nature research

Corresponding author(s): Hadrup, Sine Reker

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☒	☐	An indication of 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 statistics 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
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	The binding affinity of each of the amino acid substitutions of the original peptide sequences was predicted using NetMHCpan 4.0 Cross–reactive peptides were predicted from the Shannon PSSMs using the "Find Individual Motif Occurrences" (FIMO) software version 4.11.4 (http://meme-suite.org/meme_4.11.4/tools/fimo) The in silico structure-based pMHCs was made using MODELLER (version 9.18) and the conformation of the original peptide was optimized using the robotics-based kinematic closure protocol from Rosetta (version 2016.20). The program FoldX was used to model all single amino acid substitutions of the original peptide Flow cytometry data was analyzed in FACS DIVA software, version 8.02 Correlation analyses was conducted in GraphPadPrism version 7.03
Data analysis	Sequencing data from the DNA barcode-based screening was analyzed using Barracoda software version 1 (http://www.cbs.dtu.dk/services/barracoda)

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 upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Nature Biotechnology: doi:10.1038/nbt.4303

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

All relevant data and codes are available from the authors. TCR sequences and expression vectors must be obtained through an MTA.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	No population study or comparison is included in the manuscript. For each figure the number of samples, here in terms of TCRs are indicated. The number of samples was determined by the availability of different TCR´s for a given specificity (e.g. n=12) in figure 2.
Data exclusions	Related to Fig. 1i-k: Of a total MHC multimer library of 973 individual pMHCs, two MHC multimers were non-functioning (no reads) and four pMHCs varied greatly between duplicates. These were excluded from the analysis . These exclusion criteria were pre-established.
Replication	Replication was included wherever possible. The number of replications is given in the figure legends. Replicates were evaluated to corrolate. All replications were successful.
Randomization	Randomization is not relevant to our study. We are studying a specific structural interactions. No population cohorts are involved
Blinding	Blinding is not relevant to our study. We are studying a specific structural interactions. No population cohorts are involved

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

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

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	TCR sequences and expression vectors must be obtained from the authors through an MTA.

Antibodies

Antibodies used	CD8-PerCP (Invitrogen MHCD0831, clone 3B5), CD8-BV510 (BD 563256, clone RPA-T8), CD8-BV480 (BD 566121, clone RPA-T8), CD4-FITC (BD 345768, clone SK3), CD14-FITC (BD 345784, Clone MϕP9), CD19-FITC (BD 345776, clone 4G7), CD40-FITC (Serotec MCA1590F, clone LOB7/6), and CD16-FITC (BD 335035, clone NKP15). TNFα-PE-Cy7 (BioLegend 502930), IFNγ-APC (BD 341117),

Nature Biotechnology: doi:10.1038/nbt.4303

	and IL-2-BV421 antibody (BioLegend 500328). CD8a-BV480 (BD 566096, clone 53-6.7) and CD3-FITC (BioLegend 100206, clone 145-2C11). The dilution is given in the method section. Lot no. is not available.
Validation	Validation provided by manufacturer and each antibody was further tested and titrated using human PBMCs to ensure correct performance in the relevant setting.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	wild-type male C57BL/6 mice and OT-1 and OT-3 TCR transgenic male C57BL/6 mice (age 6-12 weeks)
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected from the wild

Human research participants

Policy information about studies involving human research participants

Population characteristics	No population study or population comparison is included in the manuscript. TCRs or T cell clones was derived from representative Merkel Cell Carcinoma patients and published previous: Miller, N. J. et al. Tumor-Infiltrating Merkel Cell Polyomavirus-Specific T Cells Are Diverse and Associated with Improved Patient Survival. Cancer Immunol. Res. 5, 137–147 (2017) and Lyngaa, R. et al. T-cell responses to oncogenic merkel cell polyomavirus proteins distinguish patients with merkel cell carcinoma from healthy donors. Clin. Cancer Res. 20, 1768–78 (2014). Previous publications includes patient characteristics. Not relevant for the current study.
Recruitment	No specific recruitment for the present study. TCRs were selected from predescibed antigen specific T cell populations, taken for the publications given above)

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	Peripheral blood mononuclear cells (PBMCs) from healthy donors were isolated from whole blood by density centrifugation on Lymphoprep (Axis-Shield PoC) and cryopreserved at −150 °C in FCS (FCS; Gibco) + 10% DMSO. Mouse spleen suspensions were obtained by mashing the full spleen through a 70 µm cell strainer (Fischer Scientific). Red blood cells were lysed with RBC Lysis buffer (BioLegend) and used directly or cryopreserved at −150 °C in FCS (FCS; Gibco) + 10% DMSO.
Instrument	Cells were sorted on a BD FACSAriaFusion or acquired on a BD LSR Fortessa.
Software	FACSDiva software was used to gate and sort the population of interest
Cell population abundance	For every T cell population sorted the sorted cell fraction represented 10-80% of the total population
Gating strategy	Lymphocytes were defined within a FSC/SSC plot. Among these we gated on single (FSC-A/FSC-H), live (NIR negative), CD8 positive (PerCP, BV510 or BV480) and 'dump' (CD4, 14, 16, 19, and 40) (FITC) negative cells and sorted either all multimer-positive (PE) cells or all CD8 positive cells within this population. For the mouse splenocytes we gated on single (FSC-A/FSC-H), live (NIR negative), CD8 (BV480) / CD3 (FITC) positive cells

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

Nature Biotechnology: doi:10.1038/nbt.4303