nature portfolio

Corresponding author(s): Russell

Last updated by author(s): May 22, 2024

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.
Data analysis	CellRanger (v. 6.0, 10X Genomics): Used for processing mRNA libraries against the mouse mm10 genome (GRCm38). CITE-Seq-Count (v. 1.4.3): Utilized for processing ADT and HTO libraries. Seurat (v. 4.3): Conducted downstream data analysis, including cell filtering, demultiplexing, integration, clustering. Harmony: Used for integrating multiple datasets based on batch, infection status, timepoint, and vaccination status. Monocle (v. 3.0): Performed trajectory and pseudotime analysis. Pegasus: Used for generating a force layout embedding (FLE) to visualize cellular states. Tempora: Employed for time-dependent pathway analysis. scWGCNA: Used for weighted gene co-expression network analysis (WGCNA) on single-cell RNA-seq data. Milo: Applied for differential abundance testing of macrophage populations. scCODA: Utilized to investigate changes in cell composition across different timepoints during Mtb infection. G:Profiler: Conducted pathway enrichment analysis. ArchR: Performed integrative single-cell chromatin accessibility analysis. Custom R Scripts: Developed for additional data processing, statistical analysis, and visualization.

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 datasets supporting the conclusion of this study are available in the Gene Expression Omnibus (GEO) under accession numbers: GSE245950 (scRNA-seq) and GSE245836 (scATAC-seq). The scRNA-seq datasets for the 3-week timepoint and naïve lung were previously published and are available under accession number: GSE167232.

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	The study didn't include human participants.
Reporting on race, ethnicity, or other socially relevant groupings	The study didn't include human participants.
Population characteristics	The study didn't include human participants.
Recruitment	The study didn't include human participants.
Ethics oversight	The study didn't include human participants.

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 were determined based on previous studies and established protocols to ensure statistical power and reproducibility. For each experimental condition, groups of 3-5 mice were used, which is consistent with prior research in similar immunological studies involving murine models of tuberculosis infection and vaccination. This sample size was chosen to balance the need for sufficient statistical power to detect significant differences while adhering to ethical guidelines to minimize the use of animals.
Data exclusions	No data were excluded from the study.
Replication	To ensure the reproducibility of our experimental findings, we implemented several key measures: Biological Replicates: Experiments were conducted with multiple biological replicates. Specifically, each experimental condition included groups of 3-5 mice, and indipendent experiments were perfomred (where applicable) to confirm the consistency of the results. Standardized Protocols: All experimental procedures were standardized and strictly followed to minimize variability. This included consistent use of reagents, equipment, and standardized protocols for tissue dissociation, cell sorting, and library preparation. Data Quality Control: Quality control steps were applied to sequencing data, including filtering out low-quality cells and ensuring consistent data processing pipelines. CellRanger, CITE-Seq-Count, and Seurat software were used for data processing with consistent parameters across experiments. Validation of Findings: Key findings were validated using multiple complementary approaches. For example, cell composition differences observed in single-cell RNA-seq were cross-validated with flow cytometry where applicable. Detailed Documentation: Detailed records of all experimental procedures, including sample preparation, sequencing protocols, and data analysis steps, were maintained.

Public Data Sharing: Data and code used in the analysis are made available in public repositories, allowing other researchers to reproduce the analysis and validate the findings independently.

These measures collectively ensured that the experimental findings reported in this study are reproducible and reliable.

Randomization: Not applicable to this study: We used SPF mice with identical genetic background.

Blinding: Not applicable to this study: We used SPF mice with identical genetic background.

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

Antibodies

Antibodies used: For our scRNA-seq timepoint experiments, we used a range of TotalSeq (BioLegend) murine antibodies in our antibody cocktail mix, each at a concentration of 0.5 µg/sample. These antibodies included SiglecF (custom-made, clone S17007L), CD64 (cat. # 139325), Ly6G (cat. # 127655), CD11c (cat. # 117355), CD14 (cat. # 123333), Ly6G-Ly6C (cat. # 108459), CD63 (cat. # 143915), F4/80 (cat. # 123153), CD38 (cat. # 102733), TLR4 (cat. # 117614), CD11b (cat. # 101265), CD16/32 (cat. # 101343), CD86 (cat. # 105047), CD1d (cat. # 123529), CD3 (cat. # 100251), CD4 (cat. # 100569), and CD8a (cat. # 100773). In addition, for hashing purposes, we used BioLegend's Hashtag 1 murine (cat. # 155801), Hashtag 2 murine (cat. # 155803) antibodies.
For flow cytometry: Antibody panels and Fluorescence Minus One controls were generated as appropriate. For this study, we used fluorochrome-conjugated mAbs specific to mouse SiglecF (E50-2440; Becton Dickinson), CD64 (X54-5/7.1; BioLegend), MerTK (2B10C42; BioLegend), CD38 (90; Biolegend) and CD45 (104; Becton Dickinson), along with the following reporter strains: smyc ’::mCherry (mCherry), hsp60’::GFP (GFP), and hspx’::GFP/smyc’::mCherry.

Validation: All antibodies used in this study were obtained commercially and validated by the manufacturers.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): We didn't use cell lines.

Authentication: We didn't use cell lines.

Mycoplasma contamination: We didn't use cell lines.

Commonly misidentified lines (See ICLAC register): We didn't use cell lines.

Palaeontology and Archaeology

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). Permits should encompass collection and, where applicable, export.*

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.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

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

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	C57BL/6J WT, B6.129P2-Nos2tm1Lau/J (NOS2−/−), B6.129S7-Ifngtm1Ts/J (IFNγ−/−), B6.129S7-Rag1tm1Mom/J (Rag1-/-) mice were purchased from The Jackson Laboratory. The mice used in this study were 6–8 wk old. All mice were maintained in a specific pathogen–free animal biosafety level 3 facility at Cornell University. Animal care was in accordance with the guidelines of the Association for Assessment and Accreditation of Laboratory Animal Care. All animal procedures were approved by the Institutional Animal Care and Use Committee of Cornell University.
Wild animals	The study did not involve wild animals.
Reporting on sex	In this study, we did not record the sex of the C57BL/6J WT, B6.129P2-Nos2tm1Lau/J (NOS2−/−), B6.129S7-Ifngtm1Ts/J (IFNγ−/−), and B6.129S7-Rag1tm1Mom/J (Rag1−/−) mice used. The primary focus of our research was to investigate the general immunological responses and gene expression changes in response to Mycobacterium tuberculosis infection and BCG vaccination. Given the extensive nature of the experiments and the primary research questions, sex-based differences were not a central focus of this study. Additionally, previous studies in similar contexts have not shown significant sex-specific effects that would necessitate stratification by sex for the specific outcomes measured in this study. To ensure a balanced representation and avoid sex bias, both male and female mice were included in the study, but sex was not considered a variable in the analysis.
Field-collected samples	The study did not involved field-collected samples.
Ethics oversight	All animal procedures were approved by the Institutional Animal Care and Use Committee of Cornell University.

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.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

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	Lungs were aseptically removed and immersed in PBS containing 5% FBS and Collagenase IV (250U/mL). Samples were immediately processed using a GentleMACS tissue dissociator (Miltenyi Biotec) and maintained on ice. The dissociated lung material was subsequently strained through a 70-µM mesh, and red blood cells were lysed using ammonium-chloride-potassium (ACK) lysis buffer (Lonza).
Instrument	Cells were analyzed with a Symphony A3 (BD Biosciences).
Software	Data were analyzed using FlowJo software (version 10.9; BD).
Cell population abundance	Lung cell suspensions were processed for cell sorting and/or flow cytometry to isolate and analyze specific cell populations. The following steps were taken to determine the abundance and purity of the relevant cell populations within post-sort fractions: For flow cytometry: Single-cell suspensions from lung tissues were stained with fluorophore-conjugated antibodies specific to cell surface markers. The stained cells were analyzed using a BD Symphony A3 flow cytometer. The gates were defined based on FMO controls to accurately identify and isolate target populations. For flow sorting: Post-sort purity was assessed immediately after sorting by re-analyzing a small fraction of the sorted cells by confocal. The purity of the infected sorted populations was typically greater than 95%, as determined by re-analysis of the sorted fractions. Abundance of Cell Populations: The abundance of each relevant cell population within the post-sort/post-analyzed fractions was quantified by calculating the percentage of cells positive for each marker relative to the total number of sorted/analyzed cells. Representative flow cytometry plots and histograms were generated to visually confirm the abundance of the sorted populations.
Gating strategy	We gated based on FSC and SSC on the macrophage population, followed by doublets exclusion. We then gated on infected cells using an mCherry FMO as a control. Infected cells have been gated on macrophages using CD64 and MERTK (or a combination of the two as described in the manuscript) and subsequently alveolar and interstitial macrophages separated by expression of SiglecF. We used FMO controls for the different markers to determine the gates.

☒ 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	*Indicate task or resting state; event-related or block design.*
Design specifications	*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.*
Behavioral performance measures	*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).*

Acquisition

Imaging type(s) *Specify: functional, structural, diffusion, perfusion.*

Field strength *Specify in Tesla*

Sequence & imaging parameters *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.*

Area of acquisition *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization *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.*

Normalization template *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.*

Noise and artifact removal *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings *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).*

Effect(s) tested *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

(See Eklund et al. 2016)

Correction *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

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 *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis *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.).*

Multivariate modeling and predictive analysis *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*