nature portfolio

Corresponding author(s): Zemin Zhang

Last updated by author(s): Sep 13, 2021

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	Data were collected using Cell Ranger software (10x Genomics) version 3.1
Data analysis	We primarily used existing computational tools, and we cited all the sources of the tools we used: https://github.com/satijalab/seurat (v. 3.0) https://pypi.org/project/umap-learn (v. 0.3.7) https://github.com/Teichlab/bbknn (v. 1.5.1) https://github.com/theislab/scanpy (v. 1.4.5) https://github.com/PaulingLiu/ROGUE (v. 1.0) https://github.com/PaulingLiu/scibet (v. 1.0) https://github.com/theislab/scvelo (v. 0.2.5) https://github.com/tanaylab/metacell (v. 0.3.41) https://github.com/milaboratory/mixcr (v. 3.0.13) https://bioconductor.org/packages/release/bioc/html/GSEABase.html (v. 1.54.0) https://github.com/velocyto-team/velocyto.py (v 0.17.16) MiSeq Control Software (MCS) (v. 3.1) No new algorithms were developed for this manuscript. All code generated for analysis is available from the corresponding authors upon reasonable request.

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

All processed scRNA-seq, scTCR-seq, and bulk TCR-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE179994. An interactive web server for analyzing and visualizing the scRNA-seq data is available at http://nsclcpd1.cancer-pku.cn/. Previously published scRNA-seq, scTCR-seq, and bulk TCR-seq that were re-analyzed here are available under accession code GSE123814, GSE120575, GSE99254, GSE108989, GSE98638, GSE113590, GSE141878, GSE24536 and E-MTAB-6149. Source data have been provided as Source Data files. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

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	No statistical method was used to predetermine sample size but our sample sizes are similar to those reported in previous publications (Caushi et al., 2021 and Yost et al., 2019)
Data exclusions	No data was excluded from the analyses
Replication	The single-cell sequencing for each tumour biopsy was performed in one experimental run. This is typical for human studies because the cell counts from a given biopsy sample are usually low and the cells cannot be analyzed more than once. The mIHC experiments for each tumour sample were performed in one experimental run, with ~10 fields of view captured to verify the reproducibility. In addition, at least three tumour samples from each group (treatment responses and conditions) were collected to verify the reproducibility.
Randomization	The patients in this study received anti-PD-1 treatment and its design is not based on randomization. Previous studies have shown that it is the PD-1 pathway blockade that affect the state transition of tumour-reactive T cells (Beltra et al., 2020), and this effect is therefore unlikely associated with covariates like age and gender. In addition, the comparisons between pre- and post-treatment samples were well controlled because these samples were from the same patients and were site-matched. Therefore, the covariates were not controlled.
Blinding	This study is intended to investigate differences between treatment conditions and responses. The Investigators were not blinded to allocation during experiments and outcome assessment so that enough samples were collected for each group.

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
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used: Rabbit anti-BCA1 (1:100, Cat #246518, Clone EPR23400-92, Abcam)

Antibodies used	Rabbit anti-TIM3 (1:50, Cat #45208, Clone D5D5R, Cell Signaling Technology) Mouse anti-CD8A (1:200, Cat #70306, Clone C8/144B, Cell Signaling Technology)
Validation	All antibodies were commercially available and validated by manufacturer. Based on the information from the manufacturer's website, the validation information for species and application is listed below. Rabbit anti-BCA1 (1:100, Cat #246518, Abcam) Reacts with: Human; Suitable for: IHC-P Rabbit anti-TIM3 (1:50, Cat #45208, Cell Signaling Technology) Reacts with: Human; Suitable for: IHC-P, IP, Flow, WB Mouse anti-CD8A (1:200, Cat #70306, Cell Signaling Technology) Reacts with: Human; Suitable for: IHC-P In addition, each antibody was validated based on manufacturer's instructions using human lung tumour samples, and the BCA1 (CXCL13) antibody was further validated using tonsil samples.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Thirty-six patients who were pathologically diagnosed with NSCLC were included in this study. None of the patients was treated with checkpoint blockade inhibitors, radiotherapy, and chemotherapy prior to the first biopsy. LUAD patients receiving anti-PD-1 therapy were treated with 200 mg pembrolizumab, carboplatin (AUC = 6) and 500 mg/m2 pemetrexed every 3 weeks, and LUSC patients were treated with 200 mg pembrolizumab, carboplatin (AUC = 6) and 260 mg/m2 albumin-bound paclitaxel every 3 weeks. Clinical response was assessed for each target lesion based on the RECIST version 1.1. The covariate-relevant population characteristics of the human research participants (e.g. age, gender) could be found in the Supplementary Table 1.
Recruitment	Non-small-cell lung cancer patients receiving anti-PD-1 therapy in Chinese PLA General Hospital were consented to collect bio-specimens before and duing treatment. Given the challenge in collecting consecutive clinical tumour biopsies from NSCLC patients before and during anti-PD-1 treatment, limited samples were used in this study, thus future work based on larger cohorts of anti-PD-1–treated patients are needed to generalize our findings
Ethics oversight	This study complies with all relevant ethical regulations and was approved by the Ethics Committee of Peking University. The written informed consent was provided by all participants.

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