Supplemental Methods

Organ collection and preparation

Spleen (BMT): Single-cell suspensions were obtained by mechanically disrupting harvested spleens and filtering through 70-µm mesh filters (Thermo Fisher Scientific). Ammonium-chloride potassium (ACK) lysis was performed to remove erythrocytes prior to cell counting.
Bone marrow (BMT): Cells were flushed from femur and tibia of donor mice and filtered through 70-µm mesh filters. Erythrocytes were lysed using ACK.
Blood: Circulating immune cells were isolated from blood drawn via cardiac puncture and subsequently filtered and lysed as described above.
Spleen and liver (downstream analysis): Mice were transcardially perfused with PBS, spleen and liver harvested and mechanically disrupted to generate single-cell suspensions. After filtering through 70 µm mesh filters, erythrocytes were lysed before proceeding with downstream applications.

Flow cytometry and fluorescence-activated cell sorting (FACS)

One million to two million live spleen and liver cells were aliquoted for flow cytometry staining. For fluorescence-activated cell sorting (FACS) of splenocytes, samples from multiple mice were pooled for each cohort (syngeneic recipients and allogeneic recipients) and stained for CD4, CD1d, CCR7 and CD44. mCD1d (PBS-57) tetramer was obtained from the NIH Tetramer Core Facility to allow exclusion of NKT cells from $CD4^+$ T cells. Pooled samples for each cohort underwent FACS using BD Influx achieving 95% purity. FACS-purified samples were collected into PBS buffer containing 2% (v/v) BSA. Sorted T cell subsets were further processed for use in metabolomics analysis, extracellular flux analysis and RNA-sequencing as described below. For the analysis of glycolytic enzymes, cells isolated from spleen and liver were first stained for CD45, CD3, CD4, CD8a, CD69, CCR7 and CD44 and subsequently fixed and permeabilized and stained for the intracellular antigens FoxP3, GAPDH, GLUT1 and HK2 or their respective isotype control. Individual mouse sample flow cytometry measurements were obtained on a BD LSR II or Fortessa (Becton Dickinson, USA), and analyzed using FlowJo Software (Becton Dickinson, USA). A list of antibodies is provided in Table 1.

Supplementary Table 1: Flow cytometry antibodies

Antigen	Clone	Manufacturer
CD45	30-F11	Biolegend
CD3	17A2	Biolegend
CD4	GK1.5	eBioscience
CD8a	5H10 / 53-6.7	Invitrogen / BD Biosciences
CD69	H1.2F3	Biolegend
CD44	IM7	Biolegend
CCR7 (CD197)	4B12	BD Biosciences
CD19	6D5	Biolegend
CD11b	M1/70	Biolegend
Ly6G	1A8	Biolegend
Ly6C	HK1.4	Biolegend

FoxP3	FJK-16s	eBioscience
GAPDH	14C10	Cell Signaling Technology
Rabbit IgG isotype	DA1E	Cell Signaling Technology
GLUT1	EPR3915	Abcam
Rabbit IgG isotype	EPR25A	Abcam
HK2	OTI4C5	Novus Biologicals
Mouse IgG isotype	11711	Novus Biologicals
CD1d	PBS57 tetramer	NIH Tetramer Core facility
CD16/32 (Fc block)	2.4G2	BD Biosciences
Live/Dead Aqua viability dye		Invitrogen

Mass spectrometry analysis of intracellular metabolites

Sorted splenic $CD4^+$ T cell subsets were centrifuged (~314 x g, 5 min, room temperature (RT)), the supernatant was discarded, and the pellet was resuspended in 10 mL aqueous mannitol (5%, w/v) by gentle mixing. The cell suspension was centrifuged (~314 x g, 5 min, RT) and the supernatant (aqueous mannitol) was discarded. The pellet was then resuspended in deionized water (50 µL) and centrifugation (~314 x g, 5 min, RT) was repeated. Approximately 100-150 µL of the supernatant was discarded followed by addition of 133.3 µL methanol (mixed for 30 sec) to quench intracellular enzymes. The working internal standard solution (10 µM) used for mass spectrometry (MS) analysis was added (91.7 µL), mixed for 30 sec and samples were stored at -80 °C. The processed T cell samples were pooled and quantified relatively utilizing capillary electrophoresis time-of-flight mass spectrometry (CE-TOF-MS, Agilent Technologies, Waldbronn, Germany) with analysis performed by Human Metabolome Technologies, Inc. (HMT, Japan). The samples were thawed, combined, vortexed and 1,100 µL was transferred to a 1.5 mL microcentrifuge tube and centrifuged (2,300 x g, 5 min, 4 °C). Two ~500 µL aliquots were transferred to 5 kDa MWCO centrifugal filter units (0.5 mL, Millipore Sigma, Burlington, MA, USA) and centrifuged (9,100 x g, 4 °C) until sample filtration was complete (~2-5 hrs.). The sample filtrates were transferred to 1.5 mL polypropylene microcentrifuge tubes and vacuum dried to completeness at room temperature using a Savant DNA 110 Speed Vac® (Savant Instruments, Inc., Farmingdale, NY, USA). The sample residue containing ~ 3 million T cells for MS analysis was capped and stored at -80 °C. After shipment to HMT, the sample residues were reconstituted with 50 µL deionized water, mixed and ~100 nL injected for CE-TOF-MS analysis.

Metabolic flux assay

Mitochondrial function and glycolysis were evaluated using a Seahorse XFe96 Bioanalyzer (Agilent) per the manufacturer instructions. Subsets of $CD4^+$ T cells from the spleen were FACS-purified and plated into 96-well plates coated with CellTak (Corning, #354240) with a seeding density of 100,000 live cells per well and a minimum of 8 wells per cell subset. Oligomycin (1 µM), 2-[[4-(trifluoromethoxy)phenyl]hydrazinylidene]propanedinitrile (FCCP, 0.5 µM) and Rotenone/antimycin (1 µM/ 1 µM) were loaded into the injection ports, injected as indicated in the figures and used to measure the oxygen consumption rate and extracellular acidification rate with a total plate run time of up to 140 minutes.

Lactic acid measurement

Mouse plasma L-lactate levels were quantified via mass spectrometry in mandibular blood samples at indicated timepoints. For quantitatively measuring L-lactic acid levels in mouse plasma, we used a Waters Acquity H-Class UPLC system coupled to a Waters Xevo ESI-MS/MS. As mouse blood volume is limited, we developed a MS method that utilized only 25 to 50 µL of mouse plasma, which was collected using a fluoride/oxalate microtainer tube. Briefly, the mouse plasma and internal standard (D3-lactic acid) were added to a 3K MWCO filter (Amicon®, 0.5 mL) and centrifuged (14,000 x g, 30 min, 20 ºC). The filtrate was transferred to a Waters autosampler microvial (deactivated) and 2 µL was injected for MS analysis. The 3K MWCO filter provided clean extracts for UPLC HILIC chromatography and the electrospray ionization ESI (-) SIR (selected ion recording) mode of MS analysis. For HILIC plasma component separation, we used a linear gradient consisting of aqueous ammonium acetate (26 mM) and acetonitrile, flow rate of 0.3 mL/min, and column temperature of 30 ºC. The HILIC chromatography provided good resolution of the plasma components and excellent peak shape for L-lactic acid and the internal standard. The quantitative SIR MS method utilized five calibration standards (L-lactic acid, 0.28-5.6 mM) and a deuterated internal standard (D3-lactic acid, 1.1 mM). The UPLC-ESI (-) SIR-MS method demonstrated good sensitivity (MS response), specificity and linearity using a quadratic fit calibration model for quantifying L-lactic acid.

RNA extraction and quantification

RNA was extracted from sorted spleen $CD4^+$ T subset populations using the RNeasy® Plus Mini Kit following the manufacturers protocol for purification of total RNA from animal cells (Qiagen). In brief, the cells were lysed in Buffer RLT Plus with β-ME, homogenized with a QIAshredder spin column, and the resultant lysate was run through a gDNA Eliminator spin column to remove genomic DNA. RNA was extracted by passing the remaining lysate through a RNeasy spin column and, following washes, total RNA was resuspended in 30 µl of RNase-free water. Measurement of RNA concentration and quality control was performed using a 2100 Bioanalyzer instrument with the standard manufacturers protocol (Agilent).

Bulk RNA Library and Sequencing Method

RNA-Seq libraries were constructed from 0.1-0.6 µg total RNA. The HyperPrep RNA-Seq Kit with Riboerase (Kapa Biosystems/Roche) was used according to manufacturer's instructions to deplete rRNA and then construct libraries for sequencing on Illumina platforms. Amplification was performed using 10 or 12 cycles which was optimized for the input amounts and to minimize the chance of over-amplification. Unique barcode adapters were applied to each library. Libraries were pooled in equimolar ratio for sequencing. The pooled libraries were sequenced on multiple lanes of a HiSeq 4000 to achieve a minimum of 65 million 75 base read pairs. The data was processed using RTA version 1.18.54 and CASAVA 1.8.2.

Bulk RNA sequencing

RNA-seq NGS-datasets were processed using the CCBR Pipeliner utility (https://github.com/CCBR/Pipeliner). Briefly, reads were trimmed of low-quality bases and adapter sequences were removed using Trimmomatic v0.33[1]. Mapping of reads to the Gencode mm10 mouse reference genome (M12 release) was performed using STAR v2.5.2b in 2-pass mode[2]. Then, RSEM v1.3.0 was used to quantify gene-level expression, with counts normalized to library size as counts- per-million[3]. Finally, limma-voom v3.34.5 was used for quantile normalization and differential

expression[4]. Data for the mitochondrial genome-encoded genes for the electron transport chain were not evaluated and the genes for complex II (succinate dehydrogenase) were selected as part of the Krebs cycle genes. Sequencing data are available under GEO accession number GSE147371.

Single cell RNA sequencing (scRNAseq)

Mouse scRNAseq

Murine lymphocytes were isolated from the liver of three syngeneic and three allogeneic on day 14 after bone marrow transplantation as described above. Cells were stained for CD45 and CD4 and $CD45^+$ $CD4^+$ cells were sorted into PBS supplemented with 1% BSA. Viability for all samples was >82%. To profile the 5' single cell gene expression, single cell suspensions were then loaded onto a Chromium Single Cell Chip (10X Genomics) with a recovery target of 6,000 cells per lane according to the manufacturer's instructions. All subsequent steps of cDNA generation, library preparation and quality control were performed according to the 10X Genomics 5' single cell user guide. Four NextSeq runs and one MiSeq run were performed and the standard 10X Genomics cellranger (version 3.0.2) pipeline was used to extract Fastq's and 10X Genomics cellranger (version 3.0.2) pipeline was used to perform data processing. Sequenced reads were aligned to the 10X Genomics provided Mouse mm10 reference sequence (refdata-cellranger-mm10-3.0.0). Sequencing data are available under GEO accession number GSE153591.

Supplementary Table 2: Mouse scRNAseq output metrics summary

Sample	Estimated Number of Cells	Mean Reads per Cell	Median Genes per Cell	Sequencing Saturation	Total Genes Detected	Median UMI Counts per Cell
syn1	5,396	41,046	1,957	74.20%	17,534	5,179
syn2	5,456	45,761	1,937	78.40%	17,299	5,263
syn3	4,751	40,256	2,075	73.80%	17,137	5,624
allo1	6,226	45,438	2,178	70.80%	17,219	6,469
allo2	5,744	47,044	2,210	71.30%	16,989	6,438
allo3	5,894	46,882	2,079	73.40%	16,953	5,950

Seurat (v3.1.1) was used for the downstream analysis following the vignette for the analysis of an integrated dataset[5]. Briefly, syngeneic and allogeneic replicates were first merged and filtered for cells containing at least 200 and less than 4500 features as well as a mitochondrial gene content of less than 5%. Samples were normalized and variable genes identified using the FindVariableFeatures function with nfeatures limited to 2000. Both samples were integrated, scaled and principal components were calculated. Clustering was performed using Uniform Manifold Approximation and Projection (UMAP) with the first 20 dimensions and a resolution of 0.5. Conserved markers between each cluster for both samples were identified using the FindConservedMarker function. Differentially expressed genes between syngeneic and allogeneic cells were calculated using the FindMarker function. Module scores were calculated

using the AddModuleScore function using pre-defined gene sets of interest as indicated in Supplementary Table 3.

Supplementary Table 3: Module gene sets

Module	Genes
Glycolysis	Slc2a1, Slc2a3, Hk1, Hk2, Gpi1, Pfkp, Aldoa, Tpi1, Gapdh, Pgk1, Pgm1, Eno1, Pkm, Ldha, Slc16a1, Slc16a3
TCA	Cs, Aco2, Idh2, Ogdh, Sucla2, Suclg1, Sdha, Sdhb, Sdhc, Sdhd, Fh1, Mdh2
Th1*	4930453N24Rik,Abracl,Acot7,Actb,Actg1,Actr3,Adgre5,Agpat4,Ahnak,Anxa1 Anxa2,Anxa5,Anxa6,Ap2s1,Aprt,Arhgdib,Arl6ip5,Arpc5,Atp1a1,Atp5c1 Atp5h,Atp5l,AW112010,Bhlhe40,Calm1,Capzb,Ccl5,Ccr2,Cd2,Cd47 Cd48,Cd52,Cdc42,Cdc42ep3,Cfl1,Clic1,Clta,Coro1a,Cox17,Cox5a Cox5b,Crip1,Crot,Ctla2a,Ctsc,Ctsd,Ctsw,Cxcr6,Cyba,Cyth4 Dnajc15,Dok2,Ech1,Emp3,Eno1,Epsti1,Esyt1,Fam107b,Gabarapl2,Gapdh Gbp7,Ggh,Gimap7,Glipr2,Glrx,Gm4950,Gmfg,Gna15,Gramd3,Gzmb H2afy,H2afz,Hmgb2,Hsp90b1,Id2,Idh3a,Ifngr1,Il18r1,Il18rap,Il2rb Itga4,Itgb2,Itgb7,Klrc1,Krtcap2,Lfng,Lgals1,Lgals3,Lsp1,Ly6c2 Mrpl33,Ms4a4b,Ms4a6b,Mtpn,Myl12a,Myl6,Myo1f,Myo1g,Ndufb7,Nkg7 Nptn,Ostf1,Pglyrp1,Plac8,Plek,Podnl1,Ppib,Ppp1ca,Ppp3ca,Prelid1 Psmb3,Pycard,Rap1b,Rasgrp2,Rbm3,Reep5,Rnf138,Rnf166,Rora,Rpa2 Runx3,S100a10,S100a11,S100a13,S100a4,S100a6,S1pr4,Sec61b,Sec61g,Selplg Sept11,Serpinb6b,Serpinb9,Sh3bgrl3,Slamf1,Sms,Spn,St3gal6,Stmn1,Sub1 Tagln2,Taldo1,Tbx21,Tceb2,Thy1,Tmed2,Tmsb4x,Tpm4,Tpst2,Tspo Txn1,Txndc5,Ube2g2,Vim,Zyx
Th2	Il4ra,Icos,Maf,Nfatc1,Gfi1,Gata3,Cebpb,Irf4,Stat6,Ccl5,Asb2,Jak1,Il13,Nfatc2ip,Il4 Ccr3,Irf8,Tmed1,Ccr4,Stat5a,Pparg,Il13ra1
Dysfunction*	2010111I01Rik,Arl6ip1,AW112010,Bhlhe40,Ccl3,Ccl4,Ccl5,Ccr5,Cd160,Cd27 Cd3g,Cd7,Cst7,Ctla2a,Ctla4,Ctsb,Ctsw,Cxcr6,Cyba,Dnaja1 Efhd2,Fasl,Fyn,Gadd45b,Gimap7,Gpr65,Gzmk,H2-K1,Hspa1a,Hspa8 Hsph1,Id2,Lag3,Lilr4b,Nkg7,Nr4a2,Pdcd1,Plac8,Psmb10,Psmb8 Rgs1,Rnaset2b,Samsn1,Serpina3g,Sh2d2a,Shisa5,Slamf7,Sub1,Tapbpl,Tnfrsf1b Traf4,Ubb
Tcm*	1810058I24Rik,Ablim1,Acot7,Acp5,Actn1,Adgre5,Adk,Ahnak,Anp32a,Anxa2 Arl5c,Atp1b3,Atp2b1,Bcl2,Bin2,Ccdc28b,Ccr7,Cd9,Cdc25b,Cdc42se1 Cdkn2d,Crip1,Crip2,Dock2,Emb,Ezr,Fam177a,Fam65b,Glud1,Hint1 Id3,Il6ra,Iqgap1,Itgb1,Itpkb,Jund,Klf2,Klf3,Klf6,mt-Co1 Pbxip1,Pdlim1,Prr13,Psma6,Raf1,Rasa3,Rasgrp2,Rora,Rpa1,S100a10 S100a11,S100a4,S100a6,Sfr1,Slamf6,Spn,Srpk1,Stk24,Stk38,Tagap Tagln2,Tcf7,Tcp11l2,Tspan13,Tuba1a,Vim,Vsir,Xrn2
Tfh*	2310001H17Rik,Aldoa,Angptl2,Asap1,Batf,BC021614,Bcl2a1b,Borcs8,Cd160,Cd200 Cd3g,Cd82,Cox14,Ctsb,Cxcr5,Ddit4,Dennd2d,Eea1,Fam162a,Fyn Gapdh,Gdi2,Gimap5,Gna13,Gng2,Hif1a,Hmgb1,Icos,Ifi27l2a,Isg15 Izumo1r,Limd2,Lpp,Lrmp,Maf,Marcksl1,Matk,Mif,Mmd,Nsg2 Nt5e,P2rx7,Pdcd1,Pfkl,Pfkp,Pgam1,Pkm,Ppp1r14b,Prkca,Ptp4a2 Ptpn11,Ptprcap,Ptrh1,Rab37,Rgs10,Rnaset2a,Rnaset2b,Rpsa,Scd2,44081 Sh2d1a,Smco4,Sostdc1,Tbc1d4,Tnfaip8,Tnfsf8,Tox,Tox2,Tpi1,Trim8 Zap70,Zfp36l1

Treg*	Bmyc,Btg1,Capg,Ccnd2,Cd2,Cd74,Ctla4,Foxp3,Gbp7,Gimap7,Gpx4,H2-T22,Ifngr1,Ikzf2,Il7r,Izumo1r,Ltb,Mbnl1,Peli1,Rgs1,Samsn1,Sdf4,Sell,Serinc3,Shisa5,Tnfrsf18,Tnfrsf4
*	Gene sets defined by Ciucci et al.[6]

Human scRNAseq patient characteristics

We screened annotated clinical data from a GvHD clinical study performed at the NIH Clinical Center (NCT00520130). Two patients that matched our criteria were identified. One patient ("pre-GvHD") received an 7/8 matched unrelated donor (MUD) peripheral blood hematopoietic stem cell transplant (PBHSCT) for CML in remission. The other ("No GvHD") received an 8/8 MUD PBHSCT for relapsed mantle cell lymphoma in remission. Both patients received reduced intensity conditioning with fludarabine and cyclophosphamide and were receiving methotrexate, tacrolimus and sirolimus for GvHD prophylaxis at the time of PBMC collection. The blood samples for both patients were collected on post-transplant day +30. The patient "pre-GvHD" developed an erythematous rash and nausea on post-transplant day +35, with upper GI biopsy on day +37 consistent with GvHD, i.e. epithelial apoptosis in the duodenum, stomach antrum and body (CMV and *H.pylori* negative). Patient "No GvHD" did not have symptoms of acute GvHD, but did have ongoing CMV retinitis that preceded HSCT and for which systemic anti-virals had been discontinued on day -1 with once weekly intravitreous anti-viral injection ongoing at the time of sample collection. Since acute GvHD significantly increases chances of developing chronic GvHD we decided to exclude patients that may have had sub-clinical undiagnosed acute that later went on to develop cGvHD. Thus, we selected our "No GvHD" patient based on the fact that at 3 years post HSCT this patient showed no signs of cGvHD.

Human scRNAseq sample preparation and sequencing

Frozen PBMC samples were thawed rapidly for 90 s at 37 ˚C, slowly resuspended in RPMI supplemented with 5% fetal bovine serum and immediately put on ice. Cells were spun at 300 rcf for 10 min, resuspended in 1 ml RBC lysis buffer (Thermo Fisher Scientific, #00-4333-57) and incubated at room temperature for 5 min. Following lysis, cells were washed twice to remove ambient contaminating mRNA, reduced to a volume of 100 µl and viability was assessed using a LunaFL cell counter (Logos Biosystems). Viability for both samples was >90%. Initial sample processing for 10X single cell sequencing was performed analogous to the mouse samples as described above. Three NextSeq runs and one MiSeq runs were performed and the standard 10X Genomics cellranger (version 3.0.2.) pipeline was used to extract FASTQ files and perform data processing. Sequenced reads were aligned to the human GRCh38 reference sequence provided by 10X Genomics (refdata-cellranger-GRCh38 3.0.0.0).

Analogous to the mouse data set, Seurat (v3.1.1) was used for the downstream analysis following the vignette for the analysis of an integrated dataset. Briefly, both samples were first filtered for cells containing at least 200 and less than 4000 features as well as a mitochondrial and ribosomal gene content less than 15% and 25%, respectively. Samples were normalized and variable genes identified using the FindVariableFeatures function with nfeatures limited to 2000. Both samples were integrated, scaled and principal components were calculated. Clustering was performed using UMAP with the first 20 dimensions and a resolution of 0.5. Conserved markers between each cluster for both samples were identified using the FindConservedMarker function and cluster identities were assigned based on previously published marker genes in combination with the web-based tool Enrichr (https://amp.pharm.mssm.edu/Enrichr/). Differentially expressed genes of the no GvHD vs. pre-GvHD sample were calculated using the FindMarker function.

The R package EnhancedVolcano was used to generate the volcano plots based on the differential gene expression within a cluster. Module scores were described as above using the human gene homologs for each gene set.

Supplementary Table 4: Human scRNAseq output metrics summary

Sample	**Estimated Number of Cells**	**Mean Reads per Cell**	**Median Genes per Cell**	**Sequencing Saturation**	**Total Genes Detected**	**Median UMI Counts per Cell**
No GvHD	6,830	66,564	1,593	84.10%	19,734	4,289
Pre-GvHD	5,219	62,829	1,468	83.20%	18,625	4,272

1 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

2 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

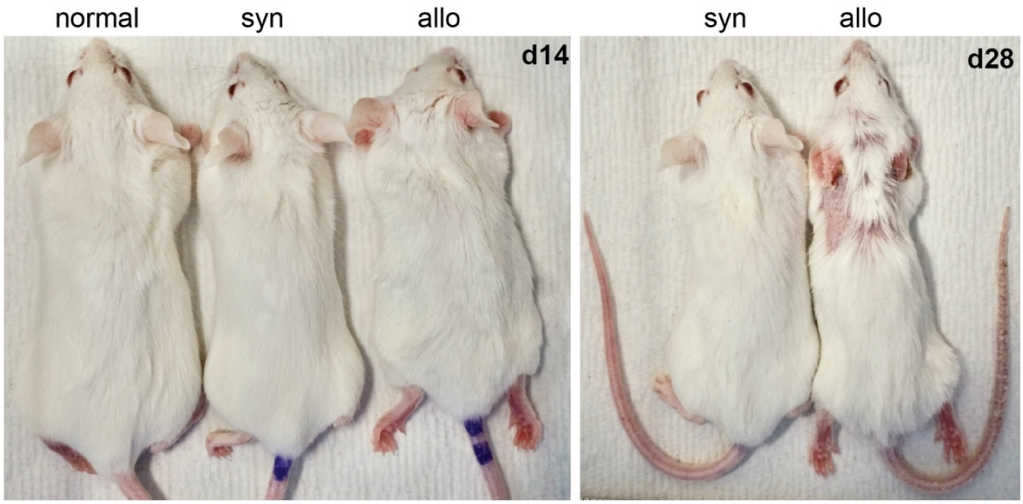
3 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

4 Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. *Ann Appl Stat* **10**, 946-963 (2016).

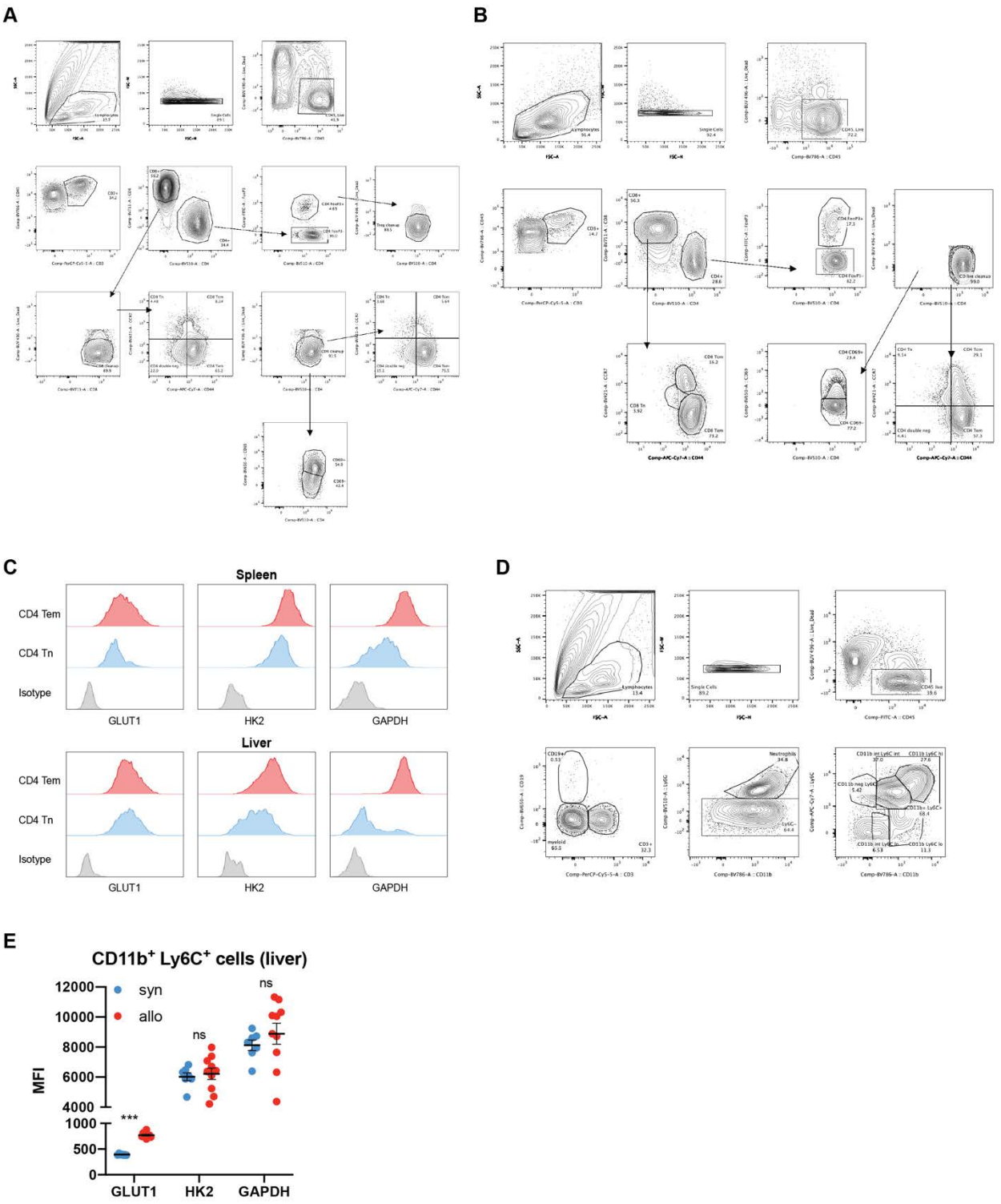
5 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).

6 Ciucci, T. *et al.* The Emergence and Functional Fitness of Memory CD4(+) T Cells Require the Transcription Factor Thpok. *Immunity* **50**, 91-105 e104 (2019).

Supplemental Figures

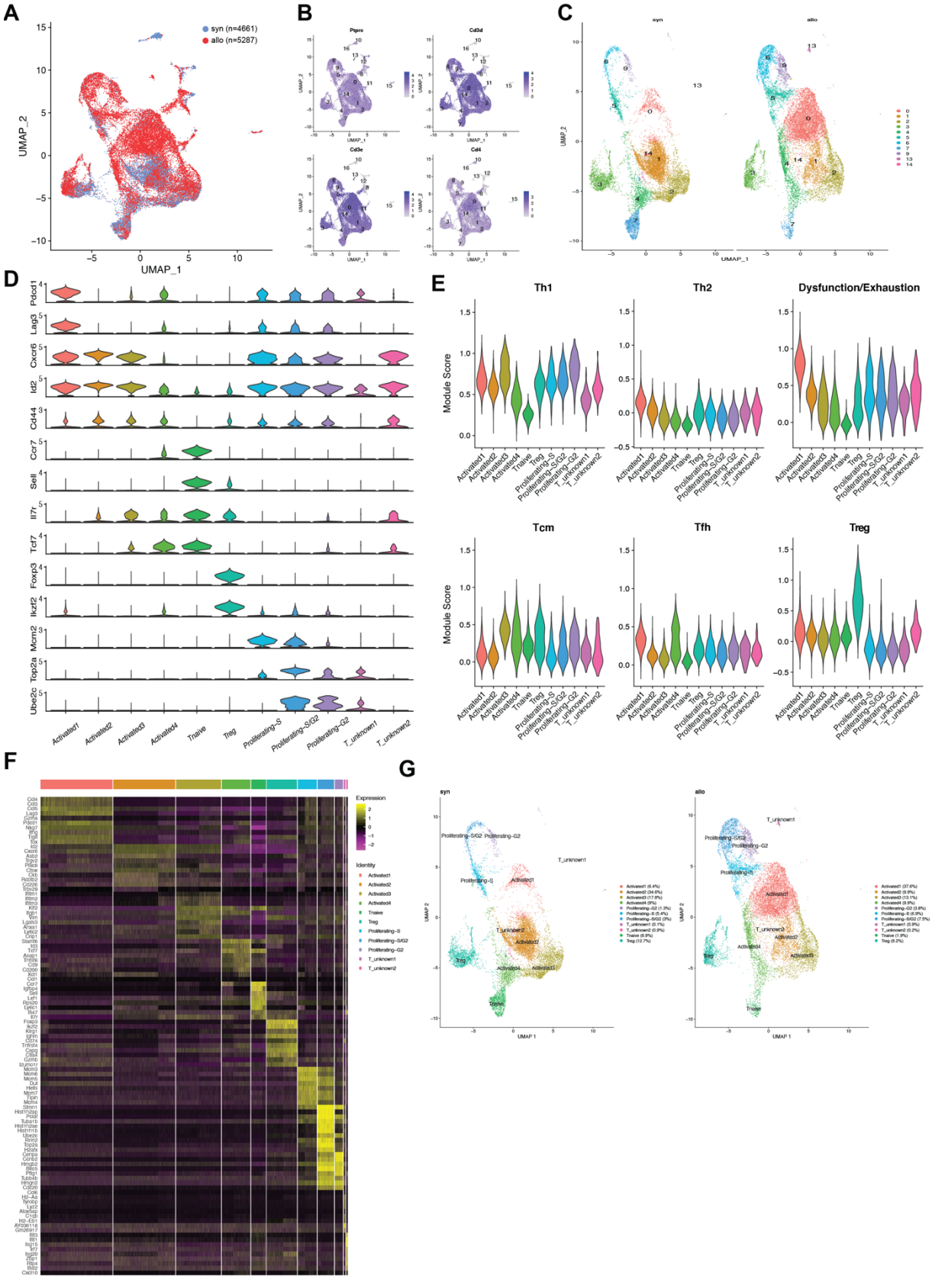


Supplemental Figure 1. Physical appearance of mice in the B10.D2 → BALB/c chronic GvHD model. A) Representative examples of mice on d14 (left) and d28 (right) after bone marrow transplantation. The normal mouse is of the same age but did receive neither irradiation nor a bone marrow transplant.

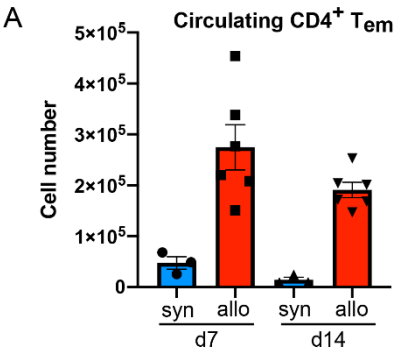


Supplementary Figure 2. Flow cytometry gating strategy and representative examples of MFI results. Flow cytometry gating strategies to quantify the MFI for GLUT1, HK2 and GAPDH in T cell subsets isolated from **A)** the spleen and **B)** the liver of syngeneic and

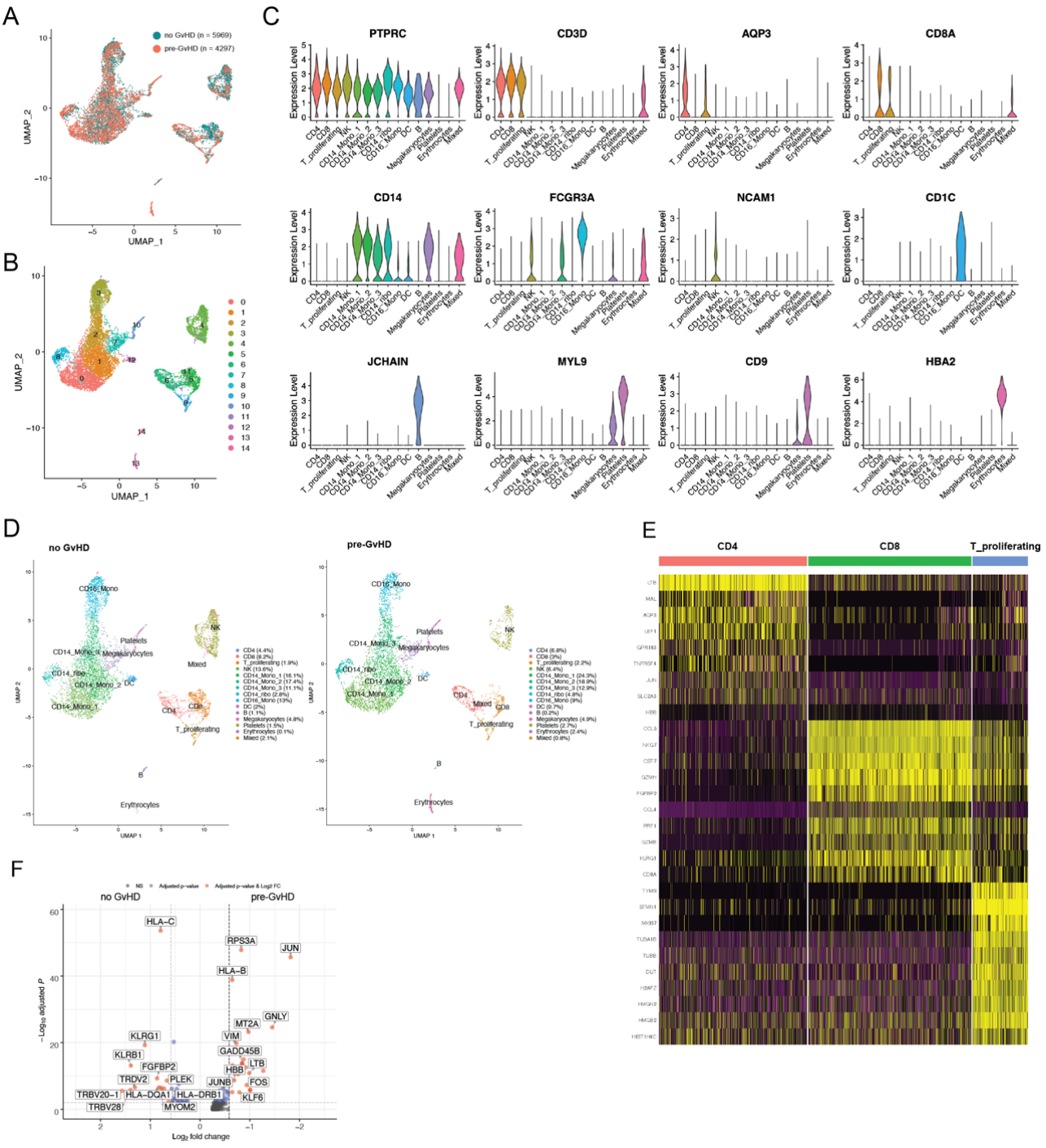
allogeneic animals on day 14 post-transplant. **C)** Representative examples of median fluorescence intensities in naïve and effector-memory CD4 T cells for GLUT1, HK2, and GAPDH. **D)** Gating strategy to quantify the MFI in monocytes/macrophages isolated from the liver of syngeneic and allogeneic animals on day 14 post-transplant. **E)** Quantification of MFI-based protein expression of GLUT1, HK2 and GAPDH in $CD45^{+}CD11b^{+}LY6G^{-}Ly6C^{+}$ cells. Data displayed as mean ± SEM. Welch's ANOVA with Dunnet's T3 multiple comparison test, n =7 (syn) / 10 (allo).


A
syn (n=4661)
allo (n=5287)
UMAP_1
UMAP_2
B
Pigre
Cd3d
Cd3e
Cd4
C
syn
allo
D
Pdcd1
Lag3
Cxcr6
Id2
Cd44
Ccr7
Sell
Il7r
Tcf7
Foxp3
Ikzf2
Mcm2
Top2a
Ube2c
Activated1
Activated2
Activated3
Activated4
Tnaive
Treg
Proliferating-S
Proliferating-S/G2
Proliferating-G2
T_unknown1
T_unknown2
E
Th1
Th2
Dysfunction/Exhaustion
Tcm
Tfh
Treg
Module Score
F
Expression
Identity
G
syn
allo
UMAP 1
UMAP 2

Supplementary Figure 3. Single-cell RNA sequencing of CD45$^+$ CD4$^+$ cells from the liver on day14 post-transplant. A) CD45$^+$ CD4$^+$ cells were isolated from the liver of syngeneic (n=3) and allogeneic animals (n=3) on day 14 post-transplant via FACS and processed separately using the 10X Genomics 5' sequencing kit. For the analysis, syngeneic and allogeneic replicates were merged, ultimately yielding 4661 (syn) and 5287 (allo) cells that were used for downstream analysis. **B, C)** Dimensional reduction resulted in 16 clusters, 11 of which expressed Cd3d, Cd3g and Cd4 and were therefore selected for further analysis. **D)** Overview of key marker expression across clusters used to assign cluster identities. **E)** Module scores for the indicated T cell subtypes as described in part by Ciucci et al., 2019. **F**) Top 10 conserved markers for each cluster. **G)** Assigned identities and cell frequencies for each cluster in syn (left) and allogeneic (right) animals.



Supplemental Figure 4. Circulating T$_{em}$ cells at early timepoints post-HSCT. To quantify circulating CD4$^+$ T$_n$ and T$_{em}$, blood was drawn via cardiac puncture into heparin-coated tubes, red blood cells were lysed, and the remaining cells were stained with antibodies against CD4, CD44, CCR7, CD25, FOXP3 and CD69 (T$_{em}$ defined as CD4$^+$, FOXP3$^-$,CCR7$^-$, CD44^{high}), n = 3-6, all data shown as mean ± SEM.



Supplemental Figure 5. Identification of cluster identities in the human PBMC single cell RNA sequencing dataset. A) UMAP visualization of the "no GvHD" and "pre-GvHD" samples and the total cell number that went into the analysis. **B)** Clustering result of the integrated data set yielding 14 separate clusters. **C)** Gene expression of known PBMC markers used to assign

cluster identities. **D)** UMAP visualization of assigned cluster identities for both samples with the cluster frequency indicated for each cell type. **E)** Heatmap illustrating the top 10 conserved genes for each of the three identified T cell clusters. **F)** Volcano plot of the differential gene expression within the CD8 T cell cluster comparing the no GvHD vs. pre-GvHD sample. Genes with a fold-change >1.5 and adjusted p-value <0.01 are labeled and highlighted in red.