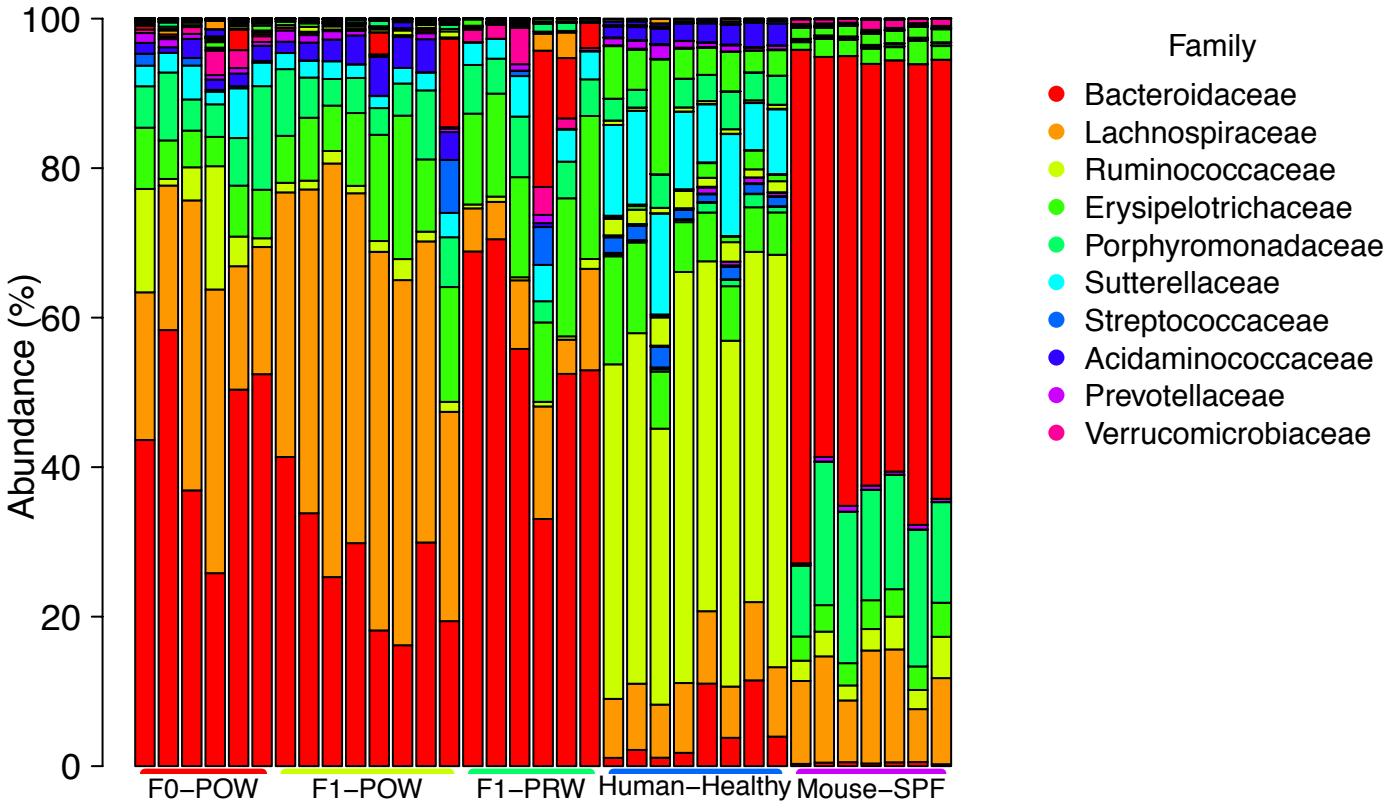

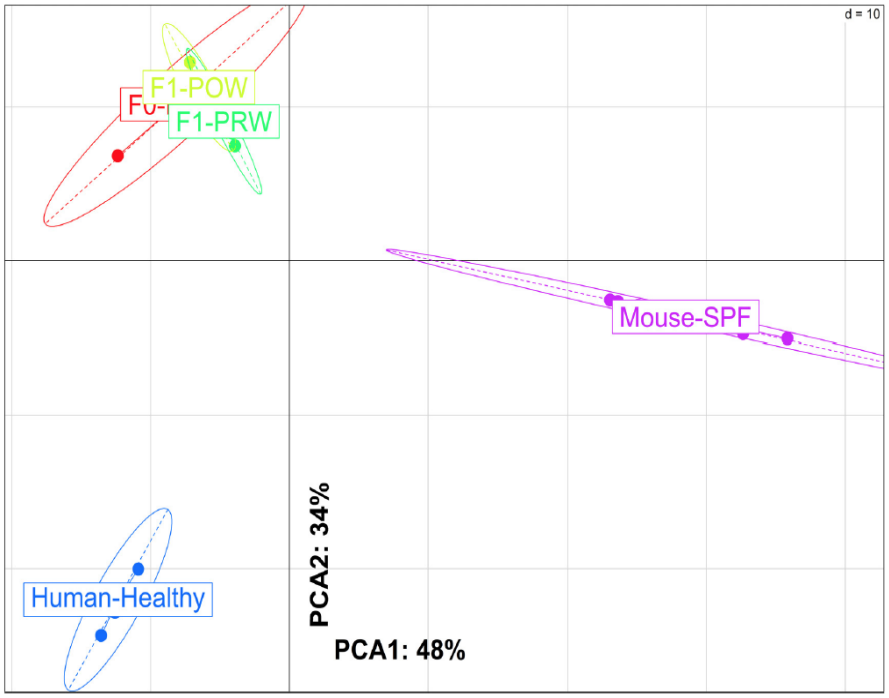

Fig. S1. Human microbiota reconstitution of breeding TAP Ii double KO mice. (A) Microbiota composition at the family level inferred for human replicates (H), SPF mice (SPF) and breeding mice reconstituted with human microbiota (F) before (F1-PRW) and after weaning (F1-POW) of the pups and mothers (FO). B) Between class analysis (BCA) computed with the principal components analysis of the whole microbiota composition dataset and constrained with H, SPF and F instrumental variables. The small distance along axis #1 indicates that the mice reconstituted with human microbial flora have some characteristics of a human flora. However, the flora has substantially drifted towards a mouse one as shown by the distance along axis #2.

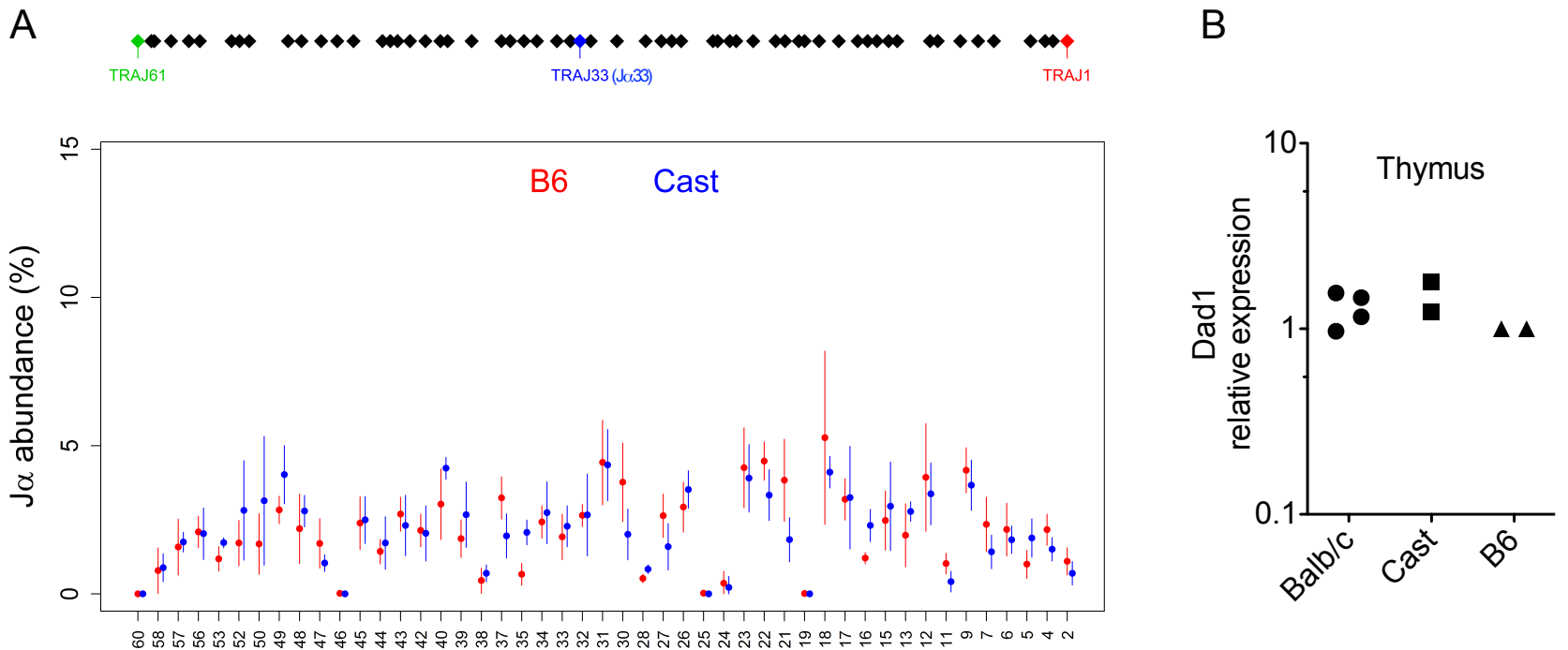

Fig. S2. **TCR Jα usage and Dad1 expression in CAST/EiJ mice**. (A) Thymic TCR Jα repertoire analysis from C57Bl/6J (red) and CAST/EiJ (blue) mice using 5'RACE and NGS sequencing. Proximal 5' to distal 3' Jα segments were plotted as mean± SD. (B) Thymic Dad1 (defender against cell death 1) mRNA expression analysis using TaqMan Gene Expression assay (Applied Biosystems). Fold change relative to C57Bl/6J is plotted, after normalizing to Cα.

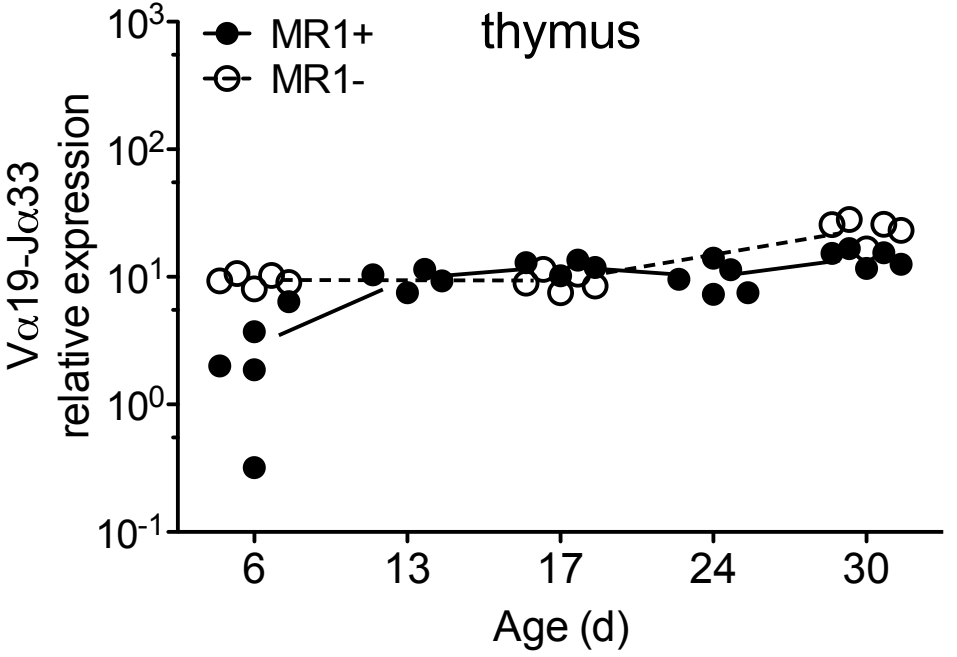

Fig. S3: No accumulation of MAIT cells in the thymus after birth. iVα19 TCRα transcript expression was quantified by RT-qPCR at the indicated time points after birth.

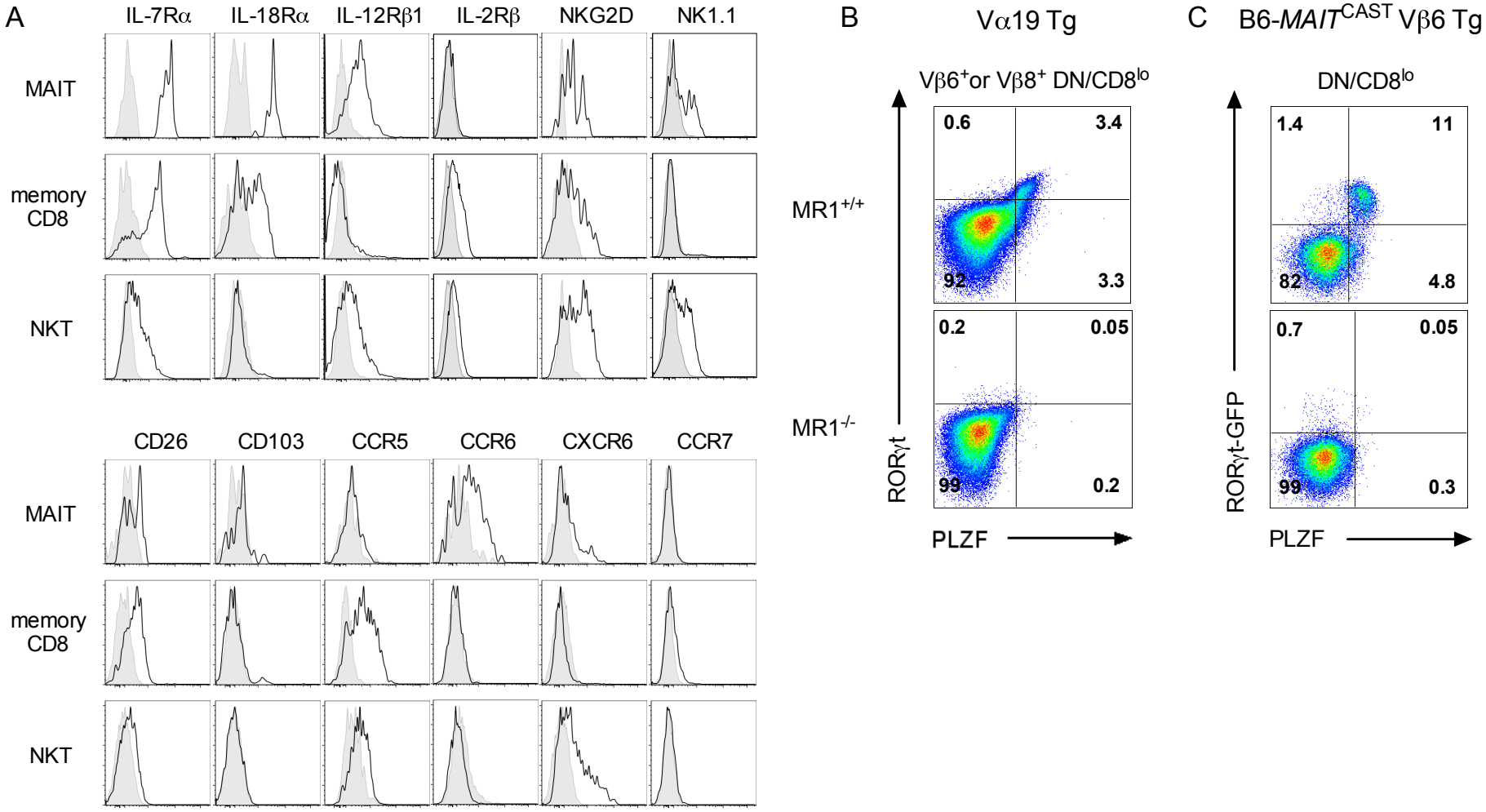

Fig. S4: Phenotype of murine MAIT cells. (A) Phenotype of MAIT cells in the liver, (B) Expression of PLZF and RORγt by Vβ6^{+} or Vβ8^{+} DN/CD8^{lo} T cells from iVα19 Tg Cα$^{-/-}$ mice on MR1$^{+/+}$ or MR1$^{-/-}$ background, C) Expression of PLZF by RORγt-GFP^{+} DN/CD8^{lo} TCRβ^{+} T cells from Vβ6 Tg B6-MAIT^{CAST} on a MR1$^{+/+}$ or MR1$^{-/-}$ background.

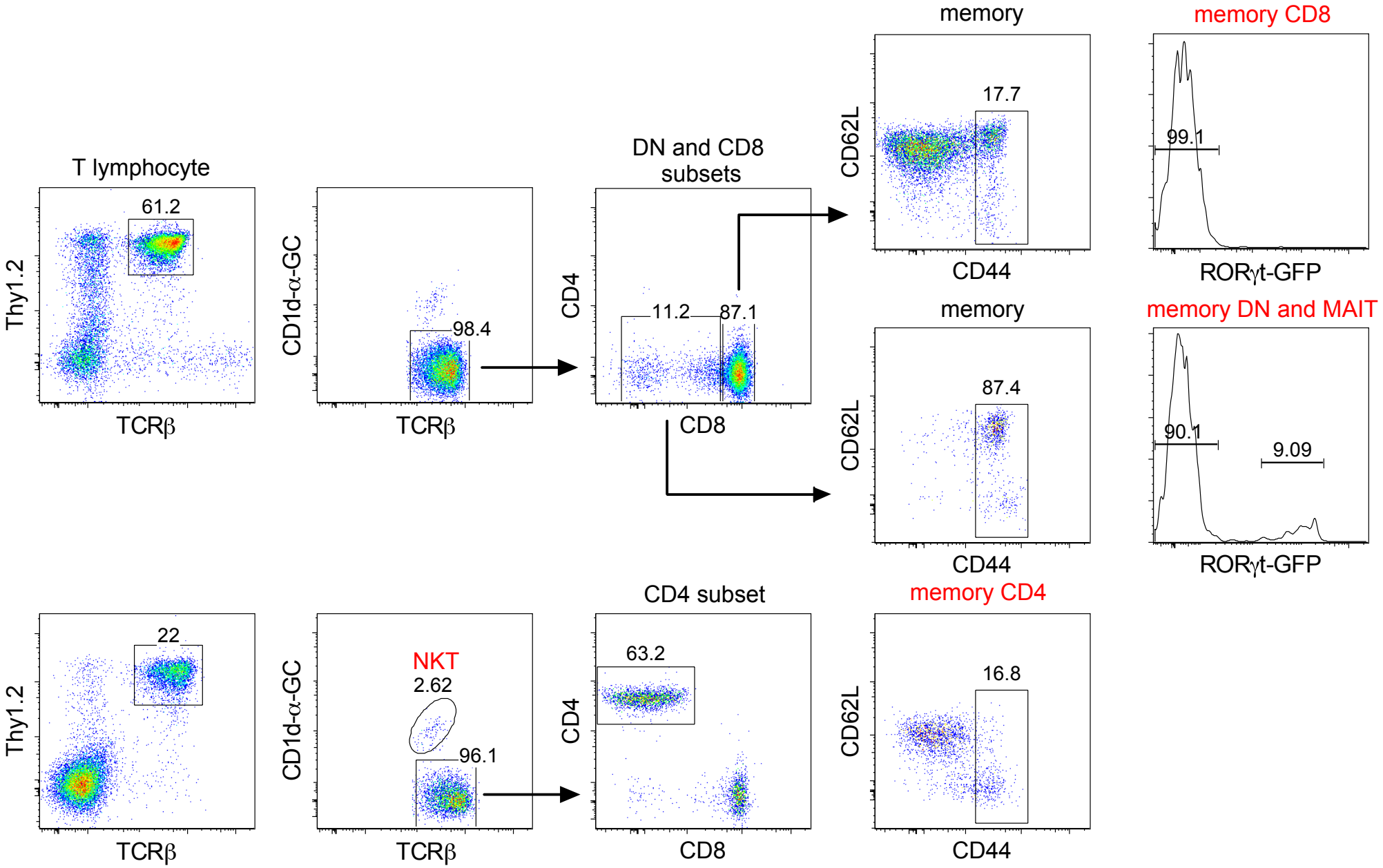

Fig. S5. Gating strategy of splenocyte sorting. Pools of 5-10 spleen were subjected to anti-CD19, -CD11c, -CD11b and -CD4 microbeads depletion to enrich DN and $CD8^+$ T cells before separation of memory CD8 ($CD8^{hi}CD44^+GFP^-$), memory DN ($CD8^{lo/neg}CD44^+GFP^-$) and MAIT ($CD8^{lo/neg}CD44^+GFP^+$) subsets by FACS sorting (top). Memory CD4 ($CD4^+CD44^+$) and NKT ($CD1d\text{-}\alpha\text{-}GC^+$) subsets were sorted from pool of splenocytes before microbead-enrichment (bottom).

Cui et al. Figure. S6

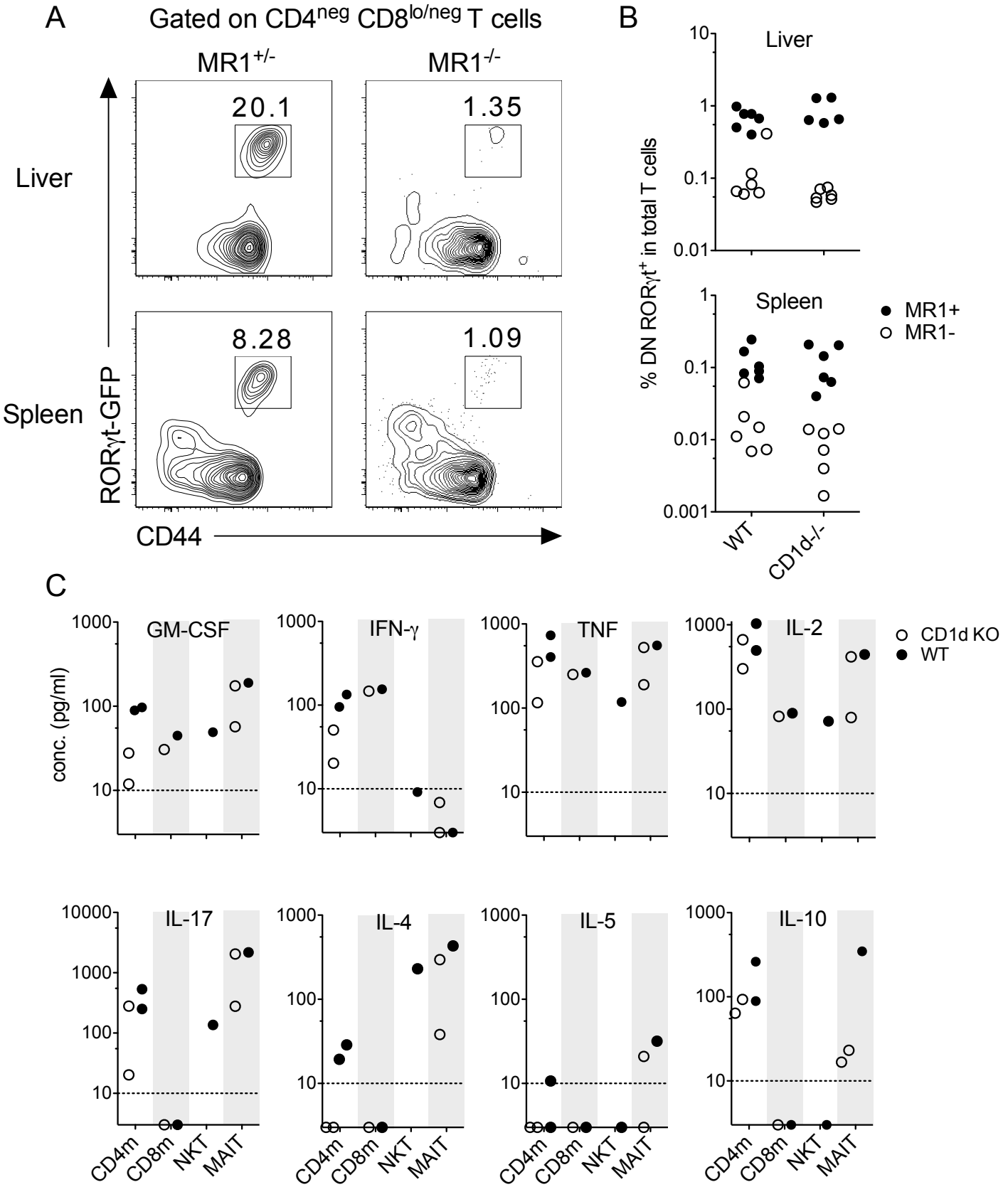

Fig. S6. Frequency and cytokine secretion pattern of MAIT cells from MR1^{+} or MR1^{-} CD1d$^{-/-}$ RORγt-GFP^{TG} B6-*MAIT*^{CAST} mice: (A, B) Frequency of MAIT cells in the liver and spleen of CD1d$^{-/-}$ RORγt-GFP^{TG} B6-*MAIT*^{CAST} mice according to the MR1 background. Representative staining gated on DNCD8^{lo} T cells (A) and cumulative data (B) from 3 independent experiments. C) Cytokine secretion of the indicated subsets isolated from CD1d^{+} and CD1d$^{-/-}$ mice as described in Fig. S5. Compilation of 2 independent experiments.

Cui et al, Figure. S7

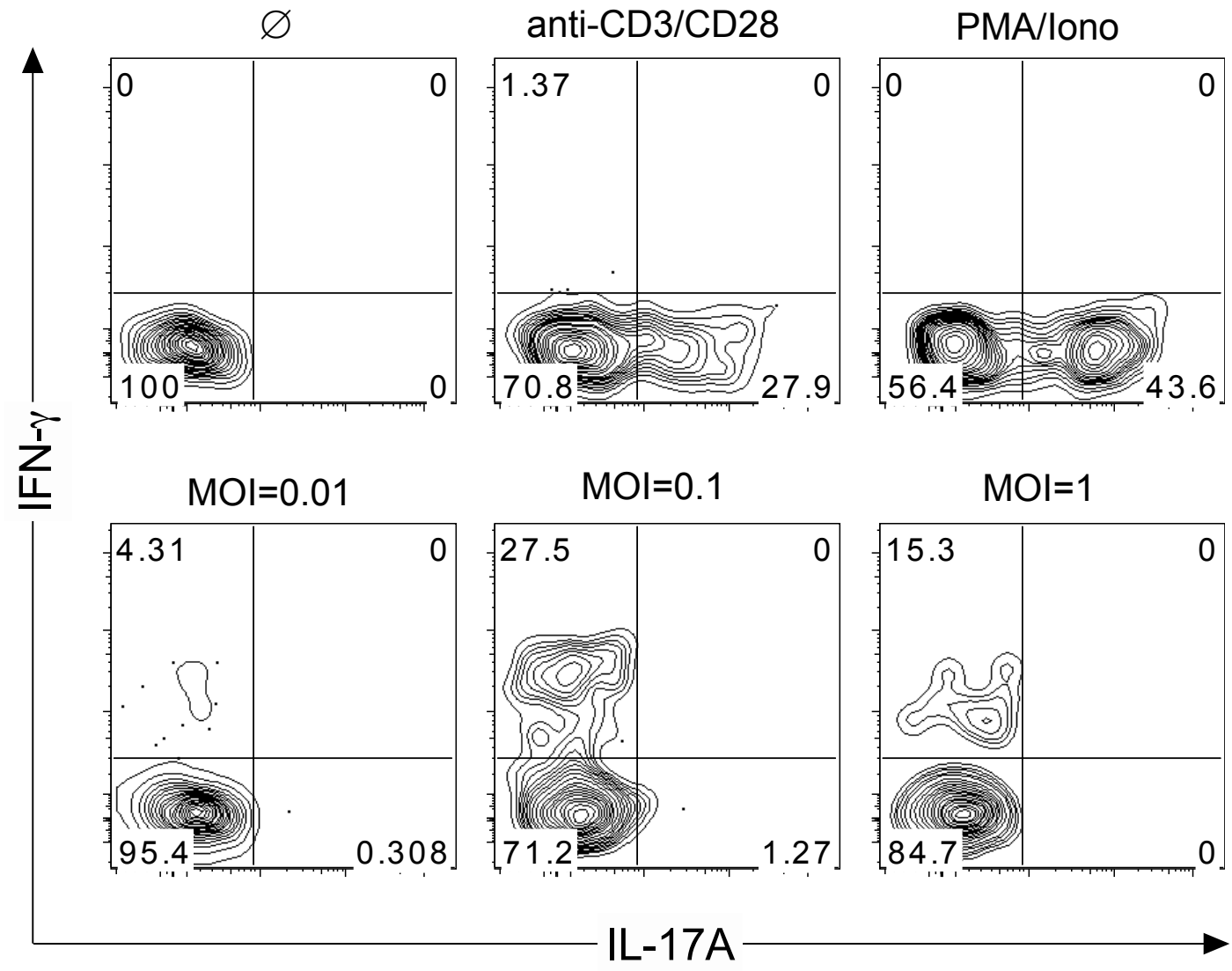

Figure S7. Cytokine production by mouse MAIT cells. Splenocytes from CD1d$^{-/-}$ Rorc(γt)-Gfp^{TG}-B6-*MAIT*^{CAST} mice were stimulated with anti-CD3/CD28 beads (9 h), PMA/Ionomycin (5 h), or PFA fixed *E.coli* (overnight, lower panel) and stained for intracellular IL-17A and IFN-γ. Representative plots from three independent experiments are shown gating on TCRβ^{+}DN/CD8^{lo}CD44^{+}GFP^{+} cells. No IL-4 staining was observed in any conditions (not shown).

Figure S8

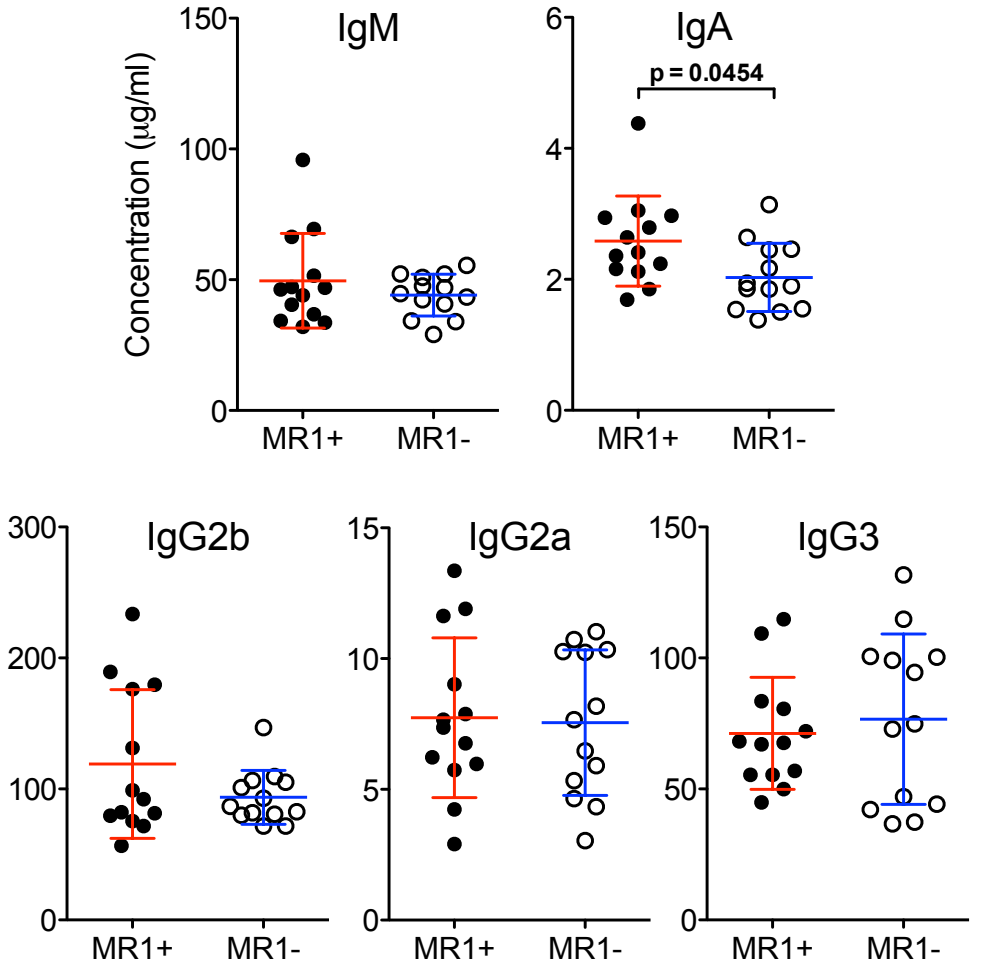

Figure S8. In the absence of MAIT cells, serum immunoglobulin isotype levels remain unchanged. ELISA was used to measure total IgM, IgG2a, IgG2b, IgG3 and IgA in sera drawn from the indicated unchallenged 7-9 week old mice. Bars indicate mean ± SD.

Cui et al supplementary materials and methods

Antibody list

All Abs were purchased from BioLegend, BD Pharmingen, eBioscience unless otherwise stated. The antibody clones used included: CD19 (1D3), TCRb (H57-597), CD90.2 (53-2.1, 30-H12), CD4 (RM4-5), CD8 (53-6.7), CD44 (IM7), CD62L (MEL-14), CD25 (PC61), CD69 (H1.2F3), mCD1d-α-GC tetramers (ProImmune), IL-7Rα (A7R34), IL-18Rα (BG/IL18RA), IL-12Rβ (114), IL-2Rβ (TM-Beta 1), NKG2D (CX5), NK1.1 (PK136), CD26 (H194-112), CD103 (2E7), CCR5 (HM-CCR5), CCR6 (29-2L17), CXCR6 (221002, R&D systems), CCR7 (4B12), PLZF (Mags.21F7), ROR(γt) (AFKJS-9), IL-17A (TC11-18H10.1), IFN-γ (XMG1.2), IL-4 (11B11).

***In vitro* stimulation**

E.coli were cultured overnight at 37°C in Luria Broth, pelleted and washed in PBS before fixation for 20 min in 1% of PFA. For intracellular cytokine staining, 2 x 10^6 splenocytes were stimulated with 10 ng/ml PMA (Sigma) and 1 μg/ml ionomycin (Sigma) for 5 h, anti-CD3/CD28 Dynabeads (1:1, Gibco) for 9 h or fixed *E.coli* overnight. Golgi Plug (1 μg/ml, BD) was added for the final 3-4 h of culture. Prior to surface marker staining, cells were incubated with live/dead violet fixable cell stain (Life Technologies) and Fc block. Cells were then fixed and permeabilized for 20 min using Cytofix/Cytoperm kit (BD) and subjected to intracellular IFN-γ, IL-4 and IL-17A staining.

ELISA

Blood samples were collected by retro-orbital bleeding from naïve B6-*MAIT*^{CAST} MR1^{+} or MR1$^{-/-}$ mice (n=13 for each genotype, 7-9 weeks old). Total serum IgM, IgG1, IgG2a, IgG2b, IgG3 and IgA titers were determined using Ready-Set-Go® ELISA kits, according to

the manufacturer's instructions (eBioscience). Data were analyzed using second-order polynomial model (Prism GarphPad 5.0) and plotted as concentration.

Fecal samples

DNA was extracted from stool as previously published (1, 2). DNA quality and quantity were checked using NanoDrop and Qubit together with BR assay.

16S rDNA Barcoding PCR

Barcoded PCR was carried out using barcoded primers targeting V4-V5 variable part of the 16S rRNA gene. iProof High fidelity master mix (Bio-Rad) was used to perform the PCR following manufacturer's recommendations. Mix for one PCR reaction was made up of 12.5 µL of iProof High fidelity master mix, 0.5 µL of both 563F and 924R primers (0.2 µM), 10 µL of RNase DNase free water and 1.5 µL of DNA (10 ng). The PCR program applied was 95°C for 5 min followed by 30 cycles (95°C for min, 41°C for min and 72°C for min) and a final step at 72°C for min. PCR reactions (25 µL) were then quantified using Qubit BR DNA assay kit. Equal amount of amplicon DNA were pooled and thoroughly mixed. The resulting amplicon pool was loaded on an agarose gel (1.5%) and electrophoresis run for 45 min at 100 V. The band corresponding to 16S amplicon was excised from the gel and purified using Qiaquick gel extraction kit (Qiagen). After gel extraction the amplicon quantity and intergrity were both checked on agarose gel electrophoresis and Qubit BR DNA assay kit. Amplicon was stored at -20°C until use.

16S rDNA sequencing and statistics

Sequencing of the amplicon pool was performed at Institut Curie NGS platform on an Ion Torrent PGM machine using the 318 chips. Raw sequencing output was demultiplexed and trimmed using in-house analysis pipeline (BCmm = 0, Pmm = 2, W = 50 bp, Q = 20, MinL = 200 bp and MaxL = 335 bp). Reads were clustered at 97% of identity using `vsearch`

pipeline (http://zenodo.org/record/15524#.VPIujy4f2J8). Chimeric OTU were identified using `UCHIME` (3) and discarded from downstream analysis. Taxonomy of representative OTU sequences was determined using RDP `classifier` (4). OTU sequences were aligned using `ssu-align` (5). The phylogenetic tree was inferred from the OTUs multiple alignments using `Fattree2` (6).

Statistical analyzes were performed using `R` program (7) and different related packages (`Ade4` (8), `Vegan` (9), `ape`, `phyloseq`). Weighted unifrac distances (10, 11) were computed using microbiota abundance table and phylogenetic tree. Statistical tests were performed using the Wilcoxon test (p-value). Multiple tests were corrected using the False Discovery Rate method (q-value).

References

1. Yu Z, Morrison M. Improved extraction of PCR-quality community DNA from digesta and fecal samples. [Internet]. *Biotechniques* 2004;36(5):808–12.

2. Ó Cuív P et al. The effects from DNA extraction methods on the evaluation of microbial diversity associated with human colonic tissue. [Internet]. *Microb. Ecol.* 2011;61(2):353–62.

3. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;27:2194–2200.

4. Cole JR et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. [Internet]. *Nucleic Acids Res.* 2009;37(Database issue):D141–5.

5. Nawrocki EP. Structural RNA Homology Search and Alignment using Covariance Models2009;

6. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. [Internet]. *PLoS One* 2010;5(3):e9490.

7. R Development Core Team. R: A Language and Environment for Statistical Computing2012;

8. Dray S, Dufour AB. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Staistical Softw.* 2007;22(4):1–20.

9. Oksanen J et al. vegan: Community Ecology Package [Internet][published online ahead of print: 2013];http://cran.r-project.org/package=vegan. cited

10. Lozupone C, Hamady M, Knight R. UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context.. *BMC Bioinformatics* 2006;7:371.

11. Lozupone C, Knight R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 2005;71:8228–8235.