

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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

The V3-V4 regions of the 16S ribosomal RNA (rRNA) gene samples obtained from participants' nasal swabs were amplified and sequenced in a single run on an Illumina HiSeq 2500 sequencing platform.
q-RT-PCR data were obtained on an ABI 7500 thermocycler (Applied Biosystems).
Cytokine concentrations in patient serum samples were determined by flow cytometry using BD FACSCanto II.

Data analysis

Analysis of 16S rRNA sequence data was performed using Quantitative Insights Into Microbial Ecology (QIIME).
Taxonomic classification from the phylum to genus level was performed based on the SILVA database (<https://www.arb-silva.de/documentation/release-128/>).
Fast Length Adjustment of Short reads (FLASH) was used to merge PE reads from sequencing.
Fold changes in qRT-PCR experiments were analyzed in Excel (Microsoft Office 2021), and statistical analysis was performed using Prism 8 software.
Cytokine concentrations were analyzed by FCAP Array v3.0.1 software and BD FACSDiva v8.0.1 software.
Blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to align the sequences of OTUs.
Epithelial thickness was measured by cellSens imaging software (Olympus).
Short tandem repeats were separated on ABI 3730XL Genetic Analyzer. The signals were then analyzed by the software GeneMapper ID v3.2.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

Raw microbiome sequencing data (related to data shown in Fig. 1, Table 1, and Extended Data Figs. 1-6) and genome sequencing data have been deposited in NCBI's Sequencing Read Archive (SRA) database under Bioproject numbers PRJNA796497 and PRJNA874865, respectively. All other data are presented in this manuscript. Source data files (see Supporting Material) contain results for all figures with quantitative data. Bacterial strains are available from Min Li (rjlimin@shsmu.edu.cn).

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](https://www.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	No statistical methods were used to pre-determine sample sizes but our sample sizes were chosen based on those reported in previous publications. For the human microbiome study, the number of analyzed individuals as compared to previous studies analyzing nasal microbiomes (Lal, D. et al. Int Forum Allergy Rhinol 7, 561-569, doi:10.1002/alr.21934 (2017), Gan, W. et al. Eur Arch Otorhinolaryngol 278, 711-718, doi:10.1007/s00405-020-06311-1 (2021), Choi, C. H. et al. Am J Rhinol Allergy 28, 281-286, doi:10.2500/ajra.2014.28.4050 (2014), Hyun, D. W. et al. Infect Immun 86, doi:10.1128/IAI.00934-17 (2018) was increased in purpose to overcome insignificant results in some of those previous studies. For animal studies, our sample sizes were chosen based on those reported data in previous publications and previous experience in similar experimental setups. (Schmalzl, A. et al. Nat Commun, 2022. 13(1): p. 5730; Kienesberger, S. et al. Nat Microbiol, 2022. 7(11): p. 1834-1848. van Heeckeren, A.M. et al. Am J Respir Crit Care Med, 2000. 161(1): p. 271-9. Liu, Q. et al. Cell Host Microbe, 2020. 27(1): p. 68-78 e5.) For in vitro experiments, sample sizes were chosen that are similar to those used in previous similar studies: Experiments in Fig. 2: van Heeckeren, A.M. et al. Am J Respir Crit Care Med, 2000. 161(1): p. 271-9.; Fig. 3: Liu, Q. et al. Cell Host Microbe, 2020. 27(1): p. 68-78 e5; Experiments in Fig. 4: Schmalzl, A. et al. Nat Commun, 2022. 13(1): p. 5730; Kienesberger, S. et al. Nat Microbiol, 2022. 7(11): p. 1834-1848.
Data exclusions	Data were only excluded in sequence processing: samples with low-quality (Q value < 25) base ratio or lower than 97% of reads.
Replication	Fig.2 was repeated in its entirety with cytokine expression and epithelial thickness results showing similar results (Extended Data Fig. 9) The experiments shown in other figures were not repeated, as they were all performed with a number of replicates deemed sufficient to achieve reliable results.
Randomization	Mice used in the experiments were litter mates, sex, and age-matched, and randomized into control and experimental groups. For some experiments, <i>S. salivarius</i> and <i>S. epidermidis</i> nasal isolates obtained from AR patients were randomly selected. Randomization is not applicable to other (in vitro) experiments.
Blinding	Blinding was not performed. The clinical study was an observational study without intervention, for which blinding is not applicable. For animal experiments, groups were given different treatments and analysis was performed by the same investigators so blinding was not possible.

Behavioural & social sciences study design

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

Study description	Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).
Research sample	State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.

Sampling strategy	Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.
Data collection	Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.
Timing	Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Non-participation	State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.
Randomization	If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

Ecological, evolutionary & environmental sciences study design

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

Study description	Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.
Research sample	Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i> , all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.
Sampling strategy	Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.
Data collection	Describe the data collection procedure, including who recorded the data and how.
Timing and spatial scale	Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Reproducibility	Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.
Randomization	Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.
Blinding	Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.
Did the study involve field work?	☐ Yes ☒ No

Field work, collection and transport

Field conditions	Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).
Location	State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).
Access and import/export	Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).
Disturbance	Describe any disturbance caused by the study and how it was minimized.

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
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

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

Antibodies

Antibodies used

Mouse anti-MUC5AC antibody (Thermo Catalog #MA512178, Clone 45M1) was used to determine expression of MUC5AC at a dilution of 1:1000.
Anti-mouse HRP-conjugated antibody (Yeasen, Catalog # 33201ES60) at a dilution of 1:5000.
Anti-beta actin antibody (CST, Catalog # 4970, Clone 13E5) was used to determine expression of beta actin at a dilution of 1:1000.
MUC5AC Rabbit antibody (CST, Catalog #36623, Clone E9V10) was used to determine expression of MUC5AC by IHC at a dilution of 1:100.
(HRP)-coupled IgG secondary antibody(Abcam, Catalog #ab205718) at a dilution of 1:2000.

Validation

The validation information can be obtained on the vendor websites with the following links:
mouse anti-MUC5AC antibody (Thermo Catalog #MA512178, Clone 45M1)
<https://www.thermofisher.cn/cn/zh/antibody/product/MUC5AC-Antibody-clone-45M1-Monoclonal/MA5-12178>
Anti-mouse HRP-conjugated antibody (Yeasen, Catalog # 33201ES60)
<http://yeasen.com/products/detail/407>
Anti-beta actin antibody (CST, Catalog # 4970, Clone 13E5)
<https://www.cellsignal.com/products/primary-antibodies/b-actin-13e5-rabbit-mab/4970?site-search-type=Products&N=4294956287&Ntt=actin&fromPage=plp>
MUC5AC Rabbit antibody (CST, Catalog #36623, Clone E9V10)
<https://www.cellsignal.com/products/primary-antibodies/muc5ac-e9v10-xp-rabbit-mab/36623?site-search-type=Products&N=4294956287&Ntt=muc5ac&fromPage=plp>
(HRP)-coupled IgG secondary antibody. (Abcam, Catalog #ab205718)
<https://www.abcam.com/goat-rabbit-igg-hl-hrp-ab205718.html>

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

A549 cells and 2650 cells were obtained from Fuheng Biology Co., Ltd

Authentication

the A549 cell line was authenticated by short tandem repeat (STR) profiling. Briefly, DNA of A549 cell lines was extracted. The twenty STRs including the amelogenin locus were amplified by six multiplex PCRs and separated on an ABI 3730XL Genetic Analyzer. The signals were then analyzed by GeneMapper software. The DNA typing of the analyzed strain matched the A549 cell line in a cell line search performed against the DSMZ database . No multiple alleles were found in this cell line. Authentication testing was last performed Jan 17,2022.

Mycoplasma contamination

No mycoplasma contamination

Commonly misidentified lines (See [ICLAC](#) register)

None of the cell lines used in this study are commonly misidentified as per ICLAC.

Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

C57BL/6J mice were purchased from GemPharmatech and bred in-house under SPF conditions. 6-week-old female mice with littermate controls were used for all experiments, except for the experiment using MUC5AC ^{-/-} mice. For the experiment with MUC5AC ^{-/-} mice and control mice, age- and gender-matched numbers of mice were used and wild-type or MUC5AC ^{-/-} mice were randomly assigned to experimental groups. Animals were housed at 19-26 °C at a humidity of 40-70% and with a light/dark cycle of 12h/12h.

Wild animals

Study did not involve wild animals.

Field-collected samples

Study did not involve field-collected samples.

Ethics oversight

All animal procedures complied with the guidelines of the Shanghai Medical Experimental Animal Care. Animal protocols were approved by the Institutional Animal Care and Use Committee of Ren Ji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China (approval number RA-2021-58). Protocols of animal studies performed in the U.S. were approved by the NIAID DIR ACUC (study protocol LB3E).

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

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

160 participants including 55 patients with AR and 105 healthy controls were enrolled in Shanghai China, between November 1 and December 30, 2018. All were between 20 and 50 years of age with a median age of 32 years. In the healthy control group, 62% were male, in the AR group, 49% were males. There was no statistically significant difference in gender between the two groups.

Recruitment

Participants were recruited through notices in the medical center of the hospital. All samples were collected between November 1 and December 30, 2018, to minimize the influence of factors associated with seasonal AR. During this period, all eligible AR patients and healthy volunteers were included in the study without artificial exclusion. There were no self-selection or other biases. Diagnostic criteria for AR were: (i) typical rhinitis symptoms including rhinorrhea, nasal obstruction and itching, and sneezing, and (ii) a positive skin prick test (SPT, wheal diameter equal to or larger than 3 mm larger than that of negative control after 15 min) to at least one allergen, or instead of a positive SPT, serologic evidence (more than 0.35 kU/l) of sensitivity to at least one of 12 aeroallergens as determined by presence of specific IgE using ImmunoCap IgE assays (Thermo-Scientific). Healthy controls did not self-report or were diagnosed by a physician with past or current rhinitis. Further exclusion criteria included receiving immunotherapy, chronic rhinosinusitis, other allergic disease (allergic skin diseases and allergic asthma), or respiratory infection within 6 weeks, or antibiotic use within one month before enrollment. There was no compensation for participants.

Ethics oversight

The human clinical study was approved by the ethics committee of Ren Ji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China (approval number KY2019-003). All subjects provided written informed consent.

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

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.

Study protocol

Note where the full trial protocol can be accessed OR if not available, explain why.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

ChIP-seq

Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session

(e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Venous blood was centrifuged to separate 0.5 ml serum.

Instrument

BD FACSCanto II

Software

FACP Array v3.0.1 software, BD FACSDiva software

Cell population abundance

No cell experiments

Gating strategy

The assay uses antibody-coated beads to capture analytes efficiently. Each bead has a unique fluorescence intensity that resolves as a unique population by Allophycocyanin (APC) and APC-CY7 on a flow cytometer. Analyte bound to a bead is detected by a second antibody with a phycoerythrin (PE) label. The PE signal is proportional to the amount of bound analyte. According to the kit instructions, we first adjusted the voltage of forward scatter (FSC) and side scatter (SSC) until the singlet bead populations fitted well. Singlet bead populations were gated by SSC-A versus FSC-A to exclude cell debris and platelets. A total of 3600 beads with around 300 in each cluster were analyzed. Different antibody-labeled bead populations were identified by APC and APC-CY7 according to their fluorescence intensities. Some serially diluted standards with known concentration were also analyzed to generate a standard curve for each analyte. After acquiring samples on a flow cytometer, the FCS files were analyzed by FACP Array™ 3.0 software which identifies each bead population, generates a standard curve, and calculates the concentration of unknown samples.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	<i>Indicate task or resting state; event-related or block design.</i>
Design specifications	<i>Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.</i>
Behavioral performance measures	<i>State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).</i>

Acquisition

Imaging type(s)	<i>Specify: functional, structural, diffusion, perfusion.</i>
Field strength	<i>Specify in Tesla</i>
Sequence & imaging parameters	<i>Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.</i>
Area of acquisition	<i>State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.</i>
Diffusion MRI	☐ Used ☐ Not used

Preprocessing

Preprocessing software	<i>Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).</i>
Normalization	<i>If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.</i>
Normalization template	<i>Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.</i>
Noise and artifact removal	<i>Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).</i>
Volume censoring	<i>Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.</i>

Statistical modeling & inference

Model type and settings	<i>Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).</i>
Effect(s) tested	<i>Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.</i>
Specify type of analysis:	☐ Whole brain ☐ ROI-based ☐ Both
Statistic type for inference (See Eklund et al. 2016)	<i>Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.</i>
Correction	<i>Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).</i>

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis
Functional and/or effective connectivity	<i>Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).</i>
Graph analysis	<i>Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).</i>

