nature portfolio

Corresponding author(s): Xiaoxiao Wan

Last updated by author(s): Aug 25, 2024

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	The T cell reativity data were collected using the Cell counting software v 5.0 (ImmunoSpot). The flow cytometry data were collected using the FACSDiva 9.0 software (BD Biosciences). The mass spectrometry data were collected using XCalibur v 4.0 (ThermoFisher Scientific).
Data analysis	The biological data were analyzed using the Graphpad Prism 10.0.0. The flow cytometry data were analyzed using the Flowjo 10.0 software (Tree Star Software). The mass spectrometry data were analyzed using PEAKS X (Bioinformatics Solutions).

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 data generated and supporting the findings of this study are available within the paper. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD030999 and 10.6019/PXD030999.

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	No human participants invovled in the study
Reporting on race, ethnicity, or other socially relevant groupings	No human participants invovled in the study
Population characteristics	No human participants invovled in the study
Recruitment	No human participants invovled in the study
Ethics oversight	No human participants invovled in the study

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	Sample sizes were not predetermined and were indicated in the figure legends. For Mass Spectrometry analysis, the samples were chosen based on flow cytometry data quantifying the number of the APCs we could isolate from thymus or islets. The sample size was then determined with the number of the APCs comparable with our previous analysis on islet peptidome. For Flow Cytometry analysis, individual mouse was used and multiple independent experiments with at least 3 mice in each group were performed. For single cell analysis, the number of the mice was determined by pilot experiments identifying the number of the cells can be isolated from one mouse.
Data exclusions	We did not exclude any data in our study.
Replication	All data were reproduced independently as stated in the figure legends
Randomization	For Mass Spectrometry analysis, mice at young age with both sexes were used. For Flow Cytometry and cell culture, sex and age matched mice were used and the mice were assigned randomly to experimental groups.
Blinding	Investigators were not blinded to group allocation during experimental setup, data collection, and the 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	For in vivo blockade treatment, rat IgG1 anti-mouse CD25 (PC61), mouse IgG2b anti-CTLA4 (9D9), and mouse IgG1 anti-TGF-β1,2,3 (1D11.16.8) were purchased from Leinco Technologies. For flow cytometry, we used the following fluorescently conjugated

antibodies purchased from BioLegend: anti-CD45 (30-F11), anti-CD11c (N418), anti-CD11b (M1/70), anti-CD3e (145-2C11), anti-CD4 (GK1.5 or RM4-5), anti-CD8a (53-6.7), anti-CD44 (IM7), anti-CD28 (37.51), anti-CD25 (PC61.5), anti-EpCAM1 (G8.8), anti-CD80 (16-10A1), anti-CD90.2 (53-2.1), anti-B220 (RA3-6B2), anti-FR4 (12A5), anti-CD69 (H1.2F3) and anti-TER-119 (TER-119); or the following antibodies purchased from other companies: anti-CD45RA (14.8, BD Pharmingen), anti-CD69 (H1.2F3, Invitrogen), anti-Ly51 (6C3, eBioscience), anti-CD73 (TY/11.8, eBioscience), and anti-Foxp3 (FJK-16s, eBioscience), anti-CD86 (GL1, eBioscience), and Rabbit anti-LAMP1 (EPR21026; Abcam). The anti-I-Ag7 antibody (AG2.42.7) and anti-insulin peptide antibodies (6F3 and AIP) were generated in our laboratory. Three hashtag antibodies were used for cell sorting for scRNA seq analysis: TotalSeq™-C0301 anti-mouse Hashtag 1 Antibody (Barcode sequence: ACCCACCAGTAAGAC); TotalSeq™-C0302 anti-mouse Hashtag 2 Antibody (Barcode sequence: GGTCGAGAGCATTCA); TotalSeq™-C0303 anti-mouse Hashtag 3 Antibody (Barcode sequence: CTTGCCGCATGTCAT). The I-Ag7-peptide tetramers are generated in Dr. Luc Teyton's laboratory.

Validation

The reactivity of anti-I-Ag7 antibody was confirmed by flow cytometry experiments showing that 95%-100% of the C3g7 APCs were stained positive by the Pacific Blue-conjugated AG2.42.7 antibody, whereas less than 1% of the cells of a I-Ag7 negative control cells line were stained positive with the antibody. Other antibodies are commercially available and validation materials are available on the appropriate websites. Tetramers were validated by the publication using the same tetramers before and specificity of each monomers are validated by stimulating antigen specific T cells to produce interleukin-2 (IL-2) or by activation marker CD69. An irrelevant T cell hybridoma is used as a negative control. C-peptide specific tetramer is confirmed by staining the cells from draining lymph nodes from the mice immunized with cognate C-peptide.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	The CD4+ T cell hybridomas and the C3.g7 B cell lymphoma line (used as APCs) were generated in the laboratory.
Authentication	The reactivity of each CD4+ T cell hybridoma to their specific antigens were tested and confirmed by their ability to produce IL-2 upon stimulation with antigen presenting cells with cognate antigens.
Mycoplasma contamination	All the cell lines were confirmed as mycoplasma free.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in the study.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	NOD/ShiLtJ (NOD), NOD.129S2(B6)-Ins2tm1Jja/GseJ (NOD.Ins2-/-), and NOD.Cg-Tg(Ins2*Y16A)1EllIns1tm1JjaIns2tm1Jja/GseJ (NOD.B16A) were originally obtained from the Jackson Laboratory. NOD.10E11 TCR transgenic mice (TCRα: TRAV5D-4/ TRAJ42; TCRβ: TRBV13-3/ TRBD2/TRBJ2-7) were generated previously.
Wild animals	This study did not include wild animals.
Reporting on sex	The sex of the mice used for the experiments was stated in the main text of the manuscript as well as in the figure legends
Field-collected samples	This study did not include samples collected from the field.
Ethics oversight	All experiments were approved by the Division of Comparative Medicine of Washington University School of Medicine in St. Louis (Accreditation number A3381-01).

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	Spleens and lymph nodes were collected and meshed through 70 µm cell strainer and red blood cells were lysed by red blood cell lysis buffer (ACK buffer) to make single cell suspensions. Thymi were collected with fat tissues removed and digested with 100 µg/ml of Liberase (Sigma-Aldrich) and 20 µg/ml of DNase I in DMEM solution at 37°C for 30 min to prepare single cell suspension. The separated cell portions after tetramer incubation were collected for staining. To block Fc-receptors engaging, the cell suspensions were incubated at 4 °C for 5 min in PBS (pH 7.4) supplemented with 2% FBS (fetal bovine serum) and

50% of FC-block (made in-house). For surface staining, cells were incubated with fluorescently labeled antibodies (1:200 [vol/vol]) at 4 °C for 20 min. Cells were then washed and analyzed by flow cytometry or subjected to FACS sorting. For staining of intracellular transcription factors, islet cells were stained with surface antibodies for 20 min, fixed, and permeabilized using the FOXP3 transcription factor staining buffer set (Thermo Fisher Scientific) per the manufacturer's instructions. Cells were incubated with the anti-FOXP3 antibody at room temperature for 30 min.

To sort the antigen specific T cells for scRNA-seq analysis, two sets of tetramers were conjugated with APC-labeled Streptavidin at a 5:1 molar ratio individually at room temperature for at least 1 hour. Thymic set included CP2, E21, DBP, ZnT8 and islet set included CP1d, G20, 2.5HIP, 6.9HIP epitopes. After conjugation, four individual thymic or islet tetramers were combined. The SLO tissues from two mice were combined for isolation and multiple tubes were used due to the number of the mice. For first experiment, 14 mice were used for thymic epitope specific T cells and 6 mice were used for islet epitope specific T cells. In second experiment, 10 mice were used for thymic epitope specific T cells and 4 mice were used for islet epitope specific T cells. After red blood cell lysis, cells were washed with MACS buffer and the CD19+ B cells and CD11b+ myeloid cells were depleted by CD19 beads (Miltenyi Biotec, 120 µl for each sample) and CD11b beads (Miltenyi Biotec, 80 µl for each sample) per instruction. After depletion, the cells were washed twice with PBS and incubated with two sets of combined tetramers at room temperature for 1 hour with shaking. After incubation, cells were washed with MACS buffer and incubated with 40 µl of anti-APC microbeads at 4 °C for 20min. Tetramer positive cells were enriched by LS columns. 10E11 CD4+ T cells were prepared by meshing the spleen through 70µm strainer, lysing the red blood cells, and enriched by Naive CD4+ T Cell Isolation Kit (Miltenyi Biotec). Before surface staining, samples were incubated with 100µl of TotalSeq™ anti-mouse Hashtag reagent (1:100 [vol/vol], BioLegend) to tag the samples: enriched cells from islet tetramer sets were tagged with TotalSeq™-C0301 anti-mouse Hashtag 1 Antibody (Barcode sequence: ACCCACCAGTAAGAC); enriched cells from thymic tetramer sets were tagged with TotalSeq™-C0302 anti-mouse Hashtag 2 Antibody (Barcode sequence: GGTCGAGAGCATTCA); 10E11 naïve CD4+ T cells were tagged with TotalSeq™-C0303 anti-mouse Hashtag 3 Antibody (Barcode sequence: CTTGCCGCATGTCAT). Then 2X antibody mixtures (1:100 [vol/vol]) for surface staining of surface markers were made. For cells with tetramer staining, PE-anti-Thy1.2, eFluor450 Pacific Blue-anti-CD73, PE-Cy7-anti-CD4, PercP-eFluro710-anti-CD8, eFluor 780-Fixable Viability Dye (Invitrogen), AF488-anti-B220, I-Ag7, CD11b were used. For 10E11 naïve T cells, APC-anti-CD62L, PE-anti-Thy1.2, BV421-anti-CD44, PE-Cy7-anti-CD4, PercP-anti-CD25, CD69, eFluor 780-Fixable Viability Dye, AF488-anti-B220, I-Ag7, CD11b were used. After 20min at 4 °C, cells were washed twice with FACS buffer. and underwent cell sorting on BD FACSAria II sorter (BD Biosciences).

Regular flow cytometry samples were analyzed by BD FACSCanto II (BD Biosciences) and FACS cell sorting was performed using the BD FACSAria II sorter (BD Biosciences). All the data were analyzed using the FlowJo 10.0 software (Tree Star).

Instrument

FACS Canto II (BD Biosciences), FACS Aria II cell sorter (BD Biosciences).

Software

Data were collected using FACS DIVA v 8.0.1 and 9.0(BD Biosciences) and were analyzed using Flowjo v. 10.00 software (Tree Star Software).

Cell population abundance

The abundance of the cells was based on the quantification by the cell counting beads (PCB100, Invitrogen) when performing flow cytometry analysis. For cell sorting, the abundance of the cells was based on the number shown by the FACS DIVA software and the eventual cell counts after scRNA seq. The post sort was performed to confirm a purity above 99%.

Gating strategy

For tetramer staining analysis, cells were gated as singlets, CD45+, Lineage negative (I-Ag7, CD11b, CD19 or B220, live/dead), CD3+ or Thy1.2+ cells. Then the cells were further gated into CD4+ or CD8+ T cells. The positive gate for CD4+ tetramer+ cells were based on the negative staining of CD8+ T cells. Then the CD4+ tetramer+ cells were further gated by CD44, Foxp3 or CD25, Foxp3 or FR4, CD73, and Foxp3. For thymic APCs, singlets, CD45-, EpCAM+, Ly51 low cells were gated as mTECs; CD45+ were gated first for other APCs, then CD11c and B220 were used for isolating B cells (CD11c- B220+) and pDC (CD11c+ B220 +), and conventional DCs (CD11c+ B220-). In conventional DC, CD8α and Sirpα were further used to gate for DC1 (CD8α +Sirpα-) and DC2 (CD8α-Sirpα+). The islet macrophages were gated as singlets, live, CD45+, F4/80+, CD11c+ cells. All the gating strategies can be found in the Supplementary information.

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