nature portfolio

Corresponding author(s): Mingxiong Guo
COMMSBIO-24-6502-C

Last updated by author(s): Jun 11, 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	Newly generated scRNA-seq data: Collected from 8 samples(KC02A, KC03A, KC05A, KC06A, KC07A, KC08A, KC09A, and KC10A) sequenced specifically for this study. Validation Cohort 1: Viral sequence data for 5 HBV-HCC samples (GSM7441022, GSM7441024, GSM7441027, GSM7441029, and GSM7441031) was downloaded from the NCBI GEO database under accession number GSE234015. Validation Cohort 2: Gene expression data for 5 HBV-positive HCC samples and 3 non-HBV HCC samples was obtained from Lu's study, available under accession number GSE149614 from the GEO database.
Data analysis	Data analysis was performed using R software (version 4.3.2) and Python (version 3.10.11). The Seurat package (v3.1.5) was used for initial preprocessing, clustering, and visualization of scRNA-seq data. To address potential ambient RNA contamination, DecontX was applied after initial quality control and doublet removal using DoubletFinder (v2.0.3). Viral read filtering and alignment were conducted using the Viral-Track pipeline (https://github.com/PierreBSC/Viral-Track) with the viruSITE database (release 2020.3). Batch correction and cell-type label refinement were performed using SCVI and SCANVI models from the scVI-tools package (https://github.com/scverse/scvi-tools). Additional R packages including dplyr, ggplot2, patchwork, ggsignif, tidyr, openxlsx, SingleR, and celldex were used for data manipulation, visualization, and annotation. No custom code or unpublished algorithms were used; all analysis was performed using publicly available tools.

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 single-cell RNA-seq data generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE282343 and are publicly accessible. External scRNA-seq datasets used for validation in this study are available in GEO under accession numbers GSE234015 (Validation cohort 1) and GSE149614 (Validation cohort 2).

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	This study did not include sex and gender-based analysis.
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity, or other socially relevant groupings were not considered in this study.
Population characteristics	The recruited population consists of eight HCC patients, aged between 47 to 66 years. Five participants were HBV-positive, and three were HBV-negative. Tumor sizes ranged from 5.5 cm to 15 cm. Key biomarkers such as ALT, AST, ALB, and AFP levels along with histopathological details indicating fibrosis/cirrhosis status were measured, as reported in Table S1
Recruitment	Participants were recruited based on their diagnosis of HCC. Selection criteria included confirmed clinical diagnosis, with participation being voluntary. Informed consent was obtained from each patient or their guardian prior to participation in the study
Ethics oversight	The HCC tumour tissues were obtained after surgical resection at Hubei Cancer Hospital, Wuhan, Hubei, China. This study was approved by the Research and Ethical Committee of Hubei Cancer Hospital (study approval No: LLHBCH2020LW-026). Informed consent was obtained from each patient or their guardian. All ethical regulations relevant to human research participants were followed.

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	A total of 8 surgical tumor specimens were collected for the discovery cohort submitted under accession number GSE282343, consisting of 5 HBV-positive and 3 non-HBV-HCC patients. Additionally, two publicly available datasets, GSE234015 (validation cohort 1) and GSE149614 (validation cohort 2), were curated to validate and substantiate the findings.
Data exclusions	No data were excluded from the analysis. All collected data met the pre-established quality control thresholds for scRNA-seq analyses.
Replication	The results from the discovery cohort were validated using two independent scRNA-seq datasets from the GEO database (GSE234015 and GSE149614). The findings were reproducible across these datasets, confirming the robustness of the results
Randomization	The samples were selected based on their availability from patients undergoing surgical resection. Since the study focused on observational analysis of tumor microenvironments, randomization was not applicable.
Blinding	Blinding was not applicable since the study involved observational analysis of tumor microenvironments. Knowing the HBV status was essential for proper grouping and interpretation during analysis.

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

CD68 (1:1000, Cat# ab955, Abcam, USA)
CD3 (1:1000, Cat# ab237721, Abcam, USA)
MT1G (1:200, Cat# CSB-PA17384A0Rb, CUSABIO Technology LLC, USA)
CD8 (1:1000, Cat# GB12068, Diagnocine, USA)
IFN-γ (1:200, Cat# abs119966, Absin Bioscience Inc, China)
Anti-HBC (Courtesy of Yuchen Xia's lab, Wuhan University)

Validation

All antibodies used in this study were validated by the respective manufacturers for use in human tissue through immunohistochemistry (IHC). Detailed validation data can be accessed on the manufacturers' websites: CD68 (1:1000, Cat# ab955, Abcam, USA), CD3 (1:1000, Cat# ab237721, Abcam, USA), MT1G (1:200, Cat# CSB-PA17384A0Rb, CUSABIO Technology LLC, USA), CD8 (1:1000, Cat# GB12068, Diagnocine, USA), and IFN-γ (1:200, Cat# abs119966, Absin Bioscience Inc, China).
Anti-HBC was validated internally by Yuchen Xia's laboratory (Wuhan University) to confirm specificity.

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.