nature research

Corresponding author(s): Yardena Samuels, Reuven Agami .

Last updated by author(s): Oct 23, 2020

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	Mass spectrometry data was acquired using Xcalibur version 4.0; Flow cytometry data was acquired using BD FACSDiva software version.8.0.1. or Attune Nxt software version 4.2. ; RNA-seq data was sequenced using Illumina NexSeq 500 (NextSeq Control Software V 2.2.0).
Data analysis	Purpose for software usage are as described Cutadapt v1.9.1 was used for adapter trimming. Bowtie2 and Tophat2 are used for sequencing data analysis purpose. HTSeq v0.5.4p3 was used for read counting. DESeq v2 was used for normalization of read count matrix and differential gene expression analysis. Samtools v2.3 was used for alignment file conversion purpose. For bed file processing, bedtools v2.25.0 was used. Heatmaps were plotted using pheatmap v1.0.12. For immunopeptidomics data analysis the following software were used: MaxQuant v1.6.0.16.; NetMHCpan 4.0; SSRCalc; GibbsCluster2.0 +; Msnbase 3.11; RStudio 3.5.1 and Python 3.6. In proteomics data analysis, we used proteome discoverer v2.3.0523 and Perseus v1.6.5.1. For 2D proteomics, we used MaxQuant v1.6.06.16. Disorderedness probability was obtained using IUPRED2A (no version information mentioned). For FACS analysis, FlowJo v.10.6.2 use used. Human genome files used for the study are downloaded from Gencode v19. Database used for tRNA information is GtRNAdb. Swissprot was used for human proteome data download. CDS sequences of GRCh38 were downloaded from Ensembl (gencode v19)R. Pseudogenes sequences were taken from gencode protein-coding transcript sequences version 34. For Gibbs analysis, IEDB database was used for reference. The transcript position of the reference point, was converted to protein coordinate using ensembldb 2.8.0. We have designed codes for the study, method for which is described completely in methods section of manuscript as is available in the github links as described in the data availability sections. R versions were routinely updated throughout the proceedings of the manuscript. PERL version used was v5.30.0.

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

Data were deposited in GEO with accession code GSE142822.
The genomic data relevant to the MD55A3 cells are found in BioProject ID PRJNA316754, identified 1M1.
Proteomics and peptidomics data were deposited in PRIDE with accession code PXD020224.
The codes used in the study are available on GITHUB with these links; https://github.com/apataskar/bump_finder_example2, https://github.com/apataskar/accessory_scripts_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	No sample size calculation was performed. Riboseq and RNAseq libraries were prepared using three different melanoma cell lines, with at least two biological replicates, based on previous reports in the field. Proteomics: Set 1, we performed proteomics analysis of triplicates for quantifying gene expression signature difference as demonstrated in Figure 2 and Extended Figure 3. Proteomics: Set 2, we performed 2D proteomic analysis for extraction of aberrant peptides on duplicates for 24 fractions in Control and IFNγ-induced conditions. For Immunopeptidomics, overall sixteen samples were analyzed: four from fresh melanoma tumor metastasis, and twelve from treated and untreated melanoma cell line, including biological replicates. For the aberrant peptide screens for immunogenicity, sample size was chosen to include as much of the naive T cell repertoire as possible from healthy donors (regular blood donation of 450 ml per donation). PBMCs from seven healthy blood donors were used. We received buffy-coats from three anonymous healthy blood donors from Oslo University Hospital Blood Bank in accordance with the institutional guidelines. Blood for the four remaining healthy blood donors was processed and HLA-typed in-house under informed consent and in accordance with the institutional guidelines
Data exclusions	No data was excluded for this study.
Replication	Sequencing data used in this study are generated at least in duplicates, and the observations presented in this study are reproducible in all replicates. The total number of replicates for all experiments is mentioned in the figure legends for each result. Immunopeptidomics analysis using treated and untreated cell lines were done with four replicates for each group. Supplementary table 1 indicates for each of the 28 aberrant peptides, in which sample/s it was identified.
Randomization	For all high throughput data analysis- randomization for allocation of samples into experimental groups is not relevant, as the design of the study requires explorative analysis for identifying discriminative features between already established experimental groups.
Blinding	For all high throughput data analysis- blinding to investigators is not relevant to this study, as the design of the study requires explorative analysis for identifying discriminative features.

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

Mouse anti-V5 Tag Monoclonal Antibody (Thermo Fisher Scientific, cat. nr. R960-25; 1:1000) Rabbit anti IDO1 Monoclonal Antibody (Cell Signaling, cat. nr. 86630; 1:1000) Mouse Anti-α-Tubulin Monoclonol Antibody (Sigma, cat. nr. T6199; 1:5000) Rabbit anti TurboGFP (Thermo Scientific, cat. Nr. PA5-22688; 1:1000) APC-conjugated anti-mouse H-2Kb bound to SIINFEKL (clone clone 25-D1.16, Biolegend, cat nr. 141606; 1:200).anti-human CD8-FITC (Biolegend, clone RPA-T8, cat no. 301050; 1:200), anti-human CD8a-BV421 (Biolegend, clone RPA-T8, cat no. 301036; 1:200) and anti-human CD137-Alexa Fluor 647 (Thermo Fisher Scientific, clone 4B4-1, cat no. A51019; 1:100).
Secondary Antibodies:
Peroxidase-conjugated Goat anti mouse antibody (Jackson ImmunoResearch #115-035-003, 1:10,000); Peroxidase-conjugated Goat anti rabbit antibody (Jackson ImmunoResearch #111-035-003, 1:10,000)
IRDye 680RD Donkey anti-Mouse (LI-COR, 926-68072; 1:10,000)
IRDye 800CW Goat anti-Rabbit (LI-COR, #926-32211, 1:10,000)

Validation

All antibodies used in the study were validation by the manufacturer, data is available at the manufacturer's website, as indicated below:
Rabbit anti IDO1 (D5J4ETM) Monoclonal antibody (cell signaling, cat. Nr. 86630)-
https://www.cellsignal.com/products/primary-antibodies/ido-d5j4e-rabbit-mab/86630?Ntk=Products&Ntt=86630
Mouse anti α-Tubulin antibody clone DM1A, (SIGMA, cat. T6199).
https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Datasheet/t6199dat.pdf
Mouse anti-V5 Tag Monoclonal Antibody (Thermo Fisher Scientific, cat. nr. R960-25) https://www.thermofisher.com/antibody/product/V5-Tag-Antibody-Monoclonal/R960-25
Rabbit anti TurboGFP (Thermo Scientific, cat. Nr. PA5-22688) https://www.thermofisher.com/antibody/product/TurboGFP-Antibody-Polyclonal/PA5-22688
APC-anti-mouse H-2Kb bound to SIINFEKL (clone clone 25-D1.16, Biolegend, cat nr. 141606)
https://www.biolegend.com/en-us/products/apc-anti-mouse-h-2kb-bound-to-siinfekl-antibody-7882
anti-human CD8-FITC (Biolegend, clone RPA-T8, cat no. 301050)
https://www.biolegend.com/en-us/search-results/fitc-anti-human-cd8a-antibody-834
anti-human CD8a-BV421 (Biolegend, clone RPA-T8, cat no. 301036)
https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-human-cd8a-antibody-7152
anti-human CD137-Alexa Fluor 647 (Thermo Fisher Scientific, clone 4B4-1, cat no. A51019)
https://www.thermofisher.com/antibody/product/CD137-Antibody-clone-4B4-1-Monoclonal/A51019
eroxidase-conjugated Goat anti mouse antibody (Jackson ImmunoResearch #115-035-003)
https://www.jacksonimmuno.com/catalog/products/115-035-003
Peroxidase-conjugated Goat anti rabbit antibody (Jackson ImmunoResearch #111-035-003)
https://www.jacksonimmuno.com/catalog/products/111-035-144
IRDye 680RD Donkey anti-Mouse (LI-COR, cat- no. 926-68072)
https://www.licor.com/bio/reagents/irdye-680rd-donkey-anti-mouse-igg-secondary-antibody
IRDye 800CW Goat anti-Rabbit (LI-COR, #926-32211, 1:10,000)
https://www.licor.com/bio/reagents/irdye-800cw-goat-anti-rabbit-igg-secondary-antibody

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Cell lines 12T and 108T were derived from pathology-confirmed metastatic melanoma tumor resections collected from patients enrolled in institutional review board (IRB)-approved clinical trials at the Surgery Branch of the National Cancer Institute, Prof. Steven Rosenberg lab. The MD55A3 cell line was given by Prof. Jennifer Wargo lab. This cell line is derived from metastatic melanoma tumor resections collected with informed patient consent under a protocol approved by the NIH Institutional Review Board (IRB) Ethics Committee and approved by the MD Anderson IRB (protocol numbers 2012-0846, LAB00-063, and 2004-0069; NCT00338377).
D10 and 888-Mel human melanoma cells and MART-1 TCR CD8 T cells were given by Prof. Daniel Peeper (PMID: 31303383).
Hek293T (ATCC CRL-11268) and A375 (ATCC CRL-1619) cells are from in-house stocks (Prof. Reuven Agami's lab).
Class I HLA-negative K562 cell line was purchased from ATCC (ATCC-CCL-243) and retrovirally transduced with HLA B*07:02.
PBMCs were isolated from the blood of healthy donors provided by the Norwegian Blood Bank. In-house donor PBMCs were isolated from the healthy donors under informed consent and HLA-typed. (Prof. Johanna Olweus´s lab).

Authentication	Cell lines 12T and 108T were authenticated by Finger printing with STR profiling (Panel: PowerPlex_16_5Nov142UAGC, Size: GS500 x35 x50 x250, Analysis Type: Fragment (Animal), Software Package: SoftGenetics GeneMarker 1.85)
Mycoplasma contamination	All cell lines tested negative for mycoplasma in a PCR-based assay.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Human research participants

Policy information about studies involving human research participants

Population characteristics	We received buffy-coats from anonymous healthy blood donors from Oslo University Hospital Blood Bank in accordance with the institutional guidelines. In addition, we processed and HLA-typed blood from recruited healthy in-house donors under informed consent. All enlisted donors are healthy persons residing in Norway during the study period. For donors provided by Oslo University Hospital Blood Bank we do not have data available for donor age, gender, ethnicity or disease profile. For recruited healthy in-house donors their name, date of birth, hemoglobin and ferritin levels, and HLA type were recorded. All the information was de-identified and each donor was assigned a unique code that connects information through a list of names stored in a secure location
Recruitment	HLA-A*03:01 positive buffy-coats from the blood bank were used without any further selection criteria for donor age, gender or ethnicity. In-house donors were selected based on their HLA type (HLA-A*03:01, HLA-B*07:02, HLA-C*07:02).
Ethics oversight	PBMCs were isolated from the blood of healthy donors donated to the Norwegian Blood Bank under informed consent. In-house donor PBMCs were isolated from the healthy donors under informed consent and HLA-typed.

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	Cells were harvested by trypsinization, and centrifugation after which the cells were immediately analyzed on FACS for tGFP positivity. For SIINFEKL positivity testing, cells were detached using versene (PBS-EDTA), after which they were stained with APC-conjugated anti-H2-Kb-bound SIINFEKL antibodies. Data was acquired after cells were thoroughly washed. In order to screen peptides for immunogenicity, CD8+ T cells were harvested after co-culture with peptide-pulsed autologous dendritic cells (PMID: 31101906), stained with FITC-conjugated anti-human CD8 antibody, APC-Cy7-conjugated Live/Dead dye and a pool of pMHC multimers, and analysed on FACS (BD LSR II) for double pMHC multimer positivity. In order to to characterize the reactivity of single cell sorted T cells, expanded T cell clones re-stimulated with peptide-loaded target cells were harvested and stained with BV421-conjugated anti-human CD8 and Alexa Fluor 647-conjugated anti-human CD137 antibodies and APC-Cy7-conjugated Live/Dead dye, and analysed on FACS (BD LSR II) for CD137 positivity. (Prof. Johanna Olweus`s lab)
Instrument	Attune Nxt, BD LSR II or BD LSR Fortessa
Software	Attune NxT Software or FACSDiva (BD LSR II and BD LSR Fortessa) were used for acquisition, FlowJo V10 software was used for data analysis
Cell population abundance	A total of 180 pMHC multimer+ single CTLs were sorted for both KCNK6 (0.05% of pMHC multimer+ CD8+ T cells) and ZNF513-reactive (0.009% of pMHC multimer+ CD8+ T cells) T cell populations using SH800 cell sorter (Sony Biotechnology) and expanded on feeder cells. Sixteen out of 180 KCNK6 pMHC multimer reactive CTL clones expanded sufficiently for subsequent analysis. Thirteen out of sixteen CTL clones stained positively with KCNK6 pMHC multimers. None of the ZNF513 pMHC multimer reactive CTL clones expanded sufficiently for subsequent analysis.
Gating strategy	Single cells were gated out using SSC-A versus SSC-H dot plots. FACS data from combinatorial pMHC multimer staining was analysed by gating on viable CD8+ T cell singlets by FSC-A/H and SSC-A/H gates, Live/Dead fixable near-IR staining and CD8 staining that was followed by Boolean gating to define double pMHC multimer positive events. FACS data from functional assay with T cell clones was analysed by gating on viable CD8+ T cell singlets by FSC-A/H and SSC-A/H gates, Live/Dead fixable near-IR staining and CD8 staining. Subsequently, activated CD8+ T cells were identified as CD137+ events, where cut-off was set based on staining of T cells incubated without target cells. (Prof. Johanna Olweus`s lab)

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