nature research

Corresponding author(s): Rafi Ahmed, Andreas Wieland

Last updated by author(s): Jul 16, 2021

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	FACS data was collected on a BD LSR II, FACS Canto II or a FACSymphony A5 (BD Biosciences) using BD FACSDiva V8.0.1. ELISPOT data was aquired using CTL Immunospot Reader. Multiplex IHC data was collected using the Vectra Polaris Imaging System.
Data analysis	FACS data was analysed with FlowJo V10.6.1. GraphPad Prism V7.0 was used for statistical analyses. scRNA-seq analyses were performed using CellRanger v3.1, Seurat v.3.1.4, R (v.3.6.2), VISION R package (v.2.1.0), and Monocle3 (v.0.2.3.0). Multiplex IHC was analysed using Phenochart (v.1.0.12), inFORM software (v2.4.8), phenoptr (v.0.2.5), and phenoptrReports (v.0.2.6).

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

The following protein sequences were employed for predicting and generating HPV peptides: E2 (Uniprot P03120), E5 (UniprotP06927), E6 (Uniprot03126), and E7 (Uniprot P03129). scRNA-seq data are available in the NCBI Gene Expression Omnibus (GEO) database under the accession number GSE180268. Other relevant data are available from the corresponding authors upon 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	Sample size for each experiment is described in the manuscript and is based on the availability of adequate samples. No sample size calculations were performed prior to the study.
Data exclusions	No data were excluded.
Replication	No data or experiments are excluded. No replication was performed on individual patient specimens given limited sample availability. However, analyses were performed over several batches of patients with similar results.
Randomization	No randomization was used in this study as no experimental treatments/groups were involved.
Blinding	Not applicable as this is a descriptive study and no interventions were tested.

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	The antibodies used for flow cytometry are detailed in Ext. Data Table 1. The following antibodies were used for multiplex IHC: anti-CD3 (clone F7.2.38, Dako, #M7254), CD8 (clone C8/144B, eBiosciences, #14-0085-82), CD20 (clone L26, Invitrogen, #14-0202-82), TCF-1 (clone C63D9, Cell Signaling Technologies, #2203S), PD-1 (clone D4W2J, Cell Signaling Technologies, #86163S), and cytokeratin (clone AE1/3, Dako, #GA053)
Validation	The primary antibodies used in this study are widely used and well validated. The mentioned antibodies are tested by immunofluorescent staining with flow cytometric analysis by the manufacturer. The following information is available through the manufacturers' websites: anti-CCR7 staining of human peripheral blood lymphocytes by flow cytometry anti-CD3 staining of human peripheral blood lymphocytes by flow cytometry anti-CD4 staining of human peripheral blood lymphocytes cells by flow cytometry anti-CD8 staining of human peripheral blood lymphocytes by flow cytometry anti-CD14 staining of human peripheral blood mononuclear cells by flow cytometry anti-CD16 staining of human peripheral blood mononuclear cells by flow cytometry anti-CD25 staining of PHA-stimulated (day-3)-stimulated human peripheral blood lymphocytes by flow cytometry anti-CD28 staining of human peripheral blood lymphocytes by flow cytometry anti-CD39 staining of human peripheral blood lymphocytes by flow cytometry anti-CD45RA staining of human peripheral blood lymphocytes by flow cytometry anti-CD45RO staining of human peripheral blood lymphocytes by flow cytometry anti-CD127 staining of human peripheral blood lymphocytes by flow cytometry anti-GranzB staining of human peripheral blood mononuclear cells by flow cytometry anti-IFN-γ staining of PMA/ionomycin-stimulated human peripheral blood mononuclear cells by flow cytometry anti-Ki-67 staining of proliferating human MOLT-4 (T lymphoblastic leukemia, ATCC CRL-1582) cell line and non-cycling peripheral

blood mononuclear cells by flow cytometry
anti-PD1 staining of PHA-stimulated (day-3)-stimulated human peripheral blood lymphocytes by flow cytometry
anti-TCF-1 staining of Jurkat cells by flow cytometry
anti-Tim-3 staining of Th1-polarized human peripheral blood lymphocytes by flow cytometry
anti-TNF-alpha staining of PMA/ionomycin-stimulated human peripheral blood lymphocytes by flow cytometry
anti-TOX staining of C57Bl/6 mice thymocytes by flow cytometry

The antibodies used for multiplex IHC analyses were validated by the manufacturers for immunohistochemistry.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Male and female patients (age 44-75) diagnosed with HNSCC and undergoing surgery were recruited. All patients had previously untreated locally advanced disease at the time of surgery, presented with a locally metastatic lymph node and were classified as Stage I by the AJCC 8th edition staging guide.
Recruitment	Patients undergoing surgery in the Department of Otolaryngology at Emory University Hospital were consented to collect specimens. Recruitment was focused on patients presenting with detectable primary tumours and metastasis to lymphnodes.
Ethics oversight	HNSCC patients undergoing surgery were recruited in accordance with an approved Emory University Institutional Review Board protocol (WINSHIP4008-17), with all patients providing informed consent.

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

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	Tumour samples (primary/metastatic) obtained from transoral robotic surgery and/or radical neck dissection were immediately collected and stored in Leibovitz's L-15 medium. Tissue samples were minced into small pieces and digested using an enzymatic cocktail of collagenase I, II, and IV, elastase and DNAse (Worthington). Lymphocytes were isolated by using a 44%/67% Percoll gradient. Peripheral blood samples were collected in sodium citrate CPT cell preparation tubes (BD) at time of surgery and processed immediately to obtain peripheral blood mononuclear cells (PBMCs) and plasma. After lysis of red blood cells using ACK Lysing buffer, PBMCs were washed 4 times with Dulbecco's phosphate buffered saline without calcium and magnesium (DPBS) plus 2% fetal calf serum (FCS). Tissue lymphocytes and PBMCs were cryopreserved in 90% FCS with 10% DMSO, and eventually stored in liquid nitrogen.
Instrument	FACS data was collected on a BD LSR-II, FACS Canto-II, FACSymphony A5 or during cell sorting on a FACSAria II
Software	FACS data was collected using BD FACSDiva and analysed with FlowJo V10.6.1
Cell population abundance	Post-sort purity was analysed for samples with more than 5000 target cells collected. In all cases purity was greater than 98%.
Gating strategy	Gating strategy is shown in Extended Data section.

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