNFa.
Accordingly, changes have been made in the Results section (Page 7, line 162) and Methods section (Page 25, lines 746-754) of the manuscript.

Reference:
Gupta, R. K. *et al.* TWEAK functions with TNF and IL-17 on keratinocytes and is a potential target for psoriasis therapy. *Sci Immunol* **6**, eabi8823 (2021)

Jiang, Y. *et al.* Suppression of TCF4 promotes a ZC3H12A-mediated self-sustaining inflammatory feedback cycle involving IL-17RA/IL-17RE epidermal signaling. *JCI Insight* **9**, (2024)

Nograles, K. E. *et al.* Th17 cytokines interleukin (IL)-17 and IL-22 modulate distinct inflammatory and keratinocyte-response pathways. *Br. J. Dermatol.* **159**, 1092–1102 (2008)

Tollenaere, M. A. X. *et al.* Signalling of multiple interleukin (IL)-17 family cytokines via IL-17 receptor A drives psoriasis-related inflammatory pathways. *Br. J. Dermatol.* **185**, 585–594 (2021)

Swindell, W. R. *et al.* Heterogeneity of inflammatory and cytokine networks in chronic plaque psoriasis. *PLoS One* **7**, e34594 (2012)

Comment 4: ***Why are the serum levels shown as log10 values(+0.1)? Is this based on the pg /ml measurements? Please clarify.***

Response:
As with our data, serum cytokine data often exhibit a skewed distribution. We therefore utilized log-transformed values in the analysis to mitigate the effects of skewed data and enhance data visualization and interpretability. To avoid negative infinite values at 0 pg/mL, we added 0.1 prior to the log transformation. This transformation, which has also been employed in previous studies of serum cytokines (Hasegawa et al., 2019; Xu et al., 2023), helps to normalize the data distribution and to highlight relative changes more clearly.

Accordingly, we have added Figure S23 to show the distribution of original cytokine and log-transformed levels. The axis title of relevant Figures 3E, 4C, 5C, S15, S16, S18, S20 was modified. The manuscript has also been modified in the Methods section (Page 23, lines 662-666).

Reference:
Hasegawa, T., Okazawa, T., Uga, H., Kurata, H. & Mori, A. Serum CXCL9 as a potential marker of Type 1 inflammation in the context of eosinophilic asthma. Allergy 74, 2515–2518 (2019)

Xu, Y.-Q. *et al.* Peripheral Blood Inflammatory Cytokine Factors Expressions are Associated with Response to Acupuncture Therapy in Postpartum Depression Patients. *J. Inflamm. Res.* **16**, 5189–5203 (2023)

Comment 5: ***It is unclear what the data quality was – how many differentially expressed genes between the different groups? Was the data of sufficient quality to differentiate between psoriasis and atopic dermatitis? This would be helpful to have. Please provide information about the quality of the data from the RNA-biopsies including RIN scores.***

Response:
We appreciate your request for more information regarding the data quality. In terms of the RNA biopsy quality, we believe the data was sufficient to assess the molecular profiles in psoriasis and AD. To address your concerns, we have provided the list and numbers of differentially expressed genes (DEGs) between AD, psoriasis, and healthy samples (revised Figure S2A, Supplementary Table 4). We identified substantial numbers of DEGs in each disease group, which were generally consistent with previous studies. However, as our objective in this study was to assess the heterogeneity in the disease population, the DEGs reflecting the differences in the averaged disease population, are only presented only as supplementary data in the revised manuscript. The manuscript has been modified in the Results section (Page 6, lines 122-125) and Methods section (Page 22, lines 645-646, 650-652).

As described in the original manuscript, we have excluded low-quality samples from the analysis, specifically removing three samples with a read count lower than $10^{6.5}$ and seven samples with a RIN lower than 2 or unmeasurable. Also, the values and distribution of RIN scores, read counts, and RNA amounts for sample and data quality were presented in the original Figure S5. In the revised manuscript, we have also included the data for these markers and the mapping rate for each sample in Supplementary Table 19. We believe this additional information will enhance the robustness of our findings.

Comment 6: ***The abstract could be made clearer. The size of the cohort is impressive, but the findings are not clearly highlighted, and it is not clear what was assessed and how. The abstract would benefit from some rewriting. In the body of the manuscript, there are similar awkward sentences and grammatical structures that could benefit from careful read-over by a native English speaker.***

Response:

Per your suggestions, we have rewritten the abstract to clearly state what was assessed, ensuring the key findings are highlighted. Additionally, our manuscript was reviewed and edited by a native English speaker, Dr. Barrows from RIKEN, and proofread by Editage, a professional English language editing service provider to enhance its readability and coherence. We have also attached certificates of proofreading in English.

Comment 7: ***I do not agree with the authors that the 1mm punch biopsy reveals robust expression profiles in AD. It appears that there is a high level of variability in the data that the authors are getting, and this reflects the way the data is analyzed and presented.***

Response:
Thank you for your valuable feedback. We understand your concern regarding the robustness and variability of the expression profiles obtained from the 1mm punch biopsy. We acknowledge that variability in clinical samples generally arises from multiple factors, including technical bias, such as sample quality and sampling bias, as well as pathophysiological heterogeneity in clinically diagnosed diseases. To address this, we used Non-negative Matrix Factorization (NMF) to decompose the original expression data. NMF has been widely used in bioinformatic analyses to unbiasedly dissect the molecular complexity of heterogenous diseases to reveal previously unrecognized pathological molecular pathways and pathogenic subtypes (Zhao et al., 2018; Bolouri et al., 2018). This approach allowed us to validate gene clusters associated with various clinical parameters while excluding gene clusters related to technical biases. We believe that the significant association between the filtered gene clusters and clinical factors, such as skin phenotypes and treatment response, underscores the quality of the 1mm punch biopsy samples and highlights the potential utility of the gene biomarkers in clinical settings.

Reference:
Zhao, L., Zhao, H. & Yan, H. Gene expression profiling of 1200 pancreatic ductal adenocarcinoma reveals novel subtypes. *BMC Cancer* **18**, 603 (2018)
Bolouri, H. *et al.* The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat. Med.* **24**, 103–112 (2018)

Comment 8: ***The Discussion is very long and should be condensed.***

Response:
We acknowledge that the discussion section needed to be shorter. We have revised and reduced it by approximately 1,000 words to ensure it is more concise and focused while

still covering all the critical points. We believe that the revised and shortened discussion is now appropriate in length while also retaining all major discussion points.

Additional clarifications

In addition to the above comments, all spelling and grammatical errors have been corrected. The genes shown in Figure 1B have been modified to more significant genes in each metagene. We have added the legends of Figure S22 and Supplementary tables, which needs to be included. The axis titles of Figure 3, S14 have been modified to make clearer that they are pretreatment levels of biomarkers. Figure legends and column titles of Supplementary tables have also been modified for greater clarity.

We look forward to hearing from you soon regarding our submission and to responding to any further questions and comments you may have.

Sincerely,
Masayuki Amagai
2024-07-19

Responses to Reviewers

Reviewers' comments are itemized below in bold font.
The authors' responses to each specific comment are in blue font with associated line numbers of new text and Figures in red.

Reviewer #3 (Remarks to the Author):
The authors have done quite extensive revisions and some of my concerns have been addressed. However, several still remain - my main concern with the data, while impressive in terms of sample size, is the quality of the transcriptomic data, the lack of validation of the Th17 response, and the definition of treatment response to dupilumab, which I feel is overstated. My remaining concerns are listed below:

Response:

Thank you for your continued feedback and thoughtful critique. We regret any confusion caused by the description of our findings and methodology in the manuscript, and we appreciate the opportunity to clarify. Our study's goal was to uncover co-expressed gene profiles in AD skin using a hypothesis-free, data-driven approach and to explore their relationship with clinical parameters. In doing so, we relied on unsupervised decomposition analysis, which we have carefully validated by referencing previously published studies on genes induced in keratinocytes and fibroblasts under various inflammatory signals. This validation confirmed that our gene sets align with known biological pathways in inflammatory skin diseases, including those associated with Th1, Th2, and Th17 responses.

Additionally, our analysis revealed novel genes that correlate with recognized gene sets, offering new insights into the potential interactions between keratinocytes and fibroblasts, which were previously unrecognized in AD skin. To further clarify our methodology and reduce any potential for misunderstanding, we have renamed all the metagenes to "SKIn-Tissue derived Metagenes" (SKITm 1-29) and made corresponding updates throughout the currently revised manuscript and figures.

We hope these revisions help address your concerns, and we are confident that the updated manuscript more clearly conveys our findings and methodology.

Comment 1:
In the Ferrucci study that the authors refer to, the cutoff for mild disease was EASI<7, which I agree is close, but I am not sure why the authors do not follow that but instead select <6.3 as the cutoff in this study. Therefore, this specific cutoff seems arbitrary and not adequately justified. How much would the findings change if the cutoff was the same as in the Ferrucci study?

Response:

We used the cutoff of 6.3 based on the cutoff for mild disease in “modified EASI” (mEASI), rather than the EASI score. The EASI is a composite score of AD symptoms (erythema, induration, lichenification and excoriation) across four anatomical regions: head/neck, upper extremities, trunk and lower extremities. mEASI is calculated by excluding the scores of the head and neck from the total EASI score and is widely used in AD clinical trials (Nakagawa et al. J Allergy Clin Immunol. 2019;144: 1575–1583). The cutoff of 6.3 for mEASI corresponds to an EASI score of 7 for mild disease (Leshem YA et al. Br J Dermatol. 2015;172: 1353–1357). The basis of cutoff of mEASI was presented in the Methods section in the originally submitted manuscript (Page 20, lines 582-593 in the current revised manuscript).

We adopted mEASI in this study because dupilumab induces facial erythema with a distinct etiology from pre-existing AD (termed as dupilumab-associated facial erythema) in 10-19.2 % of patients, which can potentially confound the EASI score's ability to reflect the true treatment effects (Jang et al. Sci Rep. 2021;11: 23539, Okiyama et al. J Eur Acad Dermatol Venereol. 2020;34: e474–e476, Jo et al. J Am Acad Dermatol. 2021;84: 1339–1347). To clarify the reasoning of using mEASI, we modified the Results section of the currently submitted manuscript (revision 2) (Page 10 lines 248 to 251).

Comment 2:

Also, a significant limitation of this definition and the data presented here is that it doesn’t consider the disease severity before treatment. Therefore, the “excellent” vs. “poor responders” is more or less meaningless, and if this was considered as an EASI response (i.e., EASI50/75/90), I'm not sure it would be defined as an excellent response. It’s very clear, based on the clinical data shown in Figure 3 and S13 that their “excellent responders” had milder disease at baseline than the other two groups and the authors acknowledge this in the results section – my concern, therefore, is that this signature doesn’t necessarily reflect, a response to dupilumab treatment but rather reflects baseline disease activity differences between these groups – it also seems counterintuitive that patients with an overall “poor” response to dupilumab would still be on the drug for 20 months as shown in Figure S13 - this is a major concern and needs to be addressed as the therapeutic stratification is a major message of this manuscript

Response:

We appreciate your careful review and acknowledge the importance of addressing baseline disease severity when interpreting treatment responses. As mentioned in the originally submitted manuscript, all patients included in the study had moderate to severe AD before starting dupilumab treatment. We followed the severity scale published by Leshem et al. (Br J Dermatol. 2015;172: 1353–1357), where mild

AD is defined as EASI 0–7, moderate as 7.1–21.0, and severe as 21.1–50. Mild patients were excluded from this study, ensuring that all enrolled patients had moderate or severe AD. Following 6 months of treatment, 14/22 patients achieved a mEASI75 (75% improvement in the EASI from baseline), and 20/22 patients achieved a mEASI50 (50% improvement in the EASI from baseline), indicating measurable improvements in the majority of patients. Notably, all "Excellent" and "Good" responders reached EASI50, as did 5 out of 7 patients in the "Poor" responder group, which explains why they continued with dupilumab treatment for an extended period.

While EASI50/75/90 is commonly used in clinical trials, recent studies suggest that the absolute value of severity scores (rather than percentage improvement) may provide a more accurate measure of treatment outcomes in real-world settings (Thyssen et al. J Dermatolog Treat. 2023;34: 2216322, Seidl et al, Dermatol Ther. 2022;35: e15193). Furthermore, AD often exhibits a fluctuating course, with periods of exacerbation and remission, even with effective systemic therapies, making it challenging to assess treatment efficacy at a single time point (Hendricks et al, Am J Clin Dermatol. 2019;20: 565–569). Therefore, we categorized patients based on their absolute severity scores over six months, following an approach from a previous study (Ferrucci et al, J Clin Med Res. 2023;12. doi:10.3390/jcm12206575). Specifically, we defined patients who achieved and maintained mild disease (defined by mEASI ≤ 6.3) within the first month as one group, those who never reached mild disease as another, and a third group for those who fell in between. These methods and data have been outlined in the Results and Methods sections of the originally submitted manuscript (Page 10 lines 252 to 257, Page 31 lines 916 to 922 in the currently revised version).

Recent studies have shown that pre-treatment clinical severity and its correlated blood biomarkers are associated with the efficacy of dupilumab (Ferrucci et al, J Clin Med Res. 2023;12. doi:10.3390/jcm12206575, Kido-Nakahara et al, J Allergy Clin Immunol Glob. 2024;3: 100317). While we acknowledge that baseline severity plays a role, our study specifically focuses on the molecular profiles in the skin associated with treatment response. Notably, in the original analysis, we have shown that only some of the gene signatures that correlate with clinical severity are associated with therapeutic efficacy. Specifically, whereas the cross-sectional analysis showed five metagenes linked to disease severity (metagenes enriched with genes associated with epidermal barrier, ECM organization, transcriptional factors, type 2 and type 17 response), only the metagenes associated with ECM organization, transcriptional factors, and type 17 response were correlated with treatment response.

To further assess the independence of these biomarkers from disease severity, as pointed out by you and other reviewers, we conducted a regression analysis adjusted for disease severity using the model: metagene expression $\sim \alpha + \beta1$ (responder group) $+ \beta2$ (disease severity) + age + sex + ε. The results were insignificant when adjusted for disease severity, likely due to a strong confounding relationship between associated metagene expression and disease severity in our patients. Further investigation is

warranted in patients with different gene expression profiles, especially in patients with predominantly high type 2 and low type 17 response profiles, to investigate whether specific endotype independently contributes to reduced treatment efficacy. As suggested in asthma (Hasegawa et al, Eur Cytokine Netw. 2017;28: 8–18), multi-inflammatory pathogenesis in AD may contribute to both severity and poor response to treatment. These limitations, along with our observations, are discussed in our previously revised manuscript (revision 1) (Page 18, lines 498 to 504 in the currently revised manuscript).

That said, we believe our findings provide valuable insights into the relationship between specific gene signatures and dupilumab treatment responses in AD.

In response to your feedback, in our currently revised manuscript (revision 2), we have clarified the criteria for defining responder groups and renamed them for greater accuracy. Specifically, “Excellent responders” are now labeled “Early responders,” and “Good responders” are labeled “Intermediate responders.” These changes have been reflected in the Results section and throughout the currently revised manuscript (Page 10 lines 252 to 306, Figures 3, 5, S14-S23).

Comment 3:

The authors have not adequately addressed my comment regarding the IL-17 response and what it means. Again, the authors are using specific modules and not showing expression of IL17A, IL17F, or IL17C as requested and not addressing what it means to have an enriched “Th17 module” if there is no detectable expression of IL17A/F. The authors should note that many of the genes in those IL-17 signatures are induced by multiple other proinflammatory mediators, that is why many of these genes are showing up across multiple modules (i.e., in Figure S9B). The overlap between IL-17A and TNF appears quite significant. Therefore, this needs further validation and the authors should show results for IL17A/F and 17C from their bulk RNA-seq data.

Response:

We apologize for any confusion in the presentation of our findings and thank the reviewer for raising this important point. Our study’s goal was to identify co-expressed gene profiles in AD skin and examine their association with clinical parameters using a hypothesis-free, data-driven approach. By applying unsupervised decomposition analysis to our whole-skin transcriptome dataset obtained in a large-scale study, we were able to uncover hidden gene interactions and explore their relevance to established biological pathways in inflammatory skin diseases (including type 1, 2, and 17 responses).

The fact that our gene sets aligned with established biological pathways in keratinocytes and fibroblasts under various inflammatory conditions (described as “type

1, 2, 17 response" in our original manuscript) reinforced the validity of our unsupervised analysis.

Regarding the specific expression of IL17A, IL17F, and IL17C, the expression levels of these cytokines in our RNA-seq data were very low, with mean TPM values of 0.051, 0.076, and 0.975, respectively, making these values unreliable for direct interpretation. In our originally submitted manuscript, we provided biological annotations of the representative genes within each metagene (as shown in Table 2 and Supplementary Table 9). Additionally, we performed enrichment analysis in the previous revision (revision 1), which demonstrated significant enrichment of genes associated with type 17 responses, as depicted in Figure S5. We agree with the reviewer's point that many genes associated with the IL-17 signature are also induced by other pro-inflammatory mediators. This overlap explains why some of these genes are shared across multiple modules.

Notably, our analysis also revealed novel genes that correlate with recognized genes, providing insights into previously unrecognized gene interactions of keratinocytes and fibroblasts in AD skin. We believe our biomarkers derived from keratinocytes and fibroblasts hold significant value for future studies in clinicopathology of AD.

We have revised and expanded the description of our methodology and results to further clarify this point in the Abstract, Results and Discussion section of the currently revised manuscript (Page 3 lines 39 to 53, Page 6 lines 132 to 189, Page 9 lines 224 to 237, Page 12 lines 329 to 331, Page 13 lines 368 to 371, Page 15 lines 405 to 407, Page 18 lines 509 to 510, Page 18 lines 525 to 526). Additionally, to emphasize the hypothesis-free nature of our analysis and that the identified gene sets represent unbiased co-expression profiles in AD skin, we have renamed the metagenes to "SKIn-Tissue derived Metagenes" (SKITm 1-29). Corresponding revisions have been made throughout the Results, Discussion, Methods sections, and the Figures in the revised manuscript (the entire manuscript in Results, Discussion and Methods section, Figures 1, 2, 3, 4, 5, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S20, S21, S22, S23).

We believe that our findings provide valuable biomarkers derived from keratinocytes and fibroblasts for future clinicopathological studies in AD.

Comment 4:

The authors state that they provide RIN numbers, but I cannot find these data. Selecting the RIN cutoff at RIN<2 is concerning. Most groups that do this type of work draw the line at RIN>6. How many of the samples were above that threshold (RIN=6) and how many below?

Response:

Thank you for raising this important point. The values and distribution of RIN scores, read counts, and RNA amounts were indeed presented in the original manuscript in Figure S7. Additionally, in response to your previous comments, we have included detailed presentation of RNA quality data, including RIN scores, in the previously submitted manuscript (revision 1) in Supplementary Table 19.

As shown, 776 out of 1036 samples had RIN scores greater than 6. To ensure the robustness of our analysis, we carefully excluded gene expression profiles that were associated with low RIN scores, as described in our originally submitted manuscript. We have shown that this approach helps to minimize biases related to sample quality and ensures that the molecular data reflect biologically relevant signals associated with the pathogenesis of AD. These steps were detailed in the Results and Methods sections in the originally submitted manuscript (Page 7 lines 151 to 159, Page 26 lines 762-769, Figure S7A-C and S8 in the current revision).

Responses to Editor

Reviewers' comments are itemized below in bold font.
The authors' responses to each specific comment are in blue font with associated line numbers of new text and Figures in red made in the Revision #1 (version -A), Revision #2 (version -B) and Revision #3 (version -C).

Reviewer #3 -A round comments
The authors have done quite extensive revisions and some of my concerns have been addressed. However, several still remain - my main concern with the data, while impressive in terms of sample size, is the quality of the transcriptomic data, the lack of validation of the Th17 response, and the definition of treatment response to dupilumab, which I feel is overstated. My remaining concerns are listed below:

Response:

Thank you for your continued feedback and thoughtful critique. We regret any confusion caused by the description of our findings and methodology in the manuscript, and we appreciate the opportunity to clarify. Regarding sample quality, we have conducted an additional analysis demonstrating the robustness of our study using skin samples with low RIN scores (new Figure S1D). Regarding the severity score applied in the study, we have further substantiated this approach with clinical photos demonstrating facial erythema in our patients, highlighting the appropriateness of mEASI for evaluating disease severity during dupilumab treatment (new Figure S14A). Regarding treatment response group classifications, we have clarified the rigorous methodology underlying the determination of treatment response groups, which aligns with previously published frameworks.

Additionally, we have demonstrated the correlation between pretreatment severity and the elevated Type 17/Type 2 skin expression ratio in our cohort, while acknowledging the limitation that the biomarkers are confounded by disease severity (new Figure S24). Regarding validation of the "type 17 response" annotation, while we emphasize that the annotation and naming of the gene set are not central to our core findings, we conducted an additional validation experiment as suggested. Specifically, we performed RT-PCR analysis on the remaining available RNA samples of AD skin tissues to quantify the expression of key cytokines, including IL17A, in the samples used for the cross-sectional analysis. This analysis confirmed a correlation between Type 17 cytokine expression levels and the RNA-seq-derived metagenes, thereby validating our annotation. These new

RT-PCR results have been incorporated into the revised manuscript (new Figure 1C). The details are explained in the response as follows.

We hope these revisions help address your concerns, and we are confident that the updated manuscript more clearly conveys our findings and methodology.

Comment #1:

In the Ferrucci study that the authors refer to, the cutoff for mild disease was EASI<7, which I agree is close, but I am not sure why the authors do not follow that but instead select <6.3 as the cutoff in this study. Therefore, this specific cutoff seems arbitrary and not adequately justified. How much would the findings change if the cutoff was the same as in the Ferrucci study?

Response:

In the longitudinal study of dupilumab-treated patients, we categorized patients based on their absolute severity scores over six months, following an approach from a previous study (Ferrucci et al, J Clin Med Res. 2023). Whereas the Ferrucci study used EASI for disease severity scale, we adopted "modified EASI" (mEASI). mEASI is calculated by excluding the scores of the head and neck from the total EASI score and widely used in Atopic Dermatitis (AD) clinical trials (Nakagawa et al. J Allergy Clin Immunol. 2019). The cutoff of 6.3 for mEASI used in our manuscript corresponds to an EASI score of 7 for mild disease (Leshem YA et al. Br J Dermatol. 2015). The basis of cutoff of 6.3 of mEASI was presented in the Methods section in the original manuscript as follows: "The suggested EASI range for a "mild" condition is 0–7, 7.1–21.0 for "moderate," and 21.1–50 for "severe," based on previous literature. Therefore, the "mild" mEASI was set at 6.3, calculated as 7×64.8/72 (based on the maximum score of EASI being 72 and that of mEASI being 64.8)." (Page 21, lines 595 to 606 of Revision #3, version -C).

The reason we adopted mEASI in our study is that dupilumab is known to induce facial erythema in 10-19.2 % of patients, termed as "dupilumab-associated facial erythema" (DFE). Whereas pruritus is the hallmark symptom of AD, DFE lacks pruritus and has a distinct etiology from pre-existing AD (Jang et al. Sci Rep. 2021, Okiyama et al. J Eur Acad Dermatol Venereol. 2020, Jo et al. J Am Acad Dermatol. 2021). Since EASI score does not differentiate based on the presence of pruritus, DFE can potentially confound the EASI score's ability to reflect the true treatment effects in dupilumab treated patients. This rationale was incorporated in the manuscript in Revision #2, version -B (Page 10, lines 256 to 259 of Revision #3, version -C).

In addition, in Revision #3 (version -C), as an example, we have presented a clinical case suspected of DFE and corresponding photo of the patient from our cohort as new Figure S14A. Although facial erythema was included in the EASI score—since the score does not differentiate based on the presence of pruritus—it is important to note that in this patient, the facial erythema at 4 months was not pruritic. Meanwhile, significant improvement was observed in the patient's pre-existing AD symptoms on the trunk and extremities. Based on these observations, we believe the facial erythema of this patient was likely due to DFE, and mEASI more accurately reflects the treatment effects of dupilumab, as it excludes facial scores that may be influenced by DFE.

As described above, none of the methodology regarding the disease severity score and the cutoff used is arbitrary. We believe that the modifications introduced in our manuscript have significantly improved the clarity and rigor of our manuscript.

[FIGURE REDACTED. Please refer to current Supplementary Figure 13.]

Reference:
Ferrucci, S. *et al.* Predictive Factors of Early Response to Dupilumab in Patients with Moderate-to-Severe Atopic Dermatitis. *J. Clin. Med. Res.* **12**, (2023). https://www.mdpi.com/2077-0383/12/20/6575

Nakagawa, H., Nemoto, O., Igarashi, A., Saeki, H., Oda, M., Kabashima, K., & Nagata, T. (2019). Phase 2 clinical study of delgocitinib ointment in pediatric patients with atopic dermatitis. *The Journal of Allergy and Clinical Immunology*, *144*(6), 1575–1583. https://www.sciencedirect.com/science/article/pii/S0091674919310450

Leshem, Y. A., Hajar, T., Hanifin, J. M., & Simpson, E. L. (2015). What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis: an interpretability study. *The British Journal of Dermatology*, *172*(5), 1353–1357. https://academic.oup.com/bjd/article-abstract/172/5/1353/6616225

Jang, D. H., Heo, S. J., Kook, H. D., Lee, D. H., Jung, H. J., Park, M. Y., & Ahn, J. (2021). A 52 weeks dupilumab treatment for moderate to severe atopic dermatitis in Korea: long-term efficacy and safety in real world. *Scientific Reports*, *11*(1), 23539. https://www.nature.com/articles/s41598-021-02950-4

Okiyama, N., Nakamura, Y., Ishitsuka, Y., Inoue, S., Kubota, N., Saito, A., Watanabe, R., Fujisawa, Y., & Igawa, K. (2020). Successful topical treatment with ketoconazole for facial rashes refractory to dupilumab in patients with atopic dermatitis: case reports. *Journal of the European Academy of Dermatology and Venereology: JEADV*, *34*(9), e474–e476. https://onlinelibrary.wiley.com/doi/10.1111/jdv.16383

Jo, C. E., Finstad, A., Georgakopoulos, J. R., Piguet, V., Yeung, J., & Drucker, A. M. (2021). Facial and neck erythema associated with dupilumab treatment: A systematic review. *Journal of the American Academy of Dermatology*, *84*(5), 1339–1347. https://www.sciencedirect.com/science/article/abs/pii/S0190962221000797

Comment #2:

Also, a significant limitation of this definition and the data presented here is that it doesn't consider the disease severity before treatment. Therefore, the "excellent"

vs. “poor responders” is more or less meaningless, and if this was considered as an EASI response (i.e., EASI50/75/90), I'm not sure it would be defined as an excellent response. It’s very clear, based on the clinical data shown in Figure 3 and S13 that their “excellent responders” had milder disease at baseline than the other two groups and the authors acknowledge this in the results section – my concern, therefore, is that this signature doesn’t necessarily reflect, a response to dupilumab treatment but rather reflects baseline disease activity differences between these groups – it also seems counterintuitive that patients with an overall “poor” response to dupilumab would still be on the drug for 20 months as shown in Figure S13 - this is a major concern and needs to be addressed as the therapeutic stratification is a major message of this manuscript.

Response:

Thank you for pointing this out. Similar critiques concerning disease severity and biomarker independence were raised in the first revision by Reviewers #1 and #2. In our longitudinal study of dupilumab-treated patients, all patients had moderate to severe AD before treatment, and mild patients were excluded from this study. As the reviewers pointed out, relatively more moderate cases had better treatment outcomes. Therefore, to further assess the independence of the biomarkers from disease severity, we conducted a regression analysis adjusted for disease severity using the model: metagene expression ~ $\alpha + \beta1$ (responder group) + $\beta2$ (disease severity) + age + sex + ε. The results were insignificant when adjusted for disease severity, likely due to a strong confounding relationship between associated metagene expression and disease severity in our patients. In our longitudinal study, as demonstrated in Figure S24 in Revision #3 (version -C), the pretreatment severity of patients was strongly correlated with an increased Type 17/Type 2 ratio, characterized by skewed SKITm10 expression (Type 17-associated genes) relative to SKITm17 expression (Type 2-associated genes). As severe patients who exhibited solely-Type 2 profile were rare, we could not investigate the independency of Type 17-associated genes as biomarkers related to treatment response. Further investigation is warranted in patients with different gene expression profiles, especially in severe patients with predominantly high type 2 and low type 17 response profiles, to investigate whether specific endotype independently contributes to reduced treatment efficacy.

That said, in our study, we have shown that only some of the gene signatures that correlate with clinical severity are associated with the therapeutic efficacy of dupilumab. Specifically, whereas the cross-sectional analysis showed five metagenes linked to disease severity (metagenes enriched with genes associated with epidermal

barrier, ECM organization, transcriptional factors, type 2 and type 17 response), only the metagenes associated with ECM organization, transcriptional factors, and type 17 response were correlated with treatment response, suggesting their specificity as biomarkers.

In response to the comments from Reviewers #1 and #2, we have accordingly documented the limitations of our findings in Revision #1, version -A, which was accepted by Reviewers #1 and #2. In Revision #3 (version -C), we have further modified the manuscript to clarify this point as follows: “Given that many biomarkers related to dupilumab response were confounded by disease severity, further investigation is necessary to assess their independence from disease severity.” (Discussion section: Page 18 lines 515 to 517 of Revision #3, version -C).

As for the method used to define the treatment responders of dupilumab, in Revision #1 (version -A), in response to feedback from Reviewers #1, #2 and #3, we revised the criteria for stratifying patients, referencing a previously published study (Ferrucci et al. J. Clin. Med. 2023). Initially, in the original manuscript, we utilized hierarchical clustering based on the severity score over six months of therapy. However, acknowledging the reviewers’ concerns regarding the potential variability in clustering results across datasets, we categorized patients based on their absolute severity scores over six months, following an approach from the Ferrucci’s study: “Excellent” responders achieved a mild disease level within the first month and maintained it over six months of dupilumab treatment. “Poor” responders failed to reach a mild level at any point, and “Good” responders were those who fell between these criteria. The cutoff of disease severity score used for defining "mild disease" is detailed in our response to Comment #1. Of note, Ferrucci’s study has also observed that pretreatment severity correlates with treatment response to dupilumab.

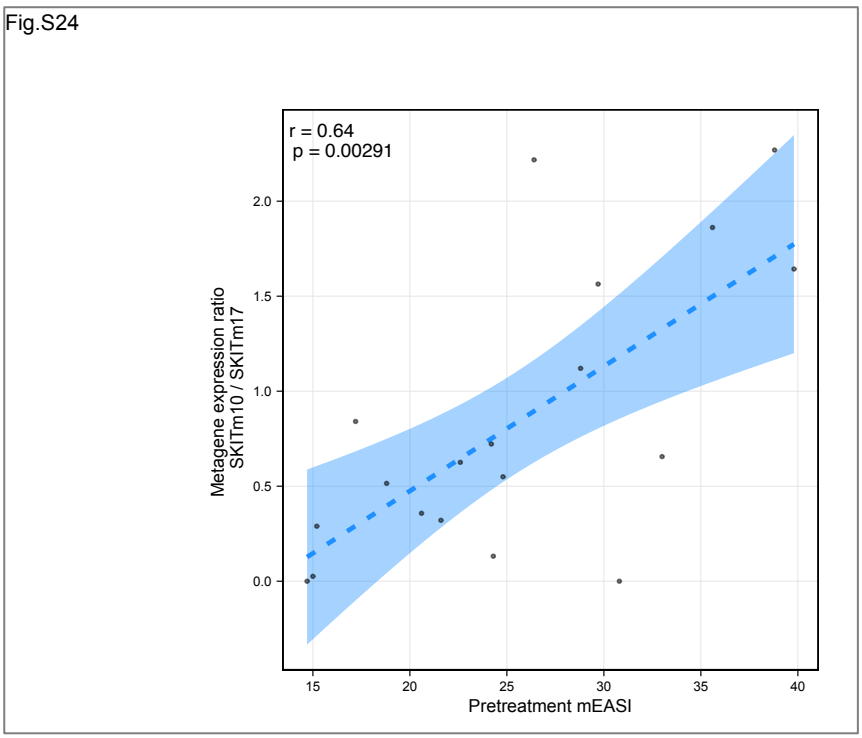
In our study, after 6 months of treatment, 14/22 patients achieved a mEASI75 (75% improvement in the EASI from baseline), and 20/22 patients achieved a mEASI50 (50% improvement in the EASI from baseline), indicating measurable improvements in the majority of patients. Notably, all "Excellent" and "Good" responders reached EASI50, as did 5 out of 7 patients in the "Poor" responder group, which explains why they continued with dupilumab treatment for an extended period. However, while EASI50/75/90 is commonly used in clinical trials, we used sequential absolute severity scores during treatment because recent studies suggest that the absolute value of severity scores (rather than percentage improvement) may provide a more accurate measure of treatment outcomes in real-world settings (Thyssen et al. J Dermatolog Treat. 2023, Seidl et al. Dermatol Ther. 2022). Furthermore, AD often exhibits a fluctuating course, with

periods of exacerbation and remission, even with effective systemic therapies, making it challenging to assess treatment efficacy by percentage improvement at a single time point (Hendricks et al. Am J Clin Dermatol. 2019).

Reviewer #3 raised additional concerns about the terminology of the response groups, suggesting that the names ("Excellent," "Good," "Poor") did not accurately reflect the criteria used for stratification. In response, we modified the terminology in Revision #2 (version -B), again referring to Ferrucci et al., renaming the groups as follows: "Excellent" became "Early," "Good" became "Intermediate," while "Poor" remained unchanged. In Revision #2 (version -B), these changes were reflected in the manuscript (the entire manuscript in Results, Discussion and Methods section, and Figures 3, 5, S14B-S23 of Revision #3, version -C).

As described above, in defining the response groups, we followed a method reported in a previous study. While this stratification approach is not widely used in clinical trials, it was not arbitrary, as it was based on prior research. Although we could not assess the independence of biomarkers from disease severity, our analysis identified specific skin biomarkers that correlate with treatment response based on the stratification. We believe our findings provide valuable insights into the relationship between specific gene signatures and dupilumab treatment responses in AD.

Fig.S24



Reference:
Ferrucci, S. *et al.* Predictive Factors of Early Response to Dupilumab in Patients with Moderate-to-Severe Atopic Dermatitis. *J. Clin. Med. Res.* **12**, (2023). https://www.mdpi.com/2077-0383/12/20/6575

Thyssen, J. P., Bieber, T., Kleyn, C. E., Nosbaum, A., Grond, S., Petto, H., Riedl, E., & Wollenberg, A. (2023). Baricitinib provides rapid and sustained improvements in absolute EASI and SCORAD outcomes in adults with moderate-to-severe atopic dermatitis. *The Journal of Dermatological Treatment*, *34*(1), 2216322. https://doi.org/10.1080/09546634.2023.2216322

Seidl, U., Pinter, A., Wilsmann-Theis, D., Poortinga, S., Morrison, K., Mrowietz, U., & Gerdes, S. (2022). Absolute Psoriasis Area and Severity Index as a valuable marker to determine initial treatment response in psoriasis patients treated with guselkumab in routine clinical care. *Dermatologic Therapy*, *35*(1), e15193. https://onlinelibrary.wiley.com/doi/10.1111/dth.15193

Hendricks, A. J., Lio, P. A., & Shi, V. Y. (2019). Management Recommendations for Dupilumab Partial and Non-durable Responders in Atopic Dermatitis. *American Journal of Clinical Dermatology*, *20*(4), 565–569. https://doi.org/10.1007/s40257-019-00436-8

Comment #3:

The authors have not adequately addressed my comment regarding the IL-17 response and what it means. Again, the authors are using specific modules and not showing expression of IL17A, IL17F, or IL17C as requested and not addressing what it means to have an enriched "Th17 module" if there is no detectable expression of IL17A/F. The authors should note that many of the genes in those IL-17 signatures are induced by multiple other proinflammatory mediators, that is why many of these genes are showing up across multiple modules (i.e., in Figure S9B). The overlap between IL-17A and TNF appears quite significant. Therefore, this needs further validation and the authors should show results for IL17A/F and 17C from their bulk RNA-seq data.

Response:

We apologize for any confusion in the presentation of our findings and thank the

reviewer for raising this important point. Our study of 1 mm punch biopsy identified transcriptional biomarkers derived from keratinocytes and fibroblasts through an unbiased, comprehensive RNA analysis and thus the biomarkers are not dependent on subjective interpretations of the gene profiles. We would like to emphasize that the primary focus and strength of our study lie in the unbiased identification of gene sets through unbiased decomposition analysis, which revealed previously unrecognized co-expression profiles of keratinocytes and fibroblasts in the skin of atopic dermatitis patients. These gene sets represent potential biomarkers of clinical severity and treatment response, independent of predefined biological annotations and known immune pathways. While the association with Type 17 responses is noteworthy, we believe the broader utility and clinical relevance of these metagenes, as well as their potential to capture the complex interplay of skin-resident cells in disease states, are the most significant contributions of this work.

That said, we have validated the annotation our gene sets by multiple methods through the revisions. Since the expression levels of IL17A, IL17F, and IL17C in the RNA-seq data were very low, with mean TPM values of 0.051, 0.076, and 0.975, respectively, these values unreliable for direct interpretation. Therefore, in Revision #1 (version -A), in response to the reviewer#3's request for the validation of the "Type 17 response-associated genes", we performed gene set enrichment analysis to demonstrate that the gene set significantly aligns with previously published studies investigating gene profiles of keratinocytes under IL-17 stimulation (Gupta et al. Sci Immunol. 2021, Jiang et al. JCI Insight. 2024, Nograles et al. Br. J. Dermatol. 2008, Tollenaere et al. Br. J. Dermatol. 2021, Swindell et al. PLoS One. 2012). Prior studies have also classified keratinocytes according to their characteristic gene profiles, termed as "keratinocyte immunophenotypes", by cross-referencing in vitro studies investigating keratinocyte gene profiles under various inflammatory stimuli (Clayton al. Br. J. Dermatol. 2020). The fact that our gene sets aligned with established biological pathways in keratinocytes and fibroblasts under various inflammatory conditions reinforced the validity of our unsupervised analysis. The results of the enrichment analysis were presented in Revision #1 (version -A) (Figures S5A, B and C of Revision #3, version -C).

In Revision #2 (version -B), we further refined the manuscript by refraining from using terms such as "Th17 response" when describing the associated gene set. Instead, we chose to use phrases such as "genes reported to be induced in keratinocytes by type 17 response". Additionally, to emphasize the hypothesis-free nature of our analysis, we renamed the identified gene sets from cell type annotations (e.g., Keratinocyte/Fibroblast) to "SKIn-Tissue derived Metagenes" (SKITm 1-29), highlighting their unbiased origin as

co-expression profiles in AD skin. The modification was incorporated in the entire manuscript in Results, Discussion and Methods section, Figures 1 to 5, S3 to S24 of Revision #3, version -C.

In Revision #3 (version -C), we performed a new set of experiments by conducting qRT-PCR analysis on the remaining available RNA samples of AD skin tissues, focusing on key cytokines of Type 17 and Type 2 inflammation. Of the samples used in the cross-sectional analysis, 33 samples of lesional AD skin were used for analysis due to considerations of sample quality and the availability of remaining material for PCR analysis. The analysis revealed a statistically significant correlation between Type 17 cytokines and the metagenes that we annotated as "Type 17 response-associated". Specifically, IL17A, IL23A and CCL20 expression and SKITm10, and IL17C and SKITm15 expression was significantly correlated. We have also observed that type 2 cytokines (IL13, CCL13, CCL26) were significantly correlated with the SKITm17 metagene that we annotated as "Type 2 response", and that these cytokines did not correlate with SKITm10 and SKITm17. These new RT-PCR results is presented below and have also been incorporated into the manuscript in Revision #3, version -C (Figure 1C, Results section: Page 6, lines 151-158, Methods section: Page 23, lines 690-707 of Revision #3, version -C).

Our findings regarding the biomarkers are derived from our unbiased, comprehensive RNA analysis and our results are not dependent on subjective interpretations of the gene profiles. We believe our observation—that hypothesis-free analysis reproduced gene sets previously identified in hypothesis-driven in vitro studies—supports the quality of our data and strengthens the value of the newly identified biomarkers. Moreover, we have validated the biological annotation through multiple independent, established approaches, and we believe that the validity of our data interpretation has been further strengthened.

Fig.S5

A

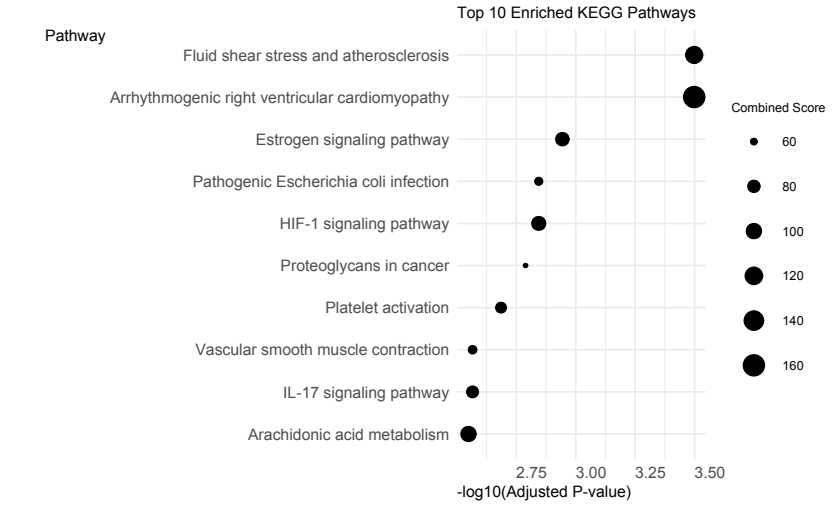


Fig.S5

B

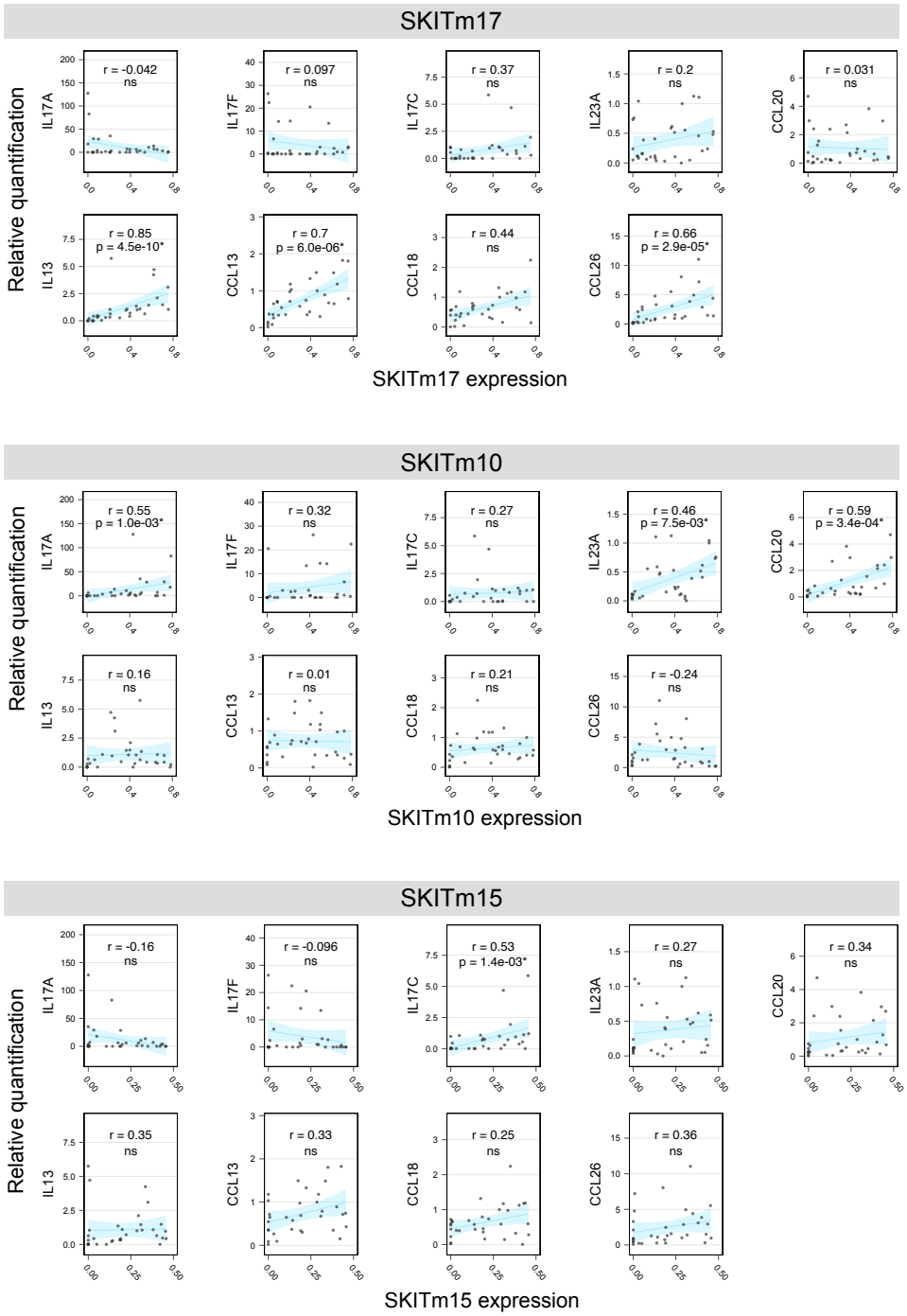
Reference	treatment	Fisher's exact test		Common genes with highly-weighted genes of SKITm10
		p-value	Odds ratio (95%CI)	
Swindell, et al.	IL-17A	3.09e-29	39.64 (23.90 – 63.78)	*SERPINB4 , S100A7A , PI3 , IVL , S100A9 , SERPINB3 , RHCG , KLK6 , CNFN , LCE3D , HBEGF , IL36G , CYP4F22 , LIPG , SLC5A1 , IL36RN , ENDOU , S100A7 , PRDM1 , SPRR2B , TCN1 , ALOX12B , FAM43A , SPRR2D , PLA2G3*
Nograles, et al.	IL-17	1.62e-10	40.26 (15.81 – 90.62)	*SERPINB4 , S100A8 , PI3 , S100A9 , RHCG , S100A7 , CRABP2 , SPRR2D*
Jiang, et al.	IL-17A	6.10e-16	38.41 (18.91 – 72.92)	*S100A7A , S100A8 , PI3 , S100A9 , SPRR2A , RHCG , LCE3A , HEPHL1 , PLA2G4D , IL36G , S100A7 , SPRR2B , SPRR2D*
	IL-17C	5.11e-05	53.78 (9.27 – 218.85)	*S100A7A , SPRR2A , IL36G*
	TNFa	1.70e-20	13.11 (8.30 – 20.06)	*S100A8 , PI3 , S100A9 , SPRR2A , RHCG , CNFN , LCE3A , LCE3D , OASL , HBEGF , LCE3E , HEPHL1 , IL36G , MXD1 , TYMP , TRPV3 , UPP1 , SLC5A1 , PRDM1 , SLC28A3 , SERPINB2 , SPRR2B , PPIF , CDSN , SPRR2D , PLA2G3 , RNF222 , CRCT1*
Gupta, et al.	IL17-A	7.19e-35	20.60 (13.87 – 30.10)	*SERPINB4 , S100A7A , S100A8 , PI3 , GJB2 , SPRR1A , S100A9 , SERPINB3 , SPRR2A , RHCG , SPRR1B , GJB6 , CNFN , TGM1 , LCE3D , OASL , HBEGF , HEPHL1 , PLA2G4D , IL36G , SULT2B1 , TYMP , SLPI , TRPV3 , SLC5A1 , IL36RN , S100A7 , PRDM1 , PRSS22 , PLA2G4E , SLC28A3 , SPRR2B , CALML3 , CRABP2 , ALOX12B , SPRR2D , MX1 , MPZL2 , RNF222 , KLK9*
	TNFa	9.80e-21	7.84 (5.34 – 11.31)	*SERPINB4 , S100A8 , PI3 , S100A9 , SPRR2A , RHCG , DSC2 , LCE3D , HBEGF , IGFL1 , IL1RN , HEPHL1 , ODC1 , IL36G , SULT2B1 , TYMP , FOSL1 , PHLDA2 , SLPI , TRPV3 , UPP1 , SLC5A1 , IL36RN , ANXA1 , PRDM1 , DRAP1 , PRSS22 , CSTB , PLA2G4E , SLC28A3 , SERPINB2 , SPRR2B , GOLGA7B , KLK10 , PNP , TUBB6 , GLTP , PPIF , SPRR2D , CDKN1A , CRCT1*
Tollenaere, et al.	IL-17A	3.05e-26	22.31 (14.02 – 34.47)	*SERPINB4 , S100A7A , SERPINB3 , SPRR2A , RHCG , LCE3A , DDIT4 , OASL , HBEGF , LCE3E , HEPHL1 , PLA2G4D , IL36G , FOSL1 , TRPV3 , UPP1 , HK2 , IL36RN , PRDM1 , PRSS22 , SPRR2B , TCN1 , PRSS27 , PPIF , FAM43A , SPRR2D , KLK9 , FCHSD1*
	IL-17C	1.00e+00	0.00 (0.00 – 851.00)	-
	IL-17F	5.24e-22	27.65 (16.12 – 45.54)	*SERPINB4 , S100A7A , GJB2 , SERPINB3 , SPRR2A , RHCG , LCE3A , HEPHL1 , PLA2G4D , IL36G , TRPV3 , UPP1 , HK2 , IL36RN , PRSS22 , SPRR2B , TCN1 , PRSS27 , SPRR2D , KLK9 , FCHSD1*

C

Reference	treatment	Fisher's exact test		Common genes with marker genes of SKITm10
		p-value	Odds ratio (95%CI)	
Swindell, et al.	IL-17A	3.59e-06	47.64 (11.44 - 149.83)	*SERPINB4, PI3, IVL, LIPG*
Nograles, et al.	IL-17	4.37e-04	75.24 (8.18 - 332.56)	*SERPINB4, PI3*
Jiang, et al.	IL-17A	1.31e-03	42.48 (4.69 - 183.02)	*LCE3A, PI3*
	IL-17C	1	0 (0 - 514.02)	-
	TNFa	2.68e-04	15.31 (3.71 - 47.72)	*UPP1, MXD1, LCE3A, PI3*
Gupta, et al.	IL17-A	4.19e-03	10.53 (1.97 - 36.58)	*SPRR1A, PI3, SERPINB4*
	TNFa	9.92e-06	12.53 (4.22 - 33.87)	*IGFL1, UPP1, IL1RN, PI3, DSC2, PHLDA2, SERPINB4*
Tollenaere, et al.	IL-17A	9.96e-04	17.63 (3.29 - 61.39)	*UPP1, LCE3A, SERPINB4*
	IL-17C	1	0 (0 - 7358.56)	-
	IL-17F	2.21e-04	29.92 (5.56 - 104.92)	*UPP1, LCE3A, SERPINB4*

Fig.1

C



Reference:

Gupta, R. K. et al. TWEAK functions with TNF and IL-17 on keratinocytes and is a potential target for psoriasis therapy. Sci Immunol 6, eabi8823 (2021). https://www.science.org/doi/10.1126/sciimmunol.abi8823

Jiang, Y. et al. Suppression of TCF4 promotes a ZC3H12A-mediated self-sustaining inflammatory feedback cycle involving IL-17RA/IL-17RE epidermal signaling. JCI Insight 9, (2024).
https://insight.jci.org/articles/view/172764

Nograles, K. E. et al. Th17 cytokines interleukin (IL)-17 and IL-22 modulate distinct inflammatory and keratinocyte-response pathways. Br. J. Dermatol. 159, 1092–1102 (2008).
https://academic.oup.com/bjd/article/159/5/1092/6641821

Tollenaere, M. A. X. et al. Signalling of multiple interleukin (IL)-17 family cytokines via IL-17 receptor A drives psoriasis-related inflammatory pathways. Br. J. Dermatol. 185, 585–594 (2021).
https://doi.org/10.1111/bjd.20090

Swindell, W. R. et al. Heterogeneity of inflammatory and cytokine networks in chronic plaque psoriasis. PLoS One 7, e34594 (2012).
https://dx.plos.org/10.1371/journal.pone.0034594

Clayton, K. et al. Machine learning applied to atopic dermatitis transcriptome reveals distinct therapy-dependent modification of the keratinocyte immunophenotype. Br. J. Dermatol. 1–10 (2020) doi:10.1111/bjd.19431.
https://academic.oup.com/bjd/article/184/5/913/6702287

Comment #4:
The authors state that they provide RIN numbers, but I cannot find these data. Selecting the RIN cutoff at RIN<2 is concerning. Most groups that do this type of work draw the line at RIN>6. How many of the samples were above that threshold (RIN=6) and how many below?
Remarks to the Editor I still have major concerns regarding this manuscript. While the value of this as a resource to the field would be high I have concerns regarding the data quality (I'm unable to find information on the RIN numbers) - that the removed samples with RIN<2, for me, indicates that the data quality is

poor overall - must studies I'm familiar with would have a cutoff of RIN<6.

Response:

Thank you for raising this important point. We acknowledge the concern regarding the inclusion of RNA samples with RIN below 6. The RNA integrity number (RIN) score indicates the integrity of RNA, and the score is known to decrease in skin samples compared to other organs. This is because skin samples are contaminated with nucleases due to environmental exposure. In addition, skin samples, particularly from 1 mm punch biopsies, are substantially composed of keratinocytes, which express high levels of endogenous ribonucleases (Reimann et al. Sci. Rep. 2019, LoCoco et al. Pain Rep. 2020). As a result, obtaining clinical skin samples with consistently high RIN values is challenging. Indeed, recently published studies using human samples such as skin and brain tissue have observed significant pathological molecular signatures using skin samples with RIN scores under 2 (Sølberg et al. J. Dermatol. Sci. 2021, Srinivasan et al. Cell Rep. 2020). In addition, our previous work published in *Nature Communications* (Sekita et al. Nat. Commun. 2023) included 219 samples that overlap with the current study, of which 50 samples had RIN values below 6. Furthermore, there is the growing claim that RIN values may not adequately reflect sample mRNA integrity, as RIN algorithm heavily weighs rRNA integrity, which only weakly associate with sample transcript quality (LoCoco et al. Pain Rep. 2020).

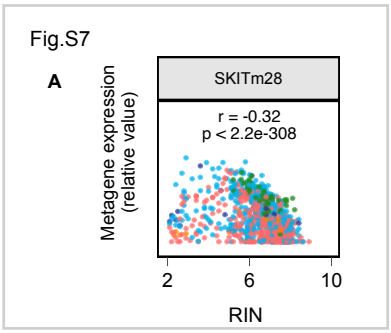
For these reasons, in our study, we set the RIN score cutoff at 2, which is lower compared to typical cutoff for RNA-seq of other tissues (such as PBMC), to avoid losing samples in our large-scale human study. However, we also recognized the potential for RIN scores to introduce bias. To address this, in the original manuscript, we conducted a correlation analysis between all gene sets and RIN score (Figure S7, the relevant portion of Figure S7 presented below), which revealed a significant correlation between SKITm28 and RIN score. To mitigate biases associated with low RIN scores, we removed gene sets linked to sample quality, including SKITm28. This approach ensured that the molecular data reflected biologically relevant signals related to AD pathogenesis, as demonstrated by UMAP analysis in the original manuscript (Figures S8, Page 7, lines 159–167; Page 27, lines 805–812 of Revision #3, version -C). Specifically, in Figure S8A, UMAP analysis was performed using all 29 metagenes. In Figure S8B, 11 sample-quality-related metagenes, including SKITm28, were excluded, and UMAP analysis was repeated using the remaining 18 metagenes. This exclusion reduced clustering biases related to sample quality, making biological patterns more discernible in Figure S8B.

In response to the reviewer's request, in addition to Figure S7 presented in the original manuscript, in Revision #1 (version -A), we presented the RIN score as

Supplementary data (Supplementary Table 19 of Revision #3, version -C).

Furthermore, in Revision #3 (version -C), as presented below, we performed an additional analysis examining the relationship between metagene expression and disease severity scores in lesional skin samples stratified by RIN scores (<6 and ≥6). The results demonstrate a consistent trend across both RIN groups, with similar correlational relationships to disease severity. The accompanying heatmap displays Pearson's correlation coefficients, where blue indicates significant negative correlations and red indicates significant positive correlations ($p < 0.01$). Notably, the correlations involving the four key metagenes (SKITm2, SKITm16, SKITm17, SKITm10) in the samples with low RIN score (<6) exhibited the same trend as observed in the regression analysis reported in the original manuscript, which included all samples with RIN scores between 2 and 10. The results have been incorporated in Revision #3, version -C (Figure S1D, Methods section: Page 22, lines 652 to 661 of Revision #3, version -C).

For the reasons outlined above, we believe that restricting samples to those with higher RIN values does not necessarily improve the quality of our data and, in fact, may diminish the overall impact of our large-scale human sample study. We hope you understand this context and allow us to continue using the current RIN value cutoff.



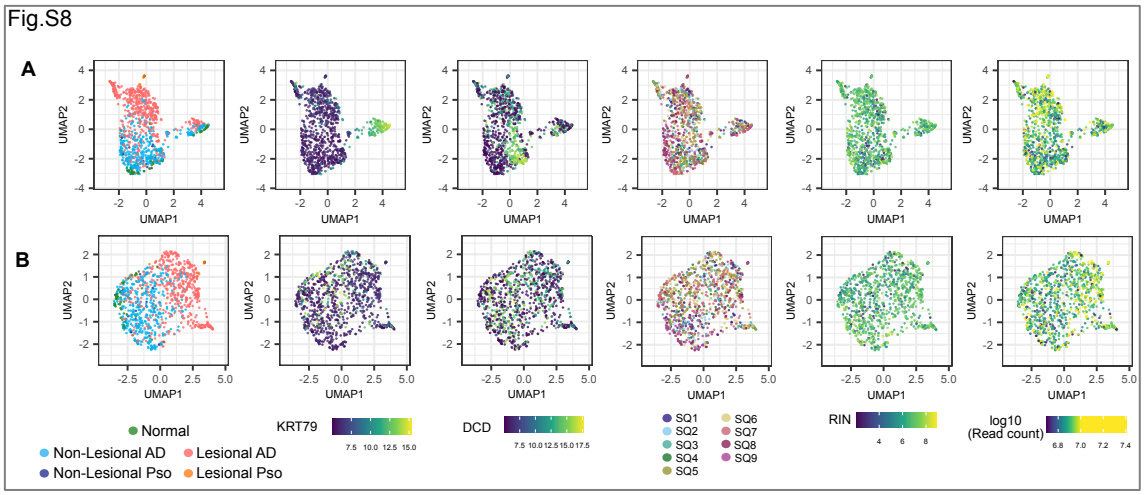
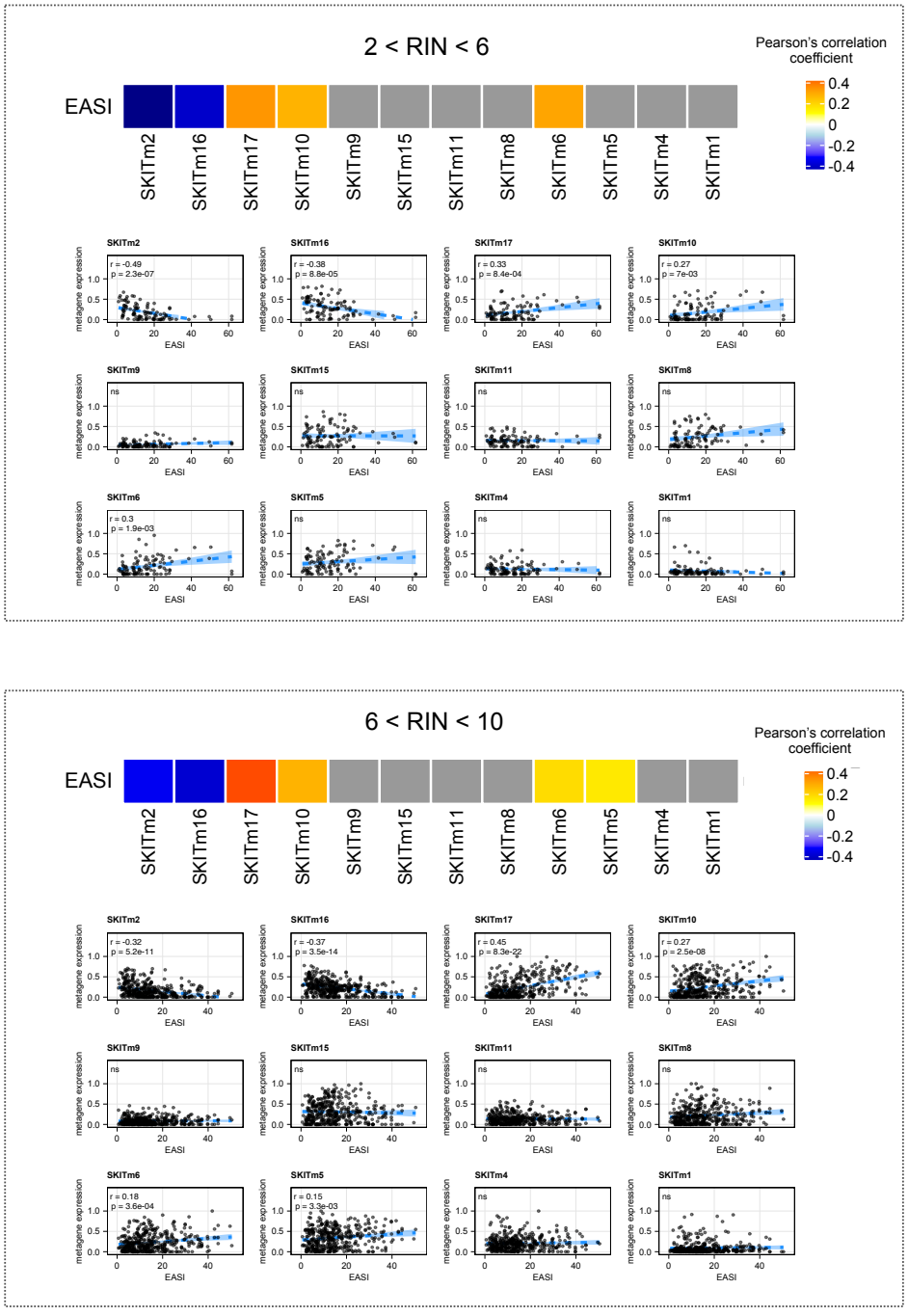


Fig.S1

D



Reference:

Reimann, E., Abram, K., Kõks, S., Kingo, K. & Fazeli, A. Identification of an optimal method for extracting RNA from human skin biopsy, using domestic pig as a model system. Sci. Rep. 9, 20111 (2019).

https://www.nature.com/articles/s41598-019-56579-5

Sølberg, J. et al. The stratum corneum transcriptome in atopic dermatitis can be assessed by tape stripping. J. Dermatol. Sci. 101, 14–21 (2021). https://www.sciencedirect.com/science/article/abs/pii/S0923181120303297

Srinivasan, K. et al. Alzheimer's patient microglia exhibit enhanced aging and unique transcriptional activation. Cell Rep. 31, 107843 (2020). https://www.cell.com/cell-reports/fulltext/S2211-1247(20)30824-X?sf124175287=1

Sekita, A. et al. Multifaceted analysis of cross-tissue transcriptomes reveals phenotype-endotype associations in atopic dermatitis. Nat. Commun. **14**, 6133 (2023). https://www.nature.com/articles/s41467-023-41857-8

LoCoco, P. M. et al. Reliable approaches to extract high-integrity RNA from skin and other pertinent tissues used in pain research. Pain Rep. 5, e818 (2020). https://journals.lww.com/10.1097/PR9.0000000000000818

Comment #5:

I have further, and major, concerns regarding the treatment response definition, which seems arbitrary (the reference they provide had a cutoff of EASI <7, but here the authors use EASI <6.3 - not a big difference, but why cite the paper if they are not using the same cutoff), and also the disease severity at baseline is different with the "Excellent responders" having significantly milder disease at baseline than the "moderate" or "poor" responders - this does not correspond to how clinical definition of response would be measured in any clinical trial or study (i.e change in EASI) - this is a fatal flaw, will lead to erroneous conclusions, and would necessitate a major revrite.

Response:

This comment is the same as comment #2. While the improvement rates of disease severity scores (such as EASI50/75/90) are commonly used in clinical trials, recent studies suggest that sequential absolute severity scores provide a more accurate measure of treatment outcomes in real-world settings. Therefore, we categorized patients based on their absolute severity scores over six months, following an approach from a previous study (Ferrucci et al. J Clin Med Res. 2023). The cutoff we used for mild disease was 6.3 of modified EASI (not EASI score). Further detailed explanation has been presented in

the response of Comment #2.

Comment #6:

Another concern is the lack of specificity of their "Th17 module" to IL-17A/F signaling...they should have the data and expression values of IL17A/IL17F in their dataset but do not show it. I'm concerned that this module is meaningless unless the authors validate it or provide data to link it back to IL-17 signaling/cytokines.

Response:

This comment is the same as comment #3. As described in detailed responses to Comment #3, the expression levels of IL17A, IL17F, and IL17C in the RNA-seq data were very low making these values unreliable for direct interpretation. As such, in Revision #3 (version -C), we performed an additional experiment by qRT-PCR to evaluate the expression levels of Type 17 cytokines including IL17A, IL17F, and IL17C in the samples and demonstrated that the cytokines evaluated by qRT-PCR correlate with the expression levels of metagenes annotated as “Type 17 response”. The additional analysis and experiments conducted, as per the reviewer’s suggestion, further strengthen the validity of our data interpretation.

Comment #7:

Lastly, there is almost nothing novel shown here - therefore, the findings are, at best, incremental in nature, with the value of this manuscript being the resource of the overall data and data availability.

Response:

Our study is the first to unravel the clinical significance of complex and heterogenous molecular pathology of AD by a large-scale, unbiased comprehensive RNA analysis of human skin tissue samples. Our hypothesis-free analysis was able to replicate findings reported in previous hypothesis-driven studies, underscoring the relevance and accuracy of our approach. We uncovered hidden co-expressed gene profiles of keratinocytes and fibroblasts in AD skin through a hypothesis-free, data-driven approach. Furthermore, critically, this research uniquely identifies gene signatures associated with both clinical phenotype and resistance to biological therapies. We believe that our findings provide valuable biomarkers derived from keratinocytes and fibroblasts for future clinicopathological studies in AD.