nature portfolio

Corresponding author(s): Jianxin Wu, Qing Huang

Last updated by author(s): Mar 20, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection in this study.
Data analysis	Established open-source software, plugins, or pipelines were used for data analysis, with detailed information provided in the manuscript. Complete references and/or web links for access are also provided below. No new code was developed for data analysis in this study. GraphPad, Version 7.0 or 8.0, RRID:SCR_002798, https://www.graphpad.com/ FIJI (ImageJ), 2.0.0, RRID: SCR_002285, https://fiji.sc/ ImageJ (OpenComet Plugin) SRC V1, RRID:SCR_001935, https://fiji.sc/ Python 3.6, https://www.python.org/downloads/R version 4.0.3, RRID:SCR_001905, https://www.r-project.org R version 3.2.0, RRID:SCR_001905, https://www.r-project.org QIIME2 (2022.11), RRID: SCR_021258, https://qiime2.org/ DADA2, RRID: SCR_023519, https://benjjneb.github.io/dada2/ Greengenes 13.8, RRID:SCR_002830, http://greengenes.lbl.gov SWISS-MODEL Repository, RRID: SCR_013032, https://swissmodel.expasy.org/

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The sequence data analyzed during the current study are available from NCBI (PRJNA 1193278, http://www.ncbi.nlm.nih.gov/bioproject/1193278). All data are available in the main text or the Supplementary Information/Source Data file. Source data is available for Figs. 1e, g, h, 2a-d, f, g, 3b-i, 4b-e, 5b-d, g, h, i and 6b, e and Supplementary Figs. 2, 3, 4, 5, 6 and 7 in the associated source data file.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	No information about sex and gender is disclosed. Subject selection based on skin condition, independent of gender.
Reporting on race, ethnicity, or other socially relevant groupings	No socially constructed or socially relevant categorical variables were included in the manuscript.
Population characteristics	All samples were collected from students of skin health at China Pharmaceutical University.
Recruitment	Healthy-skinned students living in Nanjing, China, for at least two years, with criteria for skin health assessed by the subject's self-assessment or by a dermatologist at the Affiliated Hospital of Nanjing Medical University.
Ethics oversight	Ethics Committee of The Third Affiliated Hospital of Nanjing Medical University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes have been summarized in Supplementary Table 3. Sample sizes were determined from preliminary data indicating statistically significant differences for each assay. Each experiment included one subject per group, with three or four biological samples collected per subject.
Data exclusions	No data were excluded from the analyses.
Replication	For each experiment the number of biologically independent samples is reported in the figure legend.
Randomization	Randomization and covariates were not relevant to our study design as we investigated single factors within each study.
Blinding	For analysis the subjects were blinded for the observers. This involved masking original sample identification and assigning coded IDs before data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HaCaT cell line was purchased from Zhejiang Meisen Cells Technology Co., Ltd. An immortalized human sebocyte cell line was established by transfecting secondary sebocytes with the pGMLV-SV40T-PURO lentiviral vector. The secondary sebocytes were derived from the normal skin of a 55-year-old male. Human skin samples were obtained from surgeries following the protocol approved by the Ethics Committee of the Third Affiliated Hospital of Nanjing Medical University.
Authentication	Both the HaCaT cell line and the sebocyte cell line have undergone STR authentication.
Mycoplasma contamination	Media from cell lines were randomly tested for Mycoplasma in our laboratory, and there has been no evidence for contamination by Mycoplasma.
Commonly misidentified lines (See ICLAC register)	N/A

Plants

Seed stocks	No plants were used.
Novel plant genotypes	No plants were used.
Authentication	No plants were used.