**SUPPLEMENTARY MATERIAL**

# CD8$^+$ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes

Marie Eliane Azoury[1*], Fatoumata Samassa[1*], Mijke Buitinga[2], Laura Nigi[3], Noemi Brusco[3], Aïsha Callebaut[2], Matthieu Giraud[4], Magali Irla[5], Ana Ines Lalanne[1], Alexia Carré[1], Georgia Afonso[1], Zhicheng Zhou[1], Barbara Brandao[1], Maikel L. Colli[6], Guido Sebastiani[3], Francesco Dotta[3], Maki Nakayama[7], Decio L. Eizirik[6,8], Sylvaine You[1], Sheena Pinto[9], Mark J. Mamula[10], Yann Verdier[11], Joelle Vinh[11], Soren Buus[12], Chantal Mathieu[2], Lut Overbergh[2] Roberto Mallone [1,13]

[1]Université de Paris, Institut Cochin, CNRS, INSERM, 75014 Paris, France.
[2]KU Leuven, Laboratory of Clinical and Experimental Endocrinology, 3000 Leuven, Belgium.
[3]University of Siena, Department of Medicine, Surgery and Neuroscience, Diabetes Unit and Fondazione Umberto di Mario ONLUS, Toscana Life Sciences, 53100 Siena, Italy.
[4]Université de Nantes, INSERM UMR1064, Centre de Recherche en Transplantation et Immunologie (CRTI), 44093 Nantes, France.
[5]Aix-Marseille University, CNRS, INSERM, CIML, Centre d'Immunologie de Marseille-Luminy, Marseille, France[6]Université Libre de Bruxelles, Center for Diabetes Research and Welbio, Medical Faculty, 1070 Brussels, Belgium.
[7]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO 80045, USA.
[8]Indiana Biosciences Research Institute, Indianapolis, IN 46202, USA.
[9]Deutsches Krebsforschungszentrum (DKFZ), Division of Developmental Immunology, 69120 Heidelberg, Germany.
[10]Yale University School of Medicine, New Haven, CT 06510, USA.
[11]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS USR3149, 75005 Paris, France.
[12]Panum Institute, Department of International Health, Immunology and Microbiology, D-2200 Copenhagen, Denmark.
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, 75014 Paris, France.

*M.E.A. and F.S. contributed equally to this work.

## SUPPLEMENTARY METHODS

### mTEC isolation

For RT-qPCR and proteomics, thymic lobes were harvested from 6-8-week-old C57BL/6 mice and NOD mice, cleaned of connective tissue and fat, cut into fragments, washed in RPMI and gently agitated thrice for 30 s to release excess thymocytes. At each wash, the fragments were allowed to settle for 10 min on ice, followed by removal of the thymocyte-rich supernatant. Fragments were subsequently digested for 20 min at 37°C in ultra-low attachment 6-well plates (Corning) using 5 ml of ice-cold Liberase TH (2.5 U/mL; Roche) supplemented with 100 U/ml DNase I (Applichem). Supernatants were then removed and collected in ice-cold PBS supplemented with 0.5% BSA and 2mM EDTA. This enzymatic digestion was repeated twice for 15 min with 2.5 ml Liberase. Syringe needles (18G and 25G) were used to break up tissue after the second and third digestion, respectively. Cells were filtered through a 100-µm mesh before enrichment for TECs. To this end, $2x10^8$ cells/ml were incubated in PBS/2% FCS with 2.6 µg/ml anti-CD90.2 mAb (clone 30-H12, Sigma) on a rotating shaker for 30 min at 4°C, washed and plated ($1x10^7$ cells/ml) on panning plates previously coated overnight with 10 µg/ml goat anti-rat Ab (Sigma) in 50 mM Tris buffer pH 9.5. After 30 min at room temperature, cells were collected, washed and incubated with the Fixable Viability Dye eFluor780 (eBioscience). Cells ($40x10^6$/ml) were then stained for CD45-PE/Cy7 (RRID:AB_2734986), MHC Class II (MHC-II)-AF488 (RRID:AB_493138 and RRID:AB_493147 for C57BL/6 and NOD mice, respectively), EPCAM-BV421 (RRID:AB_2563983), BP-1-APC (RRID:AB_2762698), *Ulex europaeus* agglutinin-1 (UEA-1-biotin; Vector Laboratories), and streptavidin-PE (eBioscience). Cells ($40x10^6$/ml) were sorted with a BD Influx (100-µm nozzle, 18 psi) equipped with 488 nm (200 mW), 640 nm (120 mW), 405 nm (100 mW) and 561 nm (150 mW) lasers. Sorted cells were collected in 1.5-ml tubes in PBS/2% FCS.

For RNAseq, mouse thymi were extracted from a pool of 4 C57BL/6 mice, cut into small pieces, agitated into RPMI medium to release thymocytes, and digested with collagenase D (1 mg/ml; Roche) and DNase I (1 mg/ml; Sigma) for 30 min at 37 °C. The remaining fragments were digested with collagenase/dispase (2mg/ml; Roche) and DNase I (2 mg/ml) at 37 °C to obtain a cell suspension. After filtration through a 70-µm mesh, cells were resuspended in PBS containing 1% fetal bovine serum (FBS) and 5 mM EDTA. The remaining thymocytes were removed by magnetic depletion with CD45 MicroBeads (Miltenyi). Cells were then stained with CD45-PerCPCy5.5 (1:50; RRID:AB_893340), Ly51-PE (1:800; RRID:AB_313365), and MHC-II I-A/I-E-APC (1:1,200; RRID:AB_469455). MHC-II$^{high/low}$ mTECs (CD45$^-$Ly51$^-$ MHC-II$^{high/low}$) were sorted into TRIzol (ThermoFisher) on a BD FACSAria III instrument.

**RT-qPCR and measurement of Padi enzymatic activity**

Total RNA was extracted from sorted mTECs or 50 islets, using the Single Cell RNA Purification Kit (Norgen) and cDNA synthesized with SuperScript VILO (Invitrogen). The RT-qPCR forward (F) and reverse (R) primers were: *Padi2*-F: 5'-CCGCCGGGTATGAAATAGTCC-3', *Padi2*-R: 5'-CGCCGGTGTACTTGACCAC-3'; housekeeping *Act*-F: 5'-AGAGGGAAATCGTGCGTGAC-3', *Act*-R: 5'-CAATAGTGATGACCTGGCCGT-3'; housekeeping *Rpl27*-F: 5'-GTCGAGATGGGCAAGTTCAT-3', *Rpl27*-R: 5'-TTCTTCACGATGACGGCTTT-3'; and predesigned *Hprt* TaqMan primers (ThermoFisher). RT-qPCR was performed using 4 pmol primers, 0.2 µl cDNA and 5 µl Fast SYBR Green Master Mix (Applied Biosystems) on a StepOnePlus RT-PCR System (Applied Biosystems). The relative fold gene expression was calculated using the delta-delta Ct method.

Padi activity was measured with the Ab-based assay for Padi activity (ABAP; ModiQuest Research).

**Liquid chromatography-tandem mass spectrometry (LC-MS/MS)**

mTECs were washed thrice with PBS at 400g for 5 min and lysed in a buffer composed of 7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris base, 1% dithiothreitol and Complete protease inhibitors (Roche). Debris were removed by centrifugation for 10 min at 13,000 rpm at 4°C. Samples were dialyzed and reduction, alkylation, protein precipitation, digestion and desalting performed as described (1). The purified peptides were vacuum-dried and dissolved in mobile phase A, containing 2% acetonitrile and 0.1% formic acid. An estimated final amount of 0.5 µg/µl was injected. The sample was separated by reversed-phase chromatography using a micropillar array column (µPAC$^{TM}$ C18, 200 cm, PharmaFluidics). A linear gradient of mobile phase B (0.1% formic acid in 98% acetonitrile) from 1% to 40% in 80 min was followed by a steep increase to 100% mobile phase B in 5 min. After 5 min at 100% mobile phase B, a steep decrease to 1% mobile phase B was achieved in 5 min and 1% mobile phase B was maintained for 35 min with a flow rate of 750 nl per minute. LC-MS/MS was performed on a Q-Exactive Plus equipped with a nanospray ion source (ThermoFisher). Full-scan spectrum (350 to 1850 m/z, resolution 70,000, automatic gain control 3E6, maximum injection time 100 ms) was followed by high-energy collision-induced dissociation (HCD) tandem mass spectra with a run time of 120 min. Peptide ions were selected for fragmentation by tandem MS as the 20 most intense peaks of a full-scan mass spectrum. HCD scans were acquired in the Orbitrap

(resolution 17,500, automatic gain control 1E5, maximum injection time 80 ms). Peptides were identified by MASCOT (Matrix Science) using SwissProt (Homo sapiens, 169,779 entries) as a database via Proteome Discoverer 2.2, incorporating Percolator for peptide validation. Oxidation (M), deamidation (N/Q), and deamidation (R) (referring to citrullination), were included as variable modifications, carbamidomethylation (C) as a fixed modification. Two missed cleavages were allowed, peptide tolerance was set at 5 ppm and MS/MS tolerance at 20 mmu. MS/MS spectra were checked manually for the presence of citrullinated residues as described (2).

**RNAseq and gene set enrichment analysis (GSEA)**

Total RNA was extracted following the TRIzol protocol, adding GlycoBlue (ThermoFisher) as an RNA carrier. This RNA was used to generate poly-A-selected transcriptome libraries using the non-directional TruSeq V3 RNA Sample Prep Kit (Illumina) following the manufacturer's protocol. Sequencing was carried out on an Illumina HiSeq 2000 machine and was paired-end (2x100 bp) for MHC-II$^{high}$ mTECs and single-end (50 bp) for MHC-II$^{low}$ mTECs. These datasets have been deposited under GEO: (GSE submitted). RNAseq data from published (3) MHC-II$^{high}$ and MHC-II$^{low}$ mTEC (GSE140815) and from *Aire*$^{-/-}$ MHC-II$^{high}$ mTEC (GSE140683) biological replicates were obtained with the same procedure.

Low-quality sequencing reads were removed using the Illumina CASAVA 1.8 pipeline and homogenized to 50-bp reads by trimming and merging the paired-end 100-bp reads. These reads were then aligned to the mm9 mouse reference genome using Bowtie (4). Read numbers were generated using the intersectBed with parameters -f 1.0 and the coverageBed programs of the BEDtool distribution (5), with a GTF annotation file that was generated from the UCSC Table Browser, in choosing mm9 mouse genome/Genes and Gene Predictions/RefSeq Genes/refGene. Reads per kilobase per million mapped reads (RPKM) values were then computed for each sample according to the transcript length of each gene and the global number of reads that map to the mm9 genome.

The enrichment of sets of genes driving PTMs between wild-type (WT) and *Aire*$^{-/-}$ MHC-II$^{hi}$ mTECs and between MHC-II$^{high}$ and MHC-II$^{low}$ mTECs was analyzed with the GSEA software (6). To this end, the sequencing read numbers corresponding to each group of biological replicates in each comparison were normalized using DESeq (7) and used to run GSEA with a "classic" scoring scheme and n=10,000 permutations. The conservative family-wise error rate statistics ($P_{FWER}$) was used to determine whether a set of genes was significantly over-represented in a group of mTECs versus another.

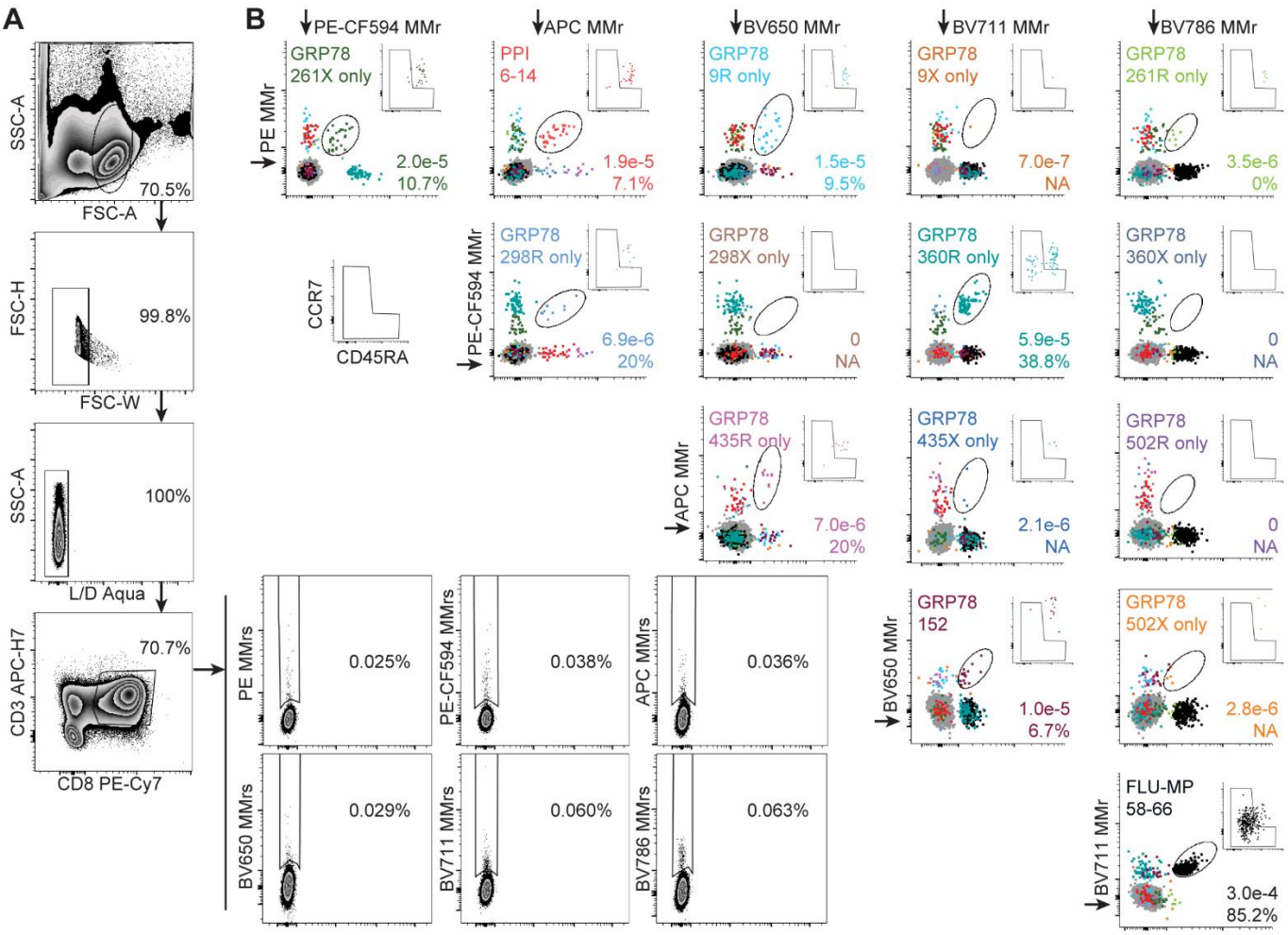
Microarray data from MHC-II$^{high}$ mTECs isolated from 3 WT and 3 *Aire*$^{-/-}$ NOD mice were obtained from the NCBI GEO database (GSE12073) (8). CEL raw data files were processed with GenePattern (9) and the ExpressionFileCreator module for normalization using the RMA method (10) with background correction. Normalized data were plotted as fold change expression. False discovery rate *p* values were computed using the NCBI GEO2R online tool. Enrichment of PTM enzyme gene sets between WT and *Aire*$^{-/-}$ NOD MHC-II$^{high}$ mTECs was analyzed using GSEA on normalized microarray expression data using a "classic" scoring scheme and nominal *p* values.

**Confocal microscopy**

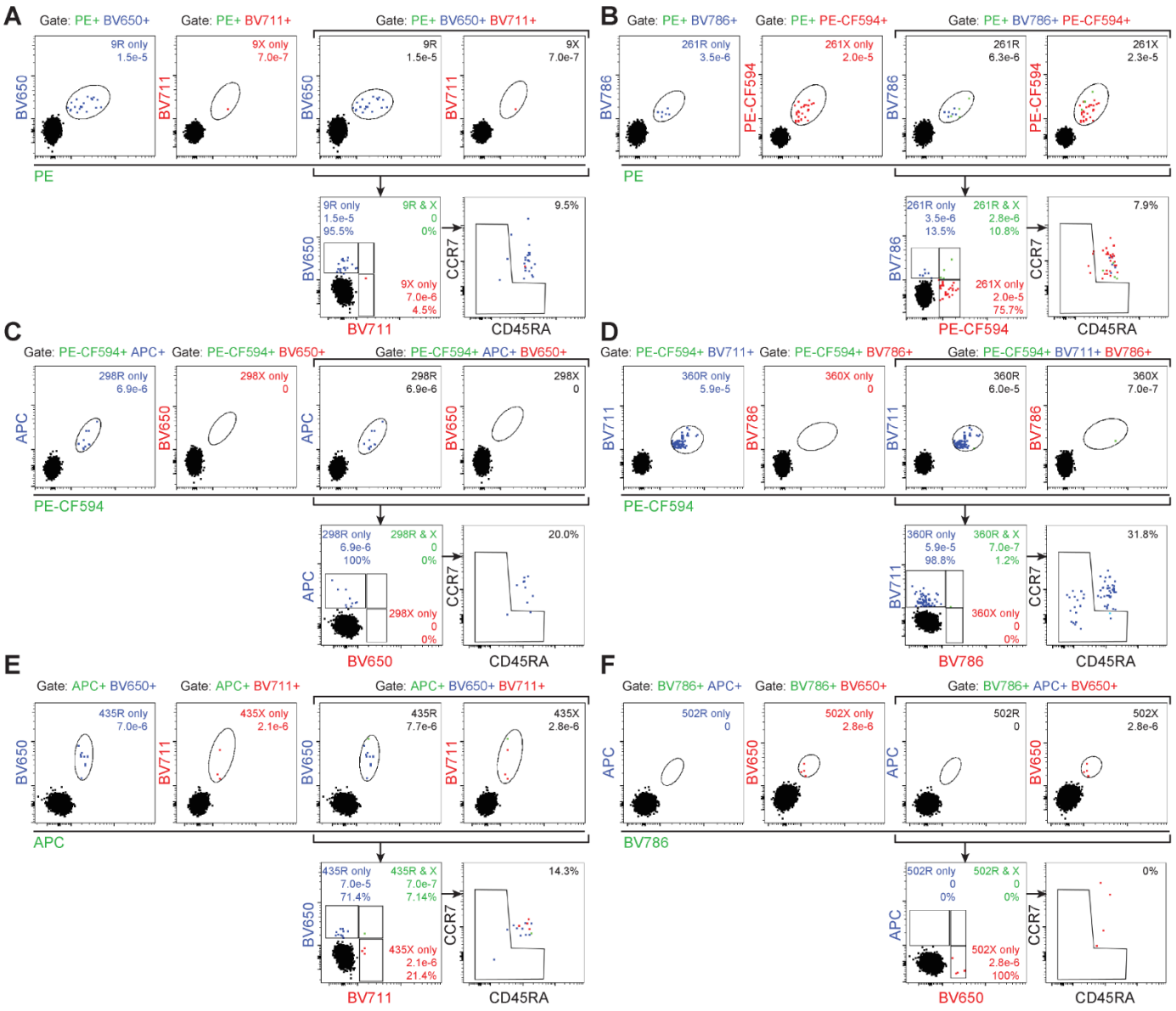
Frozen thymic tissue sections from 6-8-week-old WT and *Aire*$^{-/-}$ C57BL/6 mice and from NOD mice were fixed with 2% paraformaldehyde (Sigma) and incubated for 10 min in a saturation buffer composed of 0.1 M Tris pH 7.4, 5% mouse serum, 2% bovine serum albumin (Axday), 0.01% Triton X-100. Sections were then stained as previously described (11), with primary antibodies: rabbit anti-mouse keratin-14 (RRID:AB_2565048), mouse anti-PADI2 (RRID: AB_2881762), AF488-labeled rat anti-mouse Aire (RRID:AB_10854132); and secondary antibodies: Cy5-labeled goat anti-rabbit IgG (RRID: AB_2534032) and Cy3-labeled goat anti-mouse IgG (RRID:AB_893530). Sections were mounted with Mowiol (Calbiochem) and acquired on a LSM 780 confocal microscope (Carl Zeiss Microscopy).

**Statistics**

Significance was assessed with a cutoff value of $\alpha=0.05$ using two-tailed Student t test or Mann-Whitney U test according to normal or non-normal distribution, respectively. Variables displaying unequal variances were analyzed with the Welch t test. Paired datasets were analyzed with the Wilcoxon signed rank test.




**Supplementary Figure 1. Gating strategy for the combinatorial analysis of GRP78 MMr$^+$CD8$^+$ T cells in T1D and healthy donors. (A)** Frozen-thawed PBMCs from healthy donor C59 were magnetically depleted of CD8$^-$ cells before staining, acquisition and analysis. Cells were sequentially gated on small lymphocytes, singlets, live cells (Live/Dead Aqua$^-$), CD3$^+$CD8$^+$ T cells, and total PE$^+$, PE-CF594$^+$, APC$^+$, BV650$^+$, BV711$^+$, and BV786$^+$ MMr$^+$ T cells. Using Boolean operators, these latter gates allowed selective visualization of each double-MMr$^+$ population by including only those events positive for the corresponding fluorochrome pair (see Research Design and Methods). **(B)** The final readout obtained for the 15 peptides analyzed: 12 native/citrullinated GRP78 peptides, a GRP78$_{152\text{-}161}$ (GRP78-152) peptide without R residues, and PPI$_{6\text{-}14}$ and Flu-MP$_{58\text{-}66}$ controls). Events corresponding to each epitope-reactive population are overlaid in different colors within each plot, with MMr$^-$ events overlaid in light gray. The small dot plots on the right of each panel depict CD45RA (x axis) and CCR7 (y axis) expression in the corresponding MMr$^+$ fraction. Numbers in each panel indicate the MMr$^+$CD8$^+$ T-cell frequency out of total CD8$^+$ T cells and the percent naive (CD45RA$^+$CCR7$^+$) fraction among MMr$^+$ cells. This first gating strategy only allows to visualize CD8$^+$ T cells reactive to either the R or X isoform of each GRP78 peptides, and these T cells are therefore designated as 'R only' or 'X only', respectively. This strategy was therefore further modified to differentially visualize CD8$^+$ T cells reactive to R only, X only or both peptides (see Supplemental Figure 2).

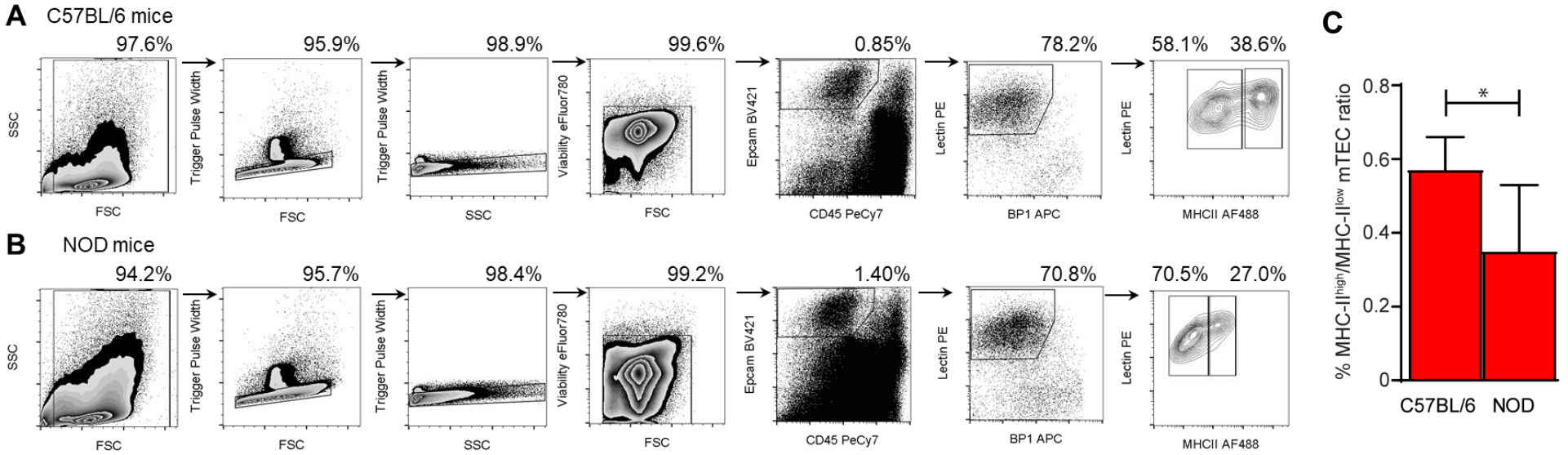



**Supplementary Figure 2. Modified combinatorial gating strategy to analyze the cross-reactivity between R-GRP78 and X-GRP78 peptide isoforms.** A representative example is shown for healthy donor C59. Each panel shows one of the peptides studied: GRP78-9 **(A)**, -261 **(B)**, -298 **(C)**, -360 **(D)**, -435 **(E)**, and -502 **(F)**.

First, $CD8^+$ T cells recognizing only the R-GRP78 (blue, 'R only') or X-GRP78 isoform (red, 'X only') were gated on double $MMr^+$ events, excluding the other four fluorochromes (first two dot plots of each panel, which are the same of Supplemental Fig. 1; $MMr^-$ events overlaid in black). For example, in panel B '261R only' $MMr^+$ cells ($PE^+BV786^+$) were visualized by gating on $PE^+BV786^+PE\text{-}CF594^-APC^-BV650^-BV711^-$ events; and '261X only' $MMr^+$ cells ($PE^+PE\text{-}CF594^+$) were visualized by gating on $PE^+PE\text{-}CF594^+APC^-BV650^-BV711^-BV786^-$ events.

Second, $CD8^+$ T cells recognizing R-GRP78 or X-GRP78 peptides (i.e. including cross-reactive T cells) were gated on triple $MMr^+$ events, i.e. by excluding three fluorochromes but not the non-shared fluorochrome coding for the MMr loaded with the other peptide isoform (third and fourth dot plots of each panel). For example, in panel B '261R' and '261X' $MMr^+$ cells (i.e. cross-reactive or not; $PE^+BV786^+$ and $PE^+PE\text{-}$

CF594$^+$, respectively) were visualized by gating on PE$^+$BV786$^+$APC$^-$BV650$^-$BV711$^-$ (i.e. not excluding PE-CF594 as before) and PE$^+$PE-CF594$^+$APC$^-$BV650$^-$BV711$^-$ (i.e. not excluding BV786), respectively.
Third, these two dot plots are combined and visualized by plotting them for the two fluorochromes not shared by the R and X isoform (first dot plot in the bottom line of each panel). In this final dot plot, single-positive events correspond to the 'R only' and 'X only' fractions (blue and red, respectively), while double-positive events correspond to the 'R and X' fraction (i.e. cross-reactive; green). Events negative for all MMr fluorochromes (PE$^-$PE-CF594$^-$APC$^-$BV650$^-$BV711$^-$BV786$^-$) are overlaid in black in each dot plot to set the double- or triple-MMr$^+$ gates. The bottom right dot plots of each panel depict CD45RA and CCR7 expression in the corresponding MMr$^+$ fractions using the same blue, red and green color code for 'R only', 'X only' and 'R and X' MMr$^+$ events, respectively.

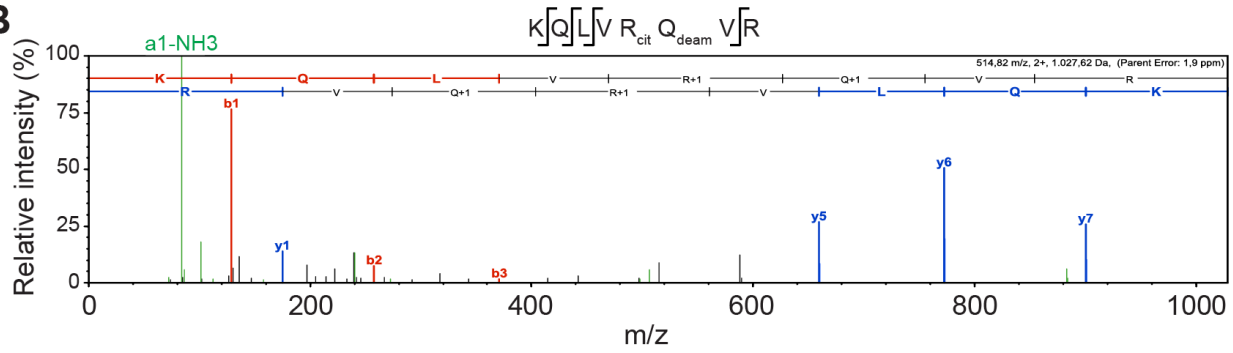



**Supplementary Figure 3. Gating strategy and distribution of MHC-II$^{low}$ and MHC-II$^{high}$ mTECs in C57BL/6 and NOD mice. (A)** mTECs from C57BL/6 mice. **(B)** mTECs from NOD mice. mTECs were gated as CD45$^{-}$ EPCAM$^{+}$ BP1$^{neg/low}$ Lectin (UEA-1)$^{high}$ and MHC-II$^{low}$ or MHC-II$^{high}$. **(C)** Ratio (median ± range) of MHC-II$^{high}$/MHC-II$^{low}$ mTECs in 6-8-week-old C57BL/6 (n=4) and NOD mice (n=5); **p*=0.032 by Mann-Whitney U test.

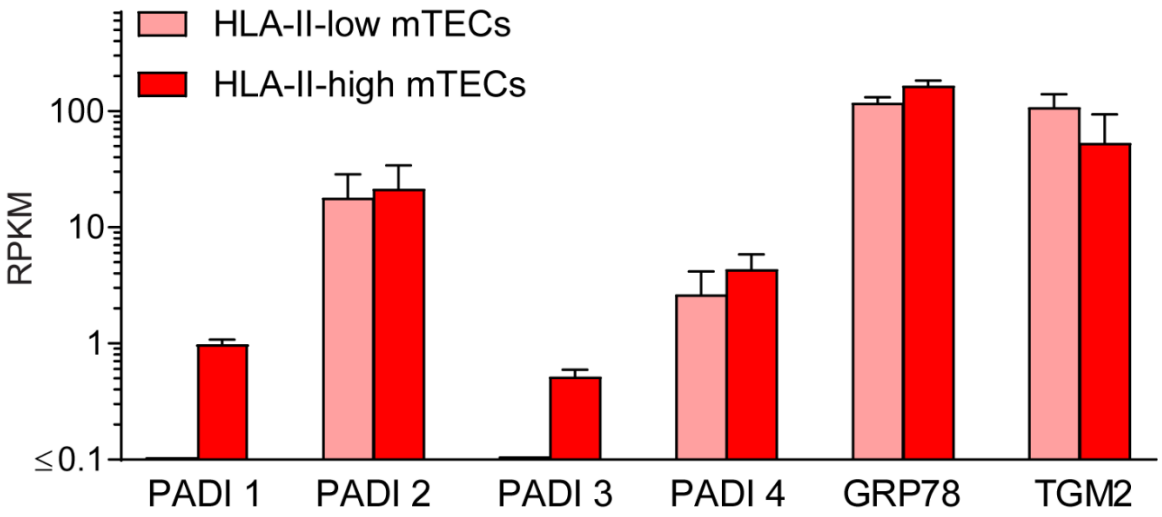
**A**

| Name | Sequence | Modifications |
|---|---|---|
| EF-hand domain-containing protein 1 | NGYAVVRRPTMGIGGDR | 1xCitrullinated [R7]; 1xOxidation [M11] |
| Hepatocyte growth factor-regulated tyrosine kinase substrate | AAEEAERQR | 1xCitrullinated [R7]; 1xDeamidated [Q8] |
| Myosin-9 | KQLVRQVR | 1xCitrullinated [R5]1xDeamidated [Q6] |

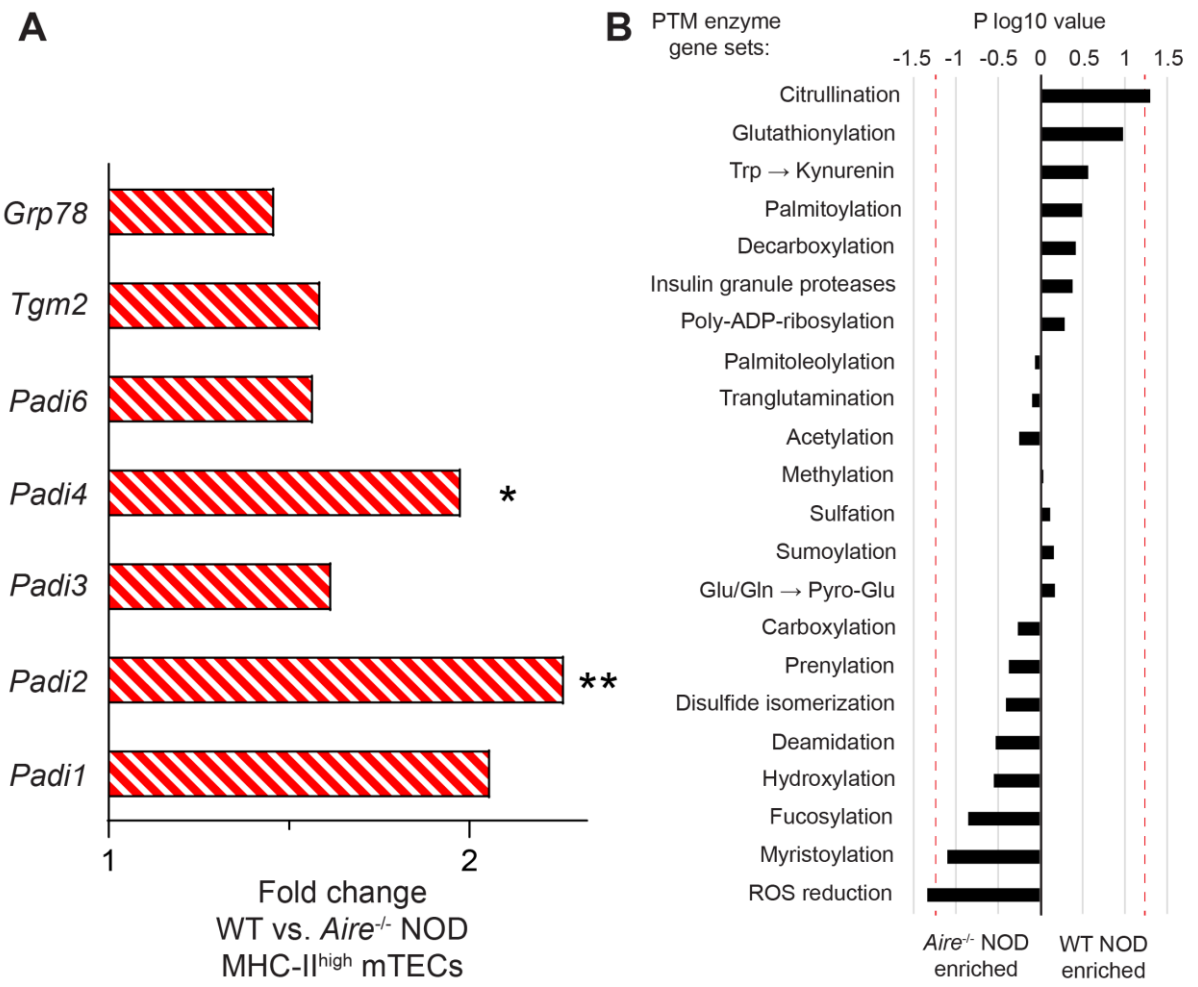
**B**



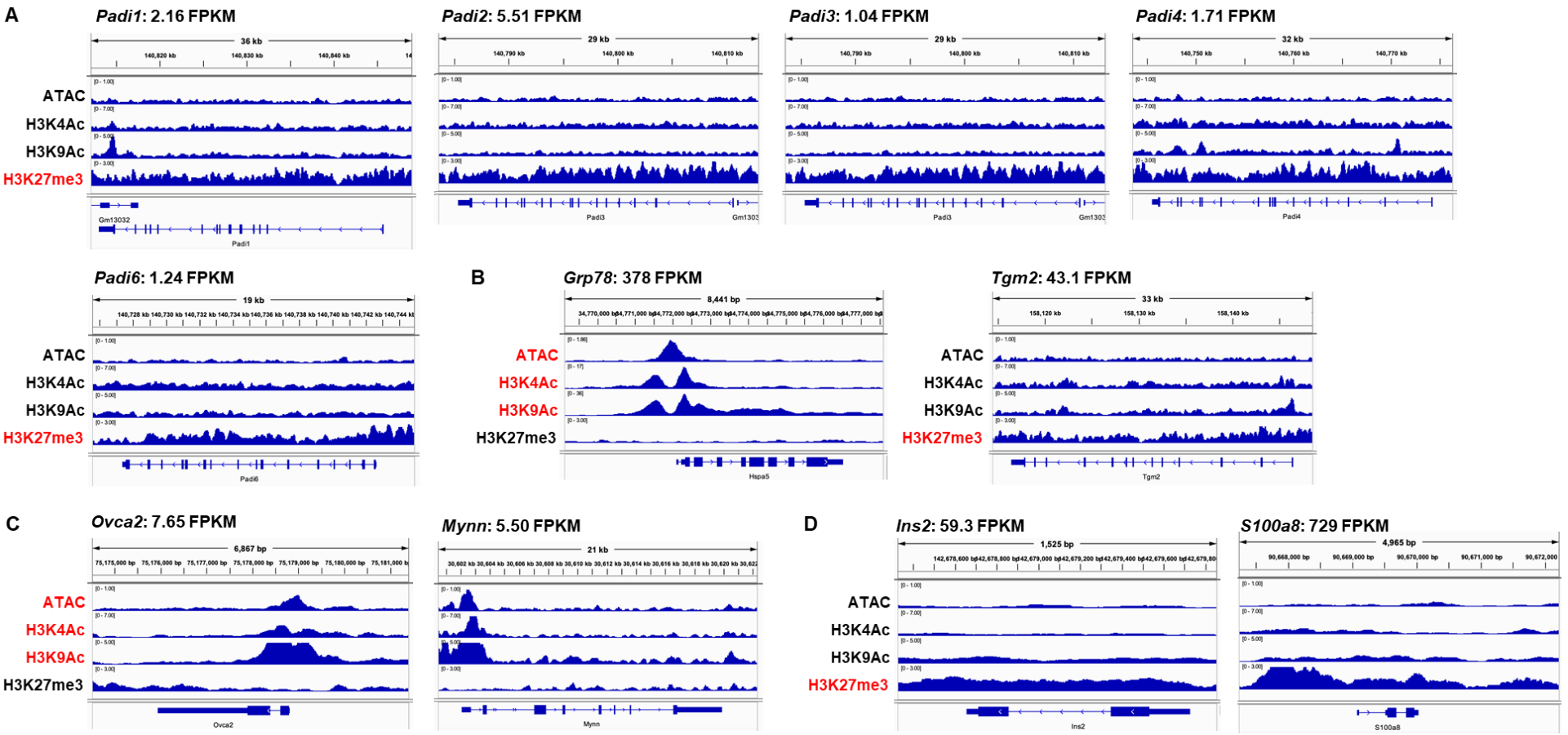

**Supplementary Figure 4. Citrullinated peptides identified in mTECs of 10-week-old C57BL/6 mice. (A)** List of the citrullinated peptides identified. **(B)** Spectrum of the R5-citrullinated myosin-9 peptide KQLVRQVR.



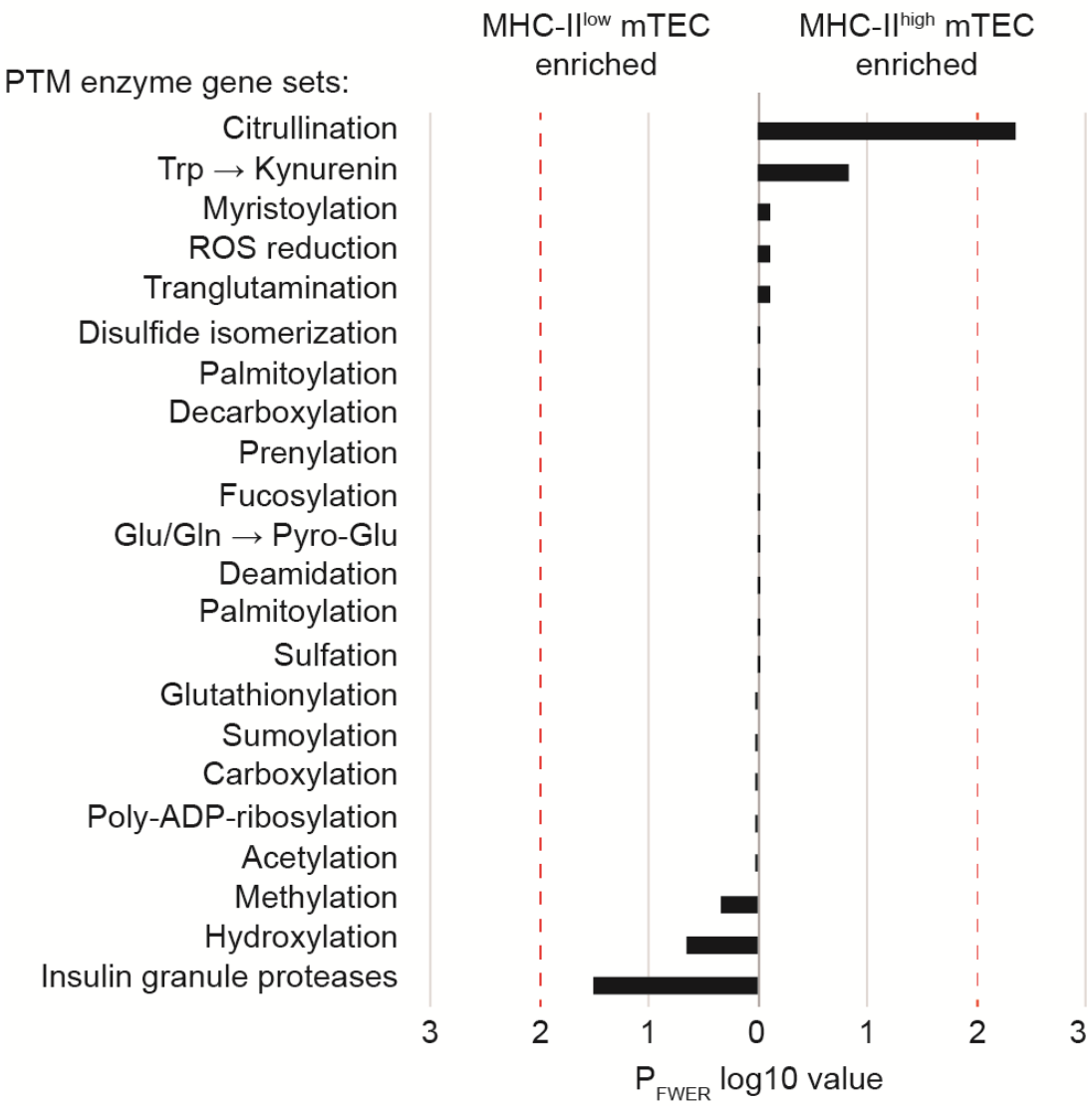

**Supplementary Figure 5. Gene expression of PADI citrullinating enzymes in human immature HLA Class II (HLA-II)$^{low}$ and mature HLA-II$^{high}$ mTECs (n=3 donors).** Data was extracted from our published RNAseq datasets (12).




**Supplementary Figure 6. Enrichment of *Padi* isoform genes and PTM enzyme gene sets in mature MHC-II$^{high}$ mTECs from WT vs. *Aire*$^{-/-}$ NOD mice (n=3/each). (A)** Fold change expression of *Padi* isoform genes in MHC-II$^{high}$ mTECs from 4-week-old WT vs. *Aire*$^{-/-}$ NOD mice analyzed by DNA microarray. **p*=0.04, ***p*=0.01 by false discovery rate method (GEO2R). **(B)** Enrichment of PTM enzyme gene sets in MHC-II$^{high}$ mTECs from WT vs *Aire*$^{-/-}$ NOD mice. Red lines indicate the 0.05 *p* cut-off value for statistical significance [for citrullination, *p*=0.05, corresponding to a WT normalized enrichment score (NES) of 1.5]. Data is from GEO GSE12073 (8).




**Supplementary Figure 7. ATAC-Seq and Chip-seq profiles of *Padi* genes in MHC-II$^{high}$ mTECs from C57BL/6 mice.** ATAC-seq and Chip-seq datasets (GSE114713) from Handel et al. (13) were analyzed. **(A)** *Padi* genes display three hallmarks of Aire-dependent expression: a closed chromatin ATAC profile (first row), near-absence of the activating chromatin modifications H3K4Ac and H3K9Ac (second and third row) and presence of the repressive modification H3K27me3 throughout the gene locus (fourth raw). **(B)** Conversely, the *Grp78* gene displays an open chromatin ATAC profile, presence of the activating chromatin modifications H3K4Ac and H3K9Ac at transcriptional start sites and absence of the repressive modification H3K27me3, in line with an Aire-independent expression. The *Tgm2* gene displays an intermediate profile. **(C-D)** Profiles of reference Aire-independent (C) and Aire-dependent genes (D) are shown as controls.




**Supplementary Figure 8. Enrichment of PTM enzyme gene sets in immature MHC-II$^{low}$ *vs* mature MHC-II$^{high}$ mTECs from WT C57BL/6 mice (n=3).** Red lines indicate the 0.01 $P_{FWER}$ cut-off value for statistical significance (for citrullination, $P_{FWER}$=0.0045, corresponding to a MHC-II$^{high}$ mTEC normalized enrichment score of 2.3).

| | PE | PE-CF594 | APC | BV650 | BV711 | BV786 |
|---|---|---|---|---|---|---|
| **PE** | | 261X | PPI6-14 | 9R | 9X | 261R |
| **PE-CF594** | | | 298R | 298X | 360R | 360X |
| **APC** | | | | 435R | 435X | 502R |
| **BV650** | | | | | 152 | 502X |
| **BV711** | | | | | | Flu MP58-66 |
| **BV786** | | | | | | |

**Supplementary Table 1.** Combinatorial HLA-A2 MMr panel used in Figure 2. See Supplementary Figures 1-2 for details.

| | nPOD case | RRID | Sex | Age (yrs) | T1D T2D (yrs) | Positive aAbs | C-peptide (ng/ml) | Pancreas MMr+ cells | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | GRP78 9R | GRP78 9X | GRP78 360R | GRP78 360X | ZnT8 186-194 | UCN3 1-9 | MelA 26-35 |
| T1D (n=10) | 6070* | SAMN15879127 | F | 23 | 7 | IA-2/mIAA | <0.05 | 0 | 16 | | 53 | 74 | 39 | 0 |
| | 6113* | SAMN15879170 | F | 13 | 2 | mIAA | <0.05 | 0 | 0 | | 0 | | | |
| | 6161* | SAMN15879217 | F | 19 | 7 | IA-2/mIAA | <0.05 | 12 | 45 | | | 124 | | 0 |
| | 6211* | SAMN15879267 | F | 24 | 4 | GAD/IA-2/ZnT8/mIAA | <0.05 | 3 | 4 | | 0 | 30 | | 0 |
| | 6212 | SAMN15879268 | M | 20 | 5 | mIAA | <0.05 | 0 | 0 | | 0 | 0 | | |
| | 6237 | SAMN15879293 | F | 18 | 12 | GAD/mIAA | <0.05 | 7 | 0 | | 0 | 267 | 8 | 0 |
| | 6242 | SAMN15879298 | M | 39 | 19 | IA-2/mIAA | <0.05 | 0 | 0 | 0 | | 66 | | 0 |
| | 6258* | SAMN15879313 | F | 39 | 37 | mIAA | <0.05 | 0 | 4 | 0 | | 118 | 19 | 0 |
| | 6325* | SAMN15879379 | F | 20 | 6 | GAD/IA-2 | 0.14 | 13 | 19 | 0 | | 28 | 28 | 0 |
| | 060217 | NA | F | 39 | 21 | GAD | NA | 0 | 41 | 0 | | | 142 | |
| aAb+ (n=9) | 6080 | SAMN15879137 | F | 69 | NA | GAD/mIAA | 1.84 | 0 | 0 | | 0 | 55 | | 25 |
| | 6101 | SAMN15879158 | M | 65 | NA | GAD | 26.18 | 0 | 0 | | 0 | 0 | | |
| | 6123 | SAMN15879180 | F | 23 | NA | GAD | 2.01 | 0 | | | 0 | 0 | | |
| | 6151 | SAMN15879207 | M | 30 | NA | GAD | 5.49 | 0 | 5 | | 0 | 28 | 9 | 0 |
| | 6154 | SAMN15879210 | F | 49 | NA | GAD | <0.05 | 0 | 0 | 0 | | 64 | 21 | 0 |
| | 6171 | SAMN15879227 | F | 4 | NA | GAD | 8.95 | 0 | 7 | | 8 | 37 | | 0 |
| | 6347 | SAMN15879401 | M | 9 | NA | mIAA | 3.26 | 6 | 49 | 7 | | 33 | | 0 |
| | 6388 | SAMN15879441 | F | 25 | NA | GAD/Miaa | 1.38 | 0 | 7 | 0 | | 34 | 0 | 0 |
| | 6397 | SAMN15879450 | F | 21 | NA | GAD | 12.8 | 4 | 4 | 0 | | 42 | 6 | 0 |
| No diabetes (n=11) | 6103 | SAMN15879160 | M | 2 | NA | — | 0.98 | 30 | 0 | | | 55 | 0 | 0 |
| | 6174 | SAMN15879230 | M | 21 | NA | — | 3.00 | 4 | 0 | | 4 | | | |
| | 6179 | SAMN15879235 | F | 20 | NA | — | 2.74 | 0 | 7 | | 0 | 96 | 0 | 0 |
| | 6182 | SAMN15879238 | M | 3 | NA | — | 2.28 | 0 | 4 | | 0 | 23 | 4 | 0 |
| | 6227 | SAMN15879283 | F | 17 | NA | — | 2.75 | 5 | 12 | | 0 | 3 | 18 | 0 |
| | 6234 | SAMN15879290 | F | 20 | NA | — | 6.89 | 0 | 0 | 0 | | 6 | 0 | 0 |
| | 6271 | SAMN15879325 | M | 17 | NA | — | 11.47 | 0 | 0 | 0 | | 0 | | |
| | 6287 | SAMN15879341 | F | 57 | NA | — | 4.75 | 0 | 0 | 0 | | 4 | | 0 |
| | 6289 | SAMN15879343 | M | 19 | NA | — | 8.05 | 0 | 0 | 0 | | 0 | | |
| | 6357 | SAMN15879410 | M | 5 | NA | — | 8.82 | | 128 | | | 0 | | |
| | 6366 | SAMN15879419 | F | 21 | NA | — | 0.41 | 0 | 35 | | 0 | 0 | | |
| T2D/pancreatitis (n (n=4+1) | 6028 | SAMN15879085 | M | 33 | 17 | — | 22.40 | | 0 | | | 0 | | |
| | 6059 | SAMN15879116 | F | 19 | 0.3 | — | 10.68 | 0 | 0 | | | 0 | | |
| | 6273 | SAMN15879327 | F | 45 | 2 | — | 3.17 | 3 | 7 | | | | | |
| | 6275 | SAMN15879329 | M | 48 | 2 | — | 3.46 | 0 | 0 | | | 0 | | |
| | | | | | | | | | | | | | | |
| | 6439 | SAMN15879492 | M | 27 | NA | — | 5.33 | 5 | 22 | | | | | |

**Supplementary Table 2. nPOD cases analyzed by *in-situ* tissue MMr staining.** The clinical characteristics of each case are reported along with the counts ($x10^{-3}$) of $MMr^{+}$ cells/$mm^2$ pancreas section area for each of the indicated peptides. Positive sections are marked in red. Case #6287 (presenting a circumscribed neuroendocrine tumor in the pancreatic pan-body region; pan-tail region analyzed here) was classified as a non-diabetic control. T1D case #060217 is from the EUnPOD biobank. Control duodenum sections were also tested for T1D cases marked with an asterisk. NA, not applicable or not available; mIAA, micro-insulin aAbs. Results for epitopes other than GRP78 are from (12, 14).

| PTM | Mouse gene | Mouse UniProt # | Human Uniprot # |
|---|---|---|---|
| Acetylation | Kat2a | Q9JHD2 | Q92830 |
| | Naa10 | Q9QY36 | P41227 |
| | Naa11 | Q3UX61 | Q9BSU3 |
| | Naa15 | Q80UM3 | Q9BXJ9 |
| | Naa16 | Q9DBB4 | Q6N069 |
| | Naa20 | P61600 | P61599 |
| | Naa25 | Q8BWZ3 | Q14CX7 |
| | Naa30 | Q8CES0 | Q147X3 |
| | Naa35 | Q6PHQ8 | Q5VZE5 |
| | Naa38 | Q9D2U5 | Q9BRA0 |
| | Naa50 | Q6PGB6 | Q9GZZ1 |
| | Naa60 | Q9DBU2 | Q9H7X0 |
| | Nat8 | Q9JIY7 | Q9UHE5 |
| | Nat9 | Q3UG98 | Q9BTE0 |
| | Sirt1 | Q923E4 | Q96EB6 |
| | Sirt2 | Q8VDQ8 | Q8IXJ6 |
| | Sirt7 | Q8BKJ9 | Q9NRC8 |
| Carboxylation | Ggcx | Q9QYC7 | A7YA96 |
| Cisteinylation (non-enzymatic) | n/a | | |
| Citrullination | Padi1 | Q9Z185 | Q9ULC6 |
| | Padi2 | Q08642 | Q9Y2J8 |
| | Padi3 | Q9Z184 | Q9ULW8 |
| | Padi4 | Q9Z183 | Q9UM07 |
| | Padi6 | Q8K3V4 | Q6TGC4 |
| Deamidation | Ntan1 | Q64311 | Q96AB6 |
| | Ntaq1 | Q80WB5 | Q96HA8 |
| Decarboxylation | Gad1 | P48318 | Q99259 |
| | Gad2 | P48320 | Q05329 |
| | Gadl1 | Q80WP8 | Q6ZQY3 |
| Fucosylation | Fut1 | O09160 | P19526 |
| | Fut10 (Sec1) | Q5F2L2 | Q6P4F1 |
| | Fut11 | Q8BHC9 | Q495W5 |
| | Fut2 (Sec2) | Q9JL27 | Q10981 |
| | Fut3 | n/a | P21217 |
| | Fut4 | Q11127 | P22083 |
| | Fut7 | Q11131 | Q11130 |
| | Fut8 | Q9WTS2 | Q9BYC5 |
| | Fut9 | O88819 | Q9Y231 |
| | Pofut1 | Q91ZW2 | Q9H488 |
| | Pofut2 | Q8VHI3 | Q9Y2G5 |
| Glutathionylation | Ggt1* | Q60928 | P19440 |
| | Ggt3p* | n/a | A6NGU5 |
| | Ggt5* | Q9Z2A9 | P36269 |
| | Ggt6 | Q6PDE7 | Q6P531 |
| | Ggt7 | Q99JP7 | Q9UJ14 |
| | Gsta1 | P13745 | P08263 |
| | Gsta2 | P10648 | P09210 |
| | Gsta3 | P30115 | Q16772 |
| | Gsta4 | P24472 | O15217 |
| | Gsta5 | n/a | Q7RTV2 |
| | Gstk1 | Q9DCM2 | Q9Y2Q3 |
| | Gstm1 | P10649 | P09488 |
| | Gstm2 | P15626 | P28161 |
| | Gstm3 | P19639 | P21266 |
| | Gstm4 | Q8R5I6 | Q03013 |
| | Gstm5 | P48774 | P46439 |
| | Gstm6 | O35660 | n/a |
| | Gstm7 | Q80W21 | n/a |
| | Gsto1 | O09131 | P78417 |
| | Gsto2 | Q8K2Q2 | Q9H4Y5 |
| | Gstp1 | P19157 | P09211 |
| | Gstp2 | P46425 | n/a |
| | Gstp3 | Q8VC73 | n/a |
| | Gstt1 | Q64471 | P30711 |
| Glutathionylation | Gstt2 | Q61133 | P0CG29 |
| | Gstt2b | n/a | P0CG30 |
| | Gstt3 | Q99L20 | n/a |
| | Gstt4 | Q9D4P7 | A0A1W2PR19 |
| | Lancl1 | O89112 | O43813 |
| | Mgst1 | Q91VS7 | P10620 |
| | Mgst2 | A2RST1 | Q99735 |
| | Mgst3 | Q9CPU4 | O14880 |
| Hydroxylation | Asph | Q8BSY0 | Q12797 |
| | Egln1 | Q91YE3 | Q9GZT9 |
| | Egln2 | Q91YE2 | Q96KS0 |
| | Egln3 | Q91UZ4 | Q9H6Z9 |
| | Jmjd6 | Q9ERI5 | Q6NYC1 |
| | Kdm8 | Q9CXT6 | Q8N371 |
| Malonylation (non-enzymatic) | n/a | | |
| Methylation | Antkmt | Q501J2 | Q9BQD7 |
| | Atpsckmt | Q9D1Z3 | Q6P4H8 |
| | Camkmt | Q3U2J5 | Q7Z624 |
| | Ehmt2 | Q9Z148 | Q96KQ7 |
| | Icmt | Q9EQK7 | O60725 |
| | Kmt5a | Q2YDW7 | Q9NQR1 |
| | Lcmt1 | A2RTH5 | Q9UIC8 |
| | Lcmt2 | Q8BYR1 | O60294 |
| | Mettl11b | B2RXM4 | Q5VVY1 |
| | Mettl18 | Q9CZ09 | O95568 |
| | Mettl21a | Q9CQL0 | Q8WXB1 |
| | Mettl21c | Q8BLU2 | Q5VZV1 |
| | Mettl21e | Q8CDZ2 | n/a |
| | Mettl22 | Q8R1C6 | Q9BUU2 |
| | N6amt1 | Q6SKR2 | Q9Y5N5 |
| | Ndufaf7 | Q9CWG8 | Q7L592 |
| | Ntmt1 (Mettl11a) | Q8R2U4 | Q9BV86 |
| | Pcmt1* | P23506 | P22061 |
| | Pcmtd1* | P59913 | Q96MG8 |
| | Prmt1 | Q9JIF0 | Q99873 |
| | Prmt2 | Q9R144 | P55345 |
| | Prmt3 | Q922H1 | O60678 |
| | Prmt4 (Carm1) | Q9WVG6 | Q86X55 |
| | Prmt5 (Anm5) | Q8CIG8 | O14744 |
| | Prmt6 (Anm6) | Q6NZB1 | Q96LA8 |
| | Prmt7 (Anm7) | Q922X9 | Q9NVM4 |
| | Prmt8 (Anm8) | Q6PAK3 | Q9NR22 |
| | Prmt9 (Anm9) | Q3U3W5 | Q6P2P2 |
| | Setd2 | E9Q5F9 | Q9BYW2 |
| | Setd4 | P58467 | Q9NVD3 |
| | Setd7 | Q8VHL1 | Q8WTS6 |

*continued*

**Supplementary Table 3. Mouse genes coding for enzymes involved in PTMs.** This list was used for the analysis of gene sets presented in Fig. 5G and Supplemental Fig. 5. PTMs involving very large sets of enzymes were excluded, namely glycosylation, phosphorylation, oxidation and ubiquitination. This list does not include enzymes described to selectively modify specific proteins or histones. Asterisks indicate enzymes indirectly involved in PTMs.

| PTM | Mouse gene | Mouse UniProt # | Human Uniprot # |
|---|---|---|---|
| **Myristoylation** | Nmt1 | O70310 | P30419 |
| | Nmt2 | O70311 | O60551 |
| **Nitration (non-enzymatic)** | n/a | | |
| **Nitrosylation (non-enzymatic)** | n/a | | |
| **Palmitoylation** | Zdhhc1 | Q8R0N9 | Q8WTX9 |
| | Zdhhc2 | P59267 | Q9UIJ5 |
| | Zdhhc3 | Q8R173 | Q9NYG2 |
| | Zdhhc4 | Q9D6H5 | Q9NPG8 |
| | Zdhhc5 | Q8VDZ4 | Q9C0B5 |
| | Zdhhc6 | Q9CPV7 | Q9H6R6 |
| | Zdhhc7 | Q91WU6 | Q9NXF8 |
| | Zdhhc8 | Q5Y5T5 | Q9ULC8 |
| | Zdhhc9 | P59268 | Q9Y397 |
| | Zdhhc11 | Q14AK4 | Q9H8X9 |
| | Zdhhc11b | n/a | P0C7U3 |
| | Zdhhc12 | Q8VC90 | Q96GR4 |
| | Zdhhc13 | Q9CWU2 | Q8IUH4 |
| | Zdhhc14 | Q8BQQ1 | Q8IZN3 |
| | Zdhhc15 | Q8BGJ0 | Q96MV8 |
| | Zdhhc16 | Q969W1 | Q969W1 |
| | Zdhhc17 | Q80TN5 | Q8IUH5 |
| | Zdhhc18 | Q5Y5T2 | Q9NUE0 |
| | Zdhhc19 | Q810M5 | Q8WVZ1 |
| | Zdhhc20 | Q5Y5T1 | Q5W0Z9 |
| | Zdhhc21 | Q9D270 | Q8IVQ6 |
| | Zdhhc22 | A0PK84 | Q8N966 |
| | Zdhhc23 | Q5Y5T3 | Q8IYP9 |
| | Zdhhc24 | Q6IR37 | Q6UX98 |
| **Palmitoleoylation** | Porcn | Q9JJJ7 | Q9H237 |
| | Notum | Q8R116 | Q6P988 |
| **Poly-ADP-ribosylation** | Parp1 | P11103 | P09874 |
| | Parp2 | O88554 | Q9UGN5 |
| | Parp3 | Q3ULW8 | Q9Y6F1 |
| | Parp4 | E9PYK3 | Q9UKK3 |
| | Parp6 | Q6P6P7 | Q2NL67 |
| | Parp8 | Q3UD82 | Q8N3A8 |
| | Parp9 | Q8CAS9 | Q8IXQ6 |
| | Parp10 | Q8CIE4 | Q53GL7 |
| | Parp11 | Q8CFF0 | Q9NR21 |
| | Parp12 | Q8BZ20 | Q9H0J9 |
| | Parp14 | Q2EMV9 | Q460N5 |
| | PARP15 | n/a | Q460N3 |
| | Parp16 | Q7TMM8 | Q8N5Y8 |
| | Tiparp (Parp7) | Q8C1B2 | Q7Z3E1 |
| | Tnks (Parp5a) | Q6PFX9 | O95271 |
| | Tnks2 (Parp5b) | Q3UES3 | Q9H2K2 |
| **Prenylation** | Fnta | Q61239 | P49354 |
| | Fntb | Q8K2I1 | P49356 |
| | Pggt1b | A0A494BAX1 | P53609 |
| | Rce1 | P57791 | Q9Y256 |

| PTM | Mouse gene | Mouse UniProt # | Human Uniprot # |
|---|---|---|---|
| **Glu/Gln → Pyro-Glu** | Qpct | Q9CYK2 | Q16769 |
| | Qpctl | Q8BH73 | Q9NXS2 |
| **Disulfide isomerization** | Creld1 | Q91XD7 | Q96HD1 |
| | Creld2 | Q9CYA0 | Q6UXH1 |
| | Ero1a* | Q8R180 | Q96HE7 |
| | Ero1b* | Q8R2E9 | Q86YB8 |
| | P4hb (Pdia1) | P09103 | P07237 |
| | Pdia2 | D3Z6P0 | Q13087 |
| | Pdia3 | P27773 | P30101 |
| | Pdia4 | P08003 | P13667 |
| | Pdia5 | Q921X9 | Q14554 |
| | Pdia6 | Q922R8 | Q15084 |
| | Qsox1 | Q8BND5 | O00391 |
| | Qsox2 | Q3TMX7 | Q6ZRP7 |
| | Tmx3 | Q8BXZ1 | Q96JJ7 |
| **ROS reduction** | Cat (Cata)* | P24270 | P04040 |
| | Cybb* | Q61093 | P04839 |
| | Mpo* | P11247 | P05164 |
| | Nos1* | Q9Z0J4 | C9J5P6 |
| | Nos2* | P29477 | P35228 |
| | Sod1* | P08228 | P00441 |
| | Sod2* | P09671 | P04179 |
| | Sod3* | O09164 | P08924 |
| **Sulfation** | Tpst1 | O70281 | O60507 |
| | Tpst2 | O88856 | O60704 |
| **Sulfone (non-enzymatic)** | n/a | | |
| **Sumoylation** | Sae1 | Q9R1T2 | Q9UBE0 |
| | Uba2 (Sae2) | Q9Z1F9 | Q9UBT2 |
| **Transglutamination** | Tgm1 | Q9JLF6 | P22735 |
| | Tgm2 | P21981 | P21980 |
| | Tgm3 | Q08189 | Q08188 |
| | Tgm4 | Q8BZH1 | P49221 |
| | Tgm5 | Q9D7I9 | O43548 |
| | Tgm6 (Tgm3l) | Q8BM11 | O95932 |
| | Tgm7 | A2ART8 | Q96PF1 |
| **Trp → Kynurenine** | Afmid | Q8K4H1 | Q63HM1 |
| | Ido1 | P28776 | P14902 |
| | Ido2 | Q8R0V5 | Q6ZQW0 |
| | Tdo2 (T23o) | P48776 | P48775 |
| **Insulin granule proteases** | Cpe | Q00493 | P16870 |
| | Cpn1 | Q9JJN5 | P15169 |
| | Cpn2 | Q9DBB9 | P22792 |
| | Ctsb | P10605 | P07858 |
| | Ctsd | P18242 | P07339 |
| | Ctsf | Q9R013 | Q9UBX1 |
| | Ctsl | P06797 | P07711 |
| | Dpp7 | Q9ET22 | Q9UHL4 |
| | Furin | P23188 | P09958 |
| | Pcsk1 | P63239 | P29120 |
| | Pcsk2 | P21661 | P16519 |
| | Tpp1 | O89023 | O14773 |

## SUPPLEMENTARY REFERENCES


1. Buitinga M, Callebaut A, Marques Camara Sodre F, Crevecoeur I, Blahnik-Fagan G, Yang ML, Bugliani M, Arribas-Layton D, Marre M, Cook DP, Waelkens E, Mallone R, Piganelli JD, Marchetti P, Mamula MJ, Derua R, James EA, Mathieu C, Overbergh L: Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. Diabetes 2018;67:2337-2348
2. Callebaut A, Derua R, Vig S, Delong T, Mathieu C, Overbergh L: Identification of Deamidated Peptides in Cytokine-Exposed MIN6 Cells through LC-MS/MS Using a Shortened Digestion Time and Inspection of MS2 Spectra. J Proteome Res 2021;20:1405-1414
3. Guyon C, Jmari N, Padonou F, Li YC, Ucar O, Fujikado N, Coulpier F, Blanchet C, Root DE, Giraud M: Aire-dependent genes undergo Clp1-mediated 3'UTR shortening associated with higher transcript stability in the thymus. eLife 2020;9
4. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25
5. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-842
6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545-15550
7. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol 2010;11:R106
8. Guerau-de-Arellano M, Martinic M, Benoist C, Mathis D: Neonatal tolerance revisited: a perinatal window for Aire control of autoimmunity. J Exp Med 2009;206:1245-1252
9. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP: GenePattern 2.0. Nat Genet 2006;38:500-501
10. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-264
11. Serge A, Bailly AL, Aurrand-Lions M, Imhof BA, Irla M: For3D: Full organ reconstruction in 3D, an automatized tool for deciphering the complexity of lymphoid organs. J Immunol Methods 2015;424:32-42
12. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, Lalanne AI, Sebastiani G, Carre A, Pinto S, Culina S, Corcos N, Bugliani M, Marchetti P, Armanet M, Diedisheim M, Kyewski B, Steinmetz LM, Buus S, You S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Dotta F, Scharfmann R, Eizirik DL, Verdier Y, Vinh J, Mallone R: Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 2018;28:946-960 e946
13. Handel AE, Shikama-Dorn N, Zhanybekova S, Maio S, Graedel AN, Zuklys S, Ponting CP, Hollander GA: Comprehensively Profiling the Chromatin Architecture of Tissue Restricted Antigen Expression in Thymic Epithelial Cells Over Development. Front Immunol 2018;9:2120
14. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, Kuranda K, Nigi L, Eugster A, Osterbye T, Maugein A, McLaren JE, Ladell K, Larger E, Beressi JP, Lissina A, Appay V, Davidson HW, Buus S, Price DA, Kuhn M, Bonifacio E, Battaglia M, Caillat-Zucman S, Dotta F, Scharfmann R, Kyewski B, Mallone R, ImMaDiab Study G: Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013