Single-cell RNA sequencing reveals heterogeneity of mucosa-associated invariant T cells in donor grafts and its diagnostic relevance in gastrointestinal graft-*versus*-host disease

Mengge Gao,[1] Kainan Zhang,[1] Siqi Li,[1] Huidong Guo,[1] Yuqian Sun,[1] Jun Kong,[1] Tingting Han,[1] Yingjun Chang,[1,2] Xiaojun Huang[1,2,3,4] and Xiaosu Zhao[1,2,4]

[1]Peking University People's Hospital, Peking University Institute of Hematology, National Clinical Research Center for Hematologic Disease, Beijing Key Laboratory of Cell and Gene Therapy for Hematologic Malignancies, Peking University; [2]Research Unit of Key Technique for Diagnosis and Treatments of Hematologic Malignancies, Chinese Academy of Medical Sciences, 2019RU029; [3]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University and [4]Collaborative Innovation Center of Hematology, Peking University, Beijing, China

Correspondence: X. Zhao
zhao.xiaosu@outlook.com

Received: December 28, 2024.
Accepted: July 22, 2025.
Early view: July 31, 2025.

https://doi.org/10.3324/haematol.2024.287114

Supplementary Methods

Human Samples

We recruited four healthy adult donors and treated them with recombinant G-CSF (filgrastim; Kirin Co., Ltd., Tokyo, Japan) at a dosage of 5 μg/kg body weight per day for 5 consecutive days. The PBSC grafts were collected before G-CSF mobilization and after five days of G-CSF treatment, and then subjected to scRNA-seq. These donors and their matched recipients underwent haploidentical allo-HSCT at our institute between August 2020 and March 2021. Table S1 details the characteristics of donors and recipients, including sex and age. For multiparameter flow cytometry (FCM) analysis, the PB of 16 healthy donors was collected before and after G-CSF administration. In addition, a clinical cohort of 80 G-CSF-mobilized PBSC grafts (G-PBSCs) from healthy donors was also analyzed (Table S2 and Table S3).

Mucosa-Associated Invariant T Cell Isolation

Human PB came from healthy donors before and after *in vivo* 5-day G-CSF application, and peripheral blood mononuclear cells (PBMCs) were separated by Ficoll density centrifugation. Erythrocytes were removed by incubating with RBC lysis buffer (BD, 555899). After neutralization, the remaining cells were collected by spinning at 1500rpm for 5 min and suspended in FACS buffer (1× PBS with 1% BSA) for subsequent staining. Cells were stained in sorting buffer with MAIT cells specific antibodies for 30 min at 4 °C. MAIT cells (CD3+CD161^{hi}Va7.2+) were isolated using fluorescence-activated cell sorting (FACS)—sorted (FACS Aria II, BD Biosciences, USA) with antibodies CD3 (Percp, Biolegend, USA), CD161 (PE-CY7, Biolegend, USA), Va7.2 (PE, Biolegend, USA) according to the manufacturer's instructions, which included one step of positive selection for CD3+ cells and next step of double positive selection for CD161^{hi} and Va7.2+ cells.

Single Cell RNA Analysis

scRNA-seq libraries were prepared using the Chromium Single Cell 3′ Reagent Kits v3 (10x Genomics), according to the manufacturer's instructions. The libraries were sequenced on the Illumina NovaSeq6000 platform (Novogene, provided by Berry

Genomics Corporation, Beijing, China).

For scRNA-seq data, raw gene expression matrices were generated for each sample by the Cell Ranger (version 6.0.2) pipeline coupled with human reference hg38. UMI counts were normalized by the NormalizeData function and top 2000 variable genes were calculated using the FindVariableFeatures function. Next, Principal components analysis (PCA) was performed on the scaled data, and based on the elbow plot, 30 principal components (PCs) were selected for clustering, default resolution (0.4) was used. The anchors were used for batch correction to avoid the batch effect across samples. The data were used in the subsequent nonlinear dimensional reduction with the RunUMAP function and cluster analysis by the FindNeighbors and FindClusters (dims=1:30, resolution=0.4) functions. Fifteen clusters were calculated, one of which had high expression of RBC-related genes or features, which was removed, and the remaining 14 clusters were subjected to a subsequent series of analyses.

Cell Type Annotation and Cluster Marker Identification

Cells were projected into two-dimensional space using UMAP and clustered by common features. Cell clusters were annotated based on the expression of cell type-associated genes, with signature genes identified using the FindAllMarkers function in Seurat. Additionally, all percentages of different cell types were calculated.

***In Vitro* Proliferation and Suppression Assays**

The PBMCs from healthy adult donors were treated with or without G-CSF *in vitro*. Each 1×10^6 PBMCs was placed in 6-well plates and cultured with 2ul G-CSF per well for 24 h, then washed with PBS and stained with antibody. CD3+CD161^{hi}Vα7.2+ MAIT and CD4+CD25- T cells were sorted by flow cytometry. Purified CD4+ T lymphocytes were labeled with 5×10^{-6} m 5,6-carboxy-fluorescein diacetate succinimidyl ester (CFSE) (BD Bioscience) for 15 min at 37 °C and washed twice with complete medium with 10% FBS (fetal bovine serum). Isolated G-CSF mobilized and non-mobilized MAIT cells were co-cultured with labeled CD4+T cells (2×10^4 cells per well) at MAIT/CD4+T ratios of 1:1, 2:1, and 4:1 in the presence of Dynabeads Human T-Activator CD3/CD28 beads (Thermo, 11132D) in RPMI

complete medium (Gibco, Invitrogen) in flat-bottomed 96-well plates. After 96 h, cells were harvested for flow cytometry analysis.

The Killing Assays of MAIT Cells On Tumor Cell Lines

Isolated MAIT cells by flow cytometry were treated with or without G-CSF for 24 h. Then, G-CSF mobilized and non-mobilized MAIT cells were co-cultured with K562 tumor cell lines at MAIT/K562 ratios of with 1:1. MAIT cells (2×10^5 cells per well) were stimulated with cytokines (IL-12/IL-18, PeproTech) and/or anti-CD3/CD28 beads (Thermo) in 96-well U plate in RPMI 1640 (Gibco, Invitrogen) added 10% FBS. After 96 h, tumor cell apoptosis was assessed by flow cytometry using 7-AAD and Annexin V staining.

GVHD Mouse Models

NOD-SCID-IL2Rγ-/- (NSG) mice were used to establish the xenogeneic GVHD model. All animals were bred and maintained under specific pathogen-free conditions in microisolator cages and used at 6–8 weeks of age in accordance with protocols approved by the local Ethics Committee. On day −1, mice received sublethal total body irradiation (180 cGy, X-ray), followed by intravenous injection of 2×10^6 human G-PBSCs via the tail vein. Human MAIT cells were assessed in peripheral blood, gut, liver, lung, and bone marrow at engraftment, GVHD onset, and day 7 post-GVHD. Mice were monitored two to three times weekly for survival, body weight, and GVHD clinical score. GVHD scoring was based on six parameters—skin integrity, weight loss, posture, activity, fur texture, and diarrhea—each graded 0 or 1, with a maximum score of 6.

CXCL16 Blocking in NSG Mice

In vivo CXCL16 blocking was done by injecting i.e. 100 μg of anti-CXCL16 antibody (R&D, AF976) or PBS twice within a week after transplantation. Mice were monitored for survival, weight, and acute GVHD score was evaluated two or three times a week. The endpoint was defined as: (1) natural death; (2) mice showing 20% weight loss and clinical symptoms of GVHD; (3) 24 days after transplantation.

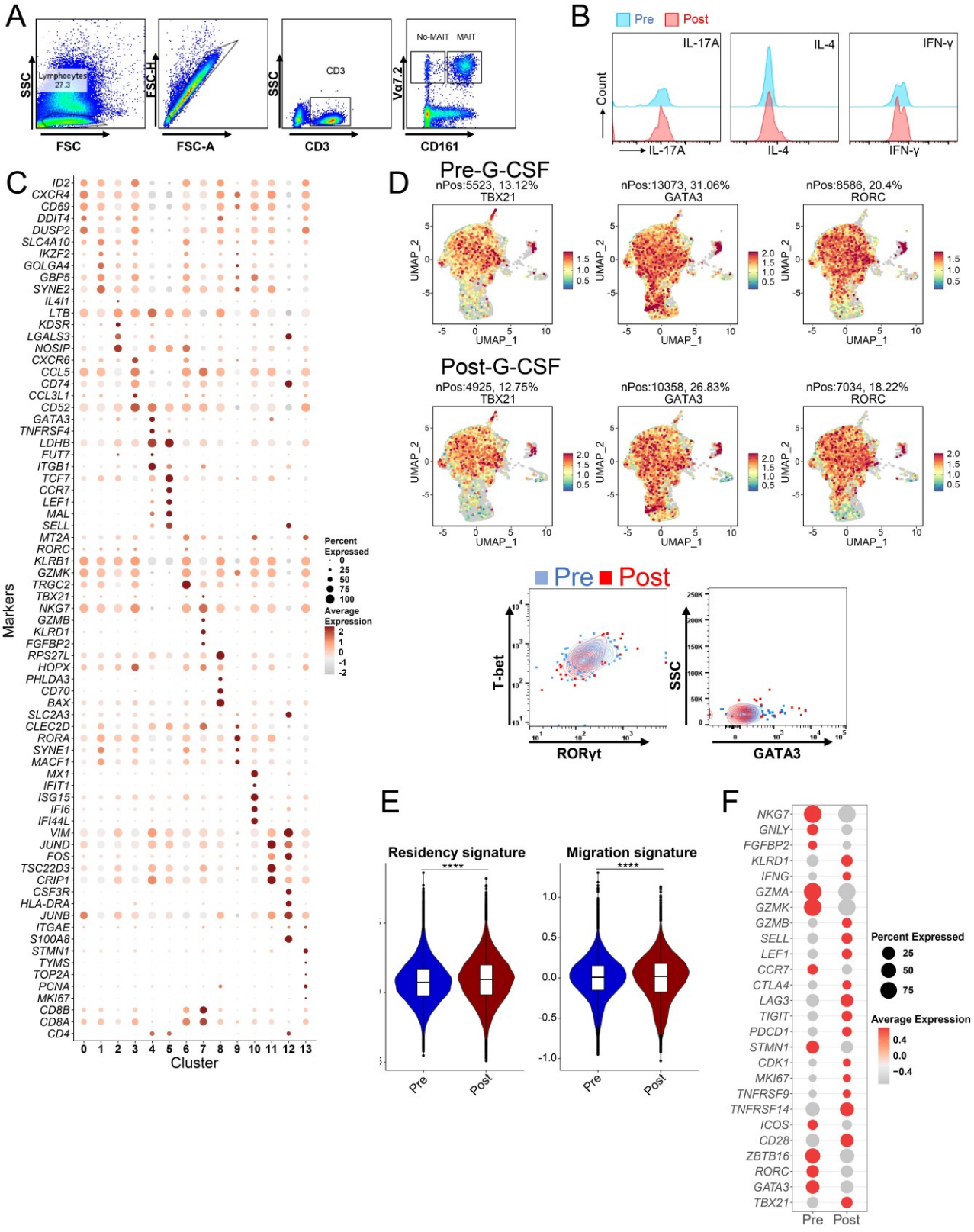
Immunofluorescence

Intestinal tissues from aGVHD NSG mice were processed for immunofluorescence analysis. Paraffin-embedded samples were sectioned into 3–4 μm slices and mounted on polylysine-coated slides. Paraffin sections were baked overnight at 58°C, dewaxed in xylene, rehydrated through a graded series of ethanol, quenched for endogenous peroxidase activity in 0.3% hydrogen peroxide for 15 mins. Antigen retrieval was performed by high-pressure cooking in citrate buffer (pH=6.0) for about 20 mins, then allowed to cool to room temperature, blocking the nonspecific antibody binding sites in 5% normal goat serum for 2 hrs. Primary antibodies, including CD3 (1:100, Abcam, ab135372, UK), CD4 (1:50, ZSGB-BIO, ZM-0418, China), CD8 (1:100, CST, CST70306, USA), and CD161 (1:100, Abcam, ab197979, UK), were applied at room temperature for 1 hour. The EnVision+ System-HRP (AEC) (K4005, Dako, Denmark) was used to detect the primary antibodies, followed by hematoxylin counterstaining (Sigma-Aldrich, USA). TMA slides were scanned using an automated microscope, and quantitative analysis was performed with Image-Pro Plus software (IPP, Media Cybernetics, USA).

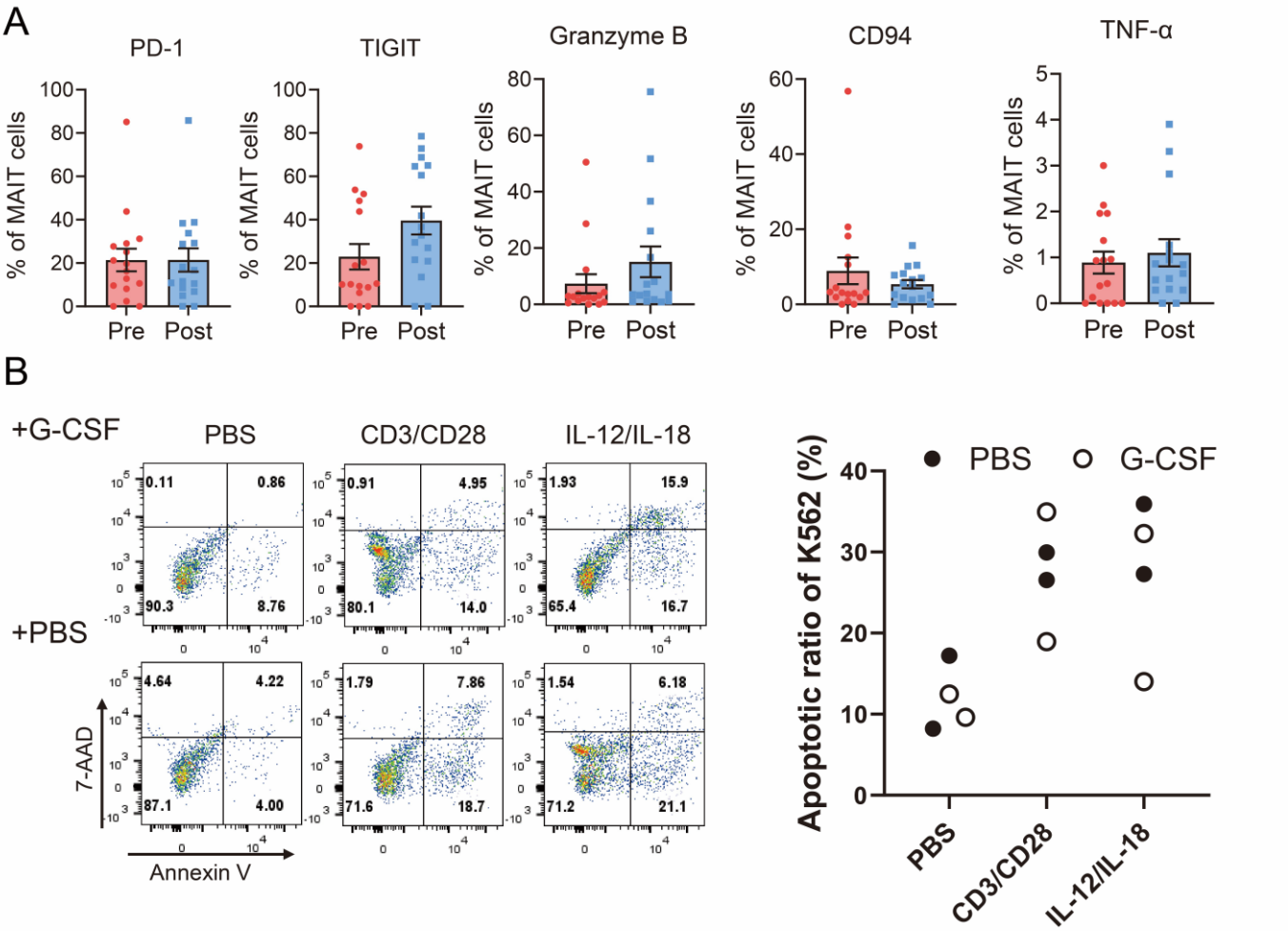
Calculation of predictive scores

The frequencies of MAIT cells and their functional subsets (CD114+, IL-17A+, IL-4+, and CXCR6+ MAIT cells) were quantified by multicolor flow cytometry, with all samples processed using consistent staining, acquisition, and analysis protocols. To determine robust thresholds, we applied bootstrap resampling (1000 iterations) to the clinical cohort, calculating the median proportion of each subset per iteration; the mean of these medians was defined as the bootstrap median threshold. For each patient, a binary score was assigned per subset: 1 point if the proportion was equal to or above the bootstrap median, and 0 if below. The sum of these five binary scores yielded a composite score ranging from 0 to 5.

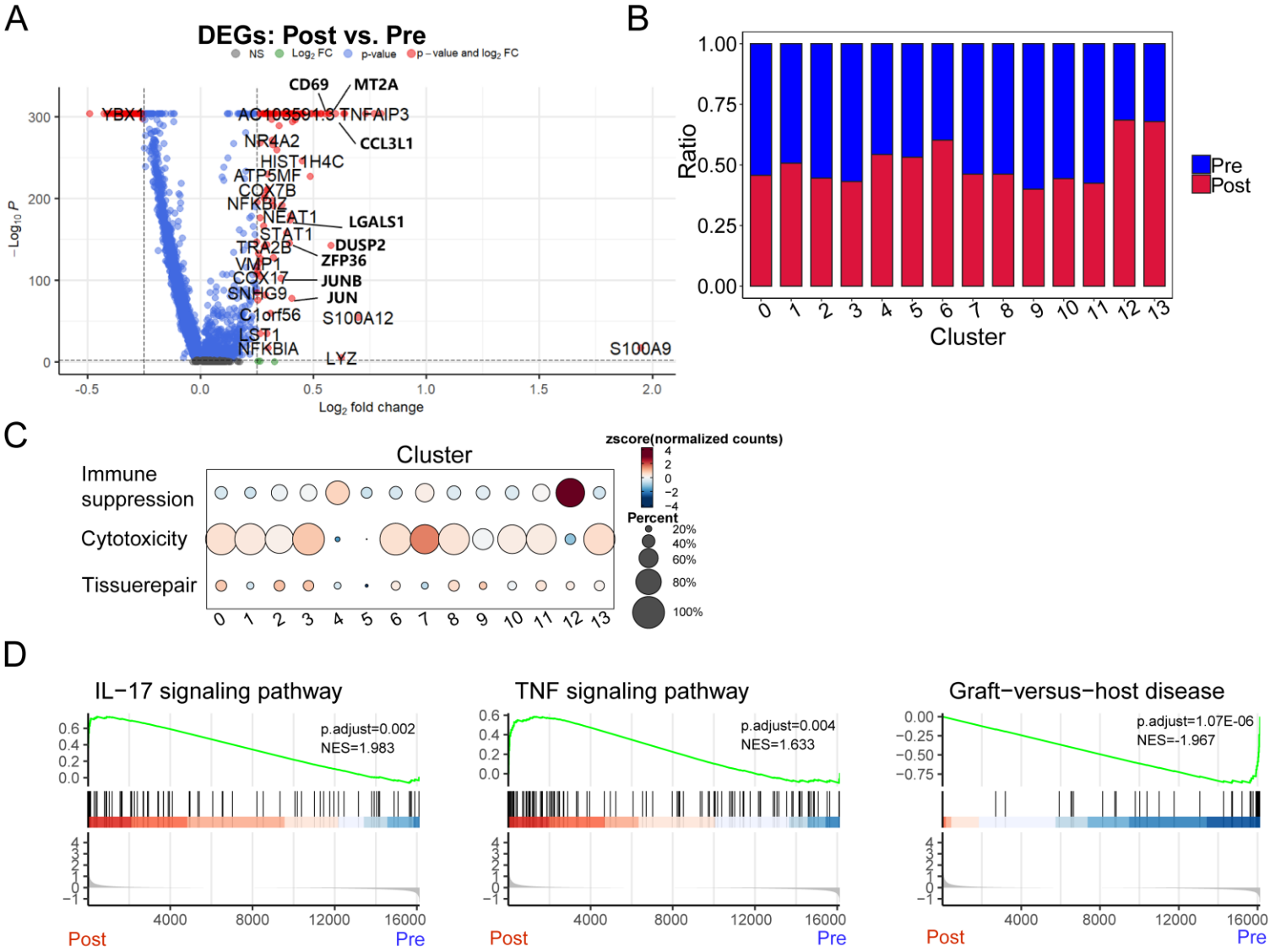
Supplementary Figures



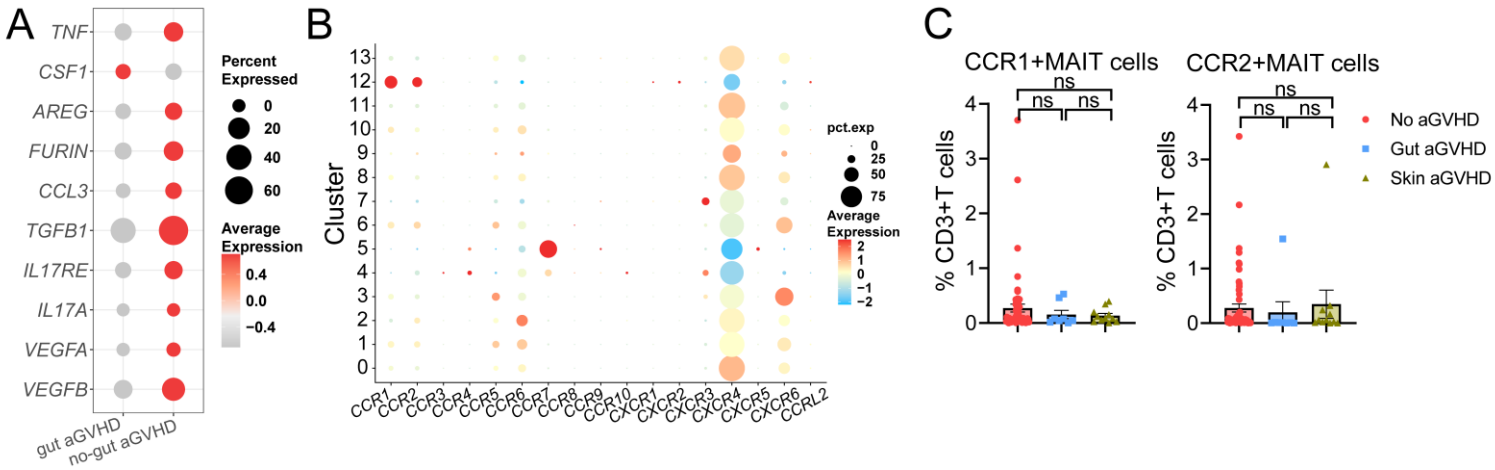
Supplementary Figure 1. Effect of G-CSF on functional molecule expression in MAIT cells. **(A)** Gate strategy of human MAIT cell. **(B)** Representative flow cytometry showing the expression levels of IL-4 (*P*=0.041), IFN-γ and IL-17A in MAIT cells from donor PBSCs (n=16) before and after G-CSF-mobilization. **(C)** Dot plot displaying key marker genes for each MAIT cell cluster among the differentially expressed genes (DEGs). Dot size and color intensity represent the proportion of expressing cells and expression levels, respectively. **(D)** UMAP plot and Flow cytometry showing the expression of transcription factors T-bet, GATA3 and RORγt for each cell. **(E)** Violin Box plot showing migration and residency signatures in MAIT cells before and after G-CSF-mobilization (****P value < 0.0001). **(F)** Dot plot showing the expression of functional molecules in MAIT cells before and after G-CSF mobilization.



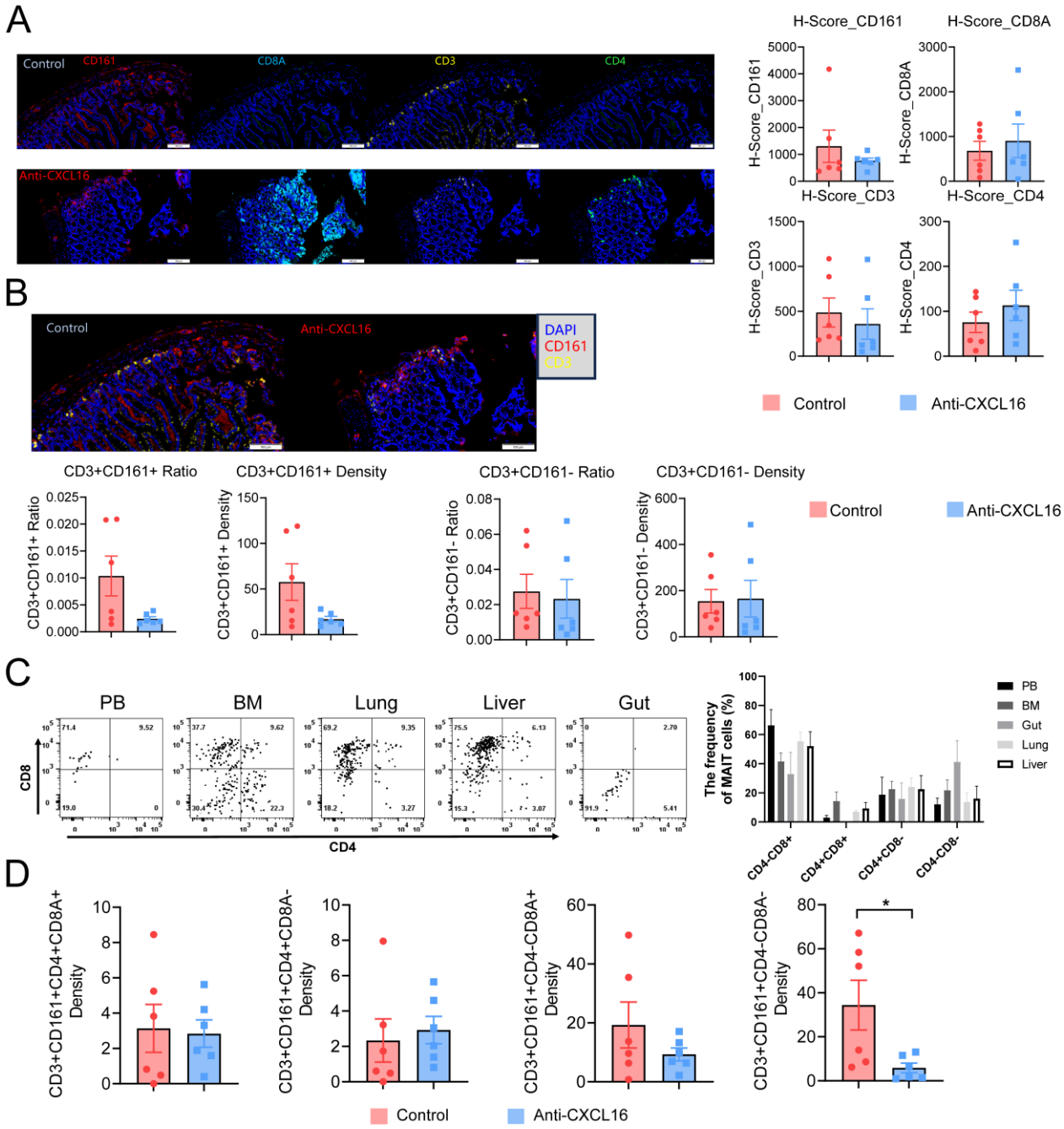
Supplementary Figure 2. (A) Flow cytometric analysis of expression of functional molecules in MAIT cells from donor PB (n=16) before and after G-CSF-mobilization. **(B)** Representative flow cytometry results showing the effect of G-CSF mobilization on the cytometry activity of MAIT cells against K562 cell line, following stimulation by cytokine (IL-12/IL-18) and TCR (CD3/CD28) pathway.



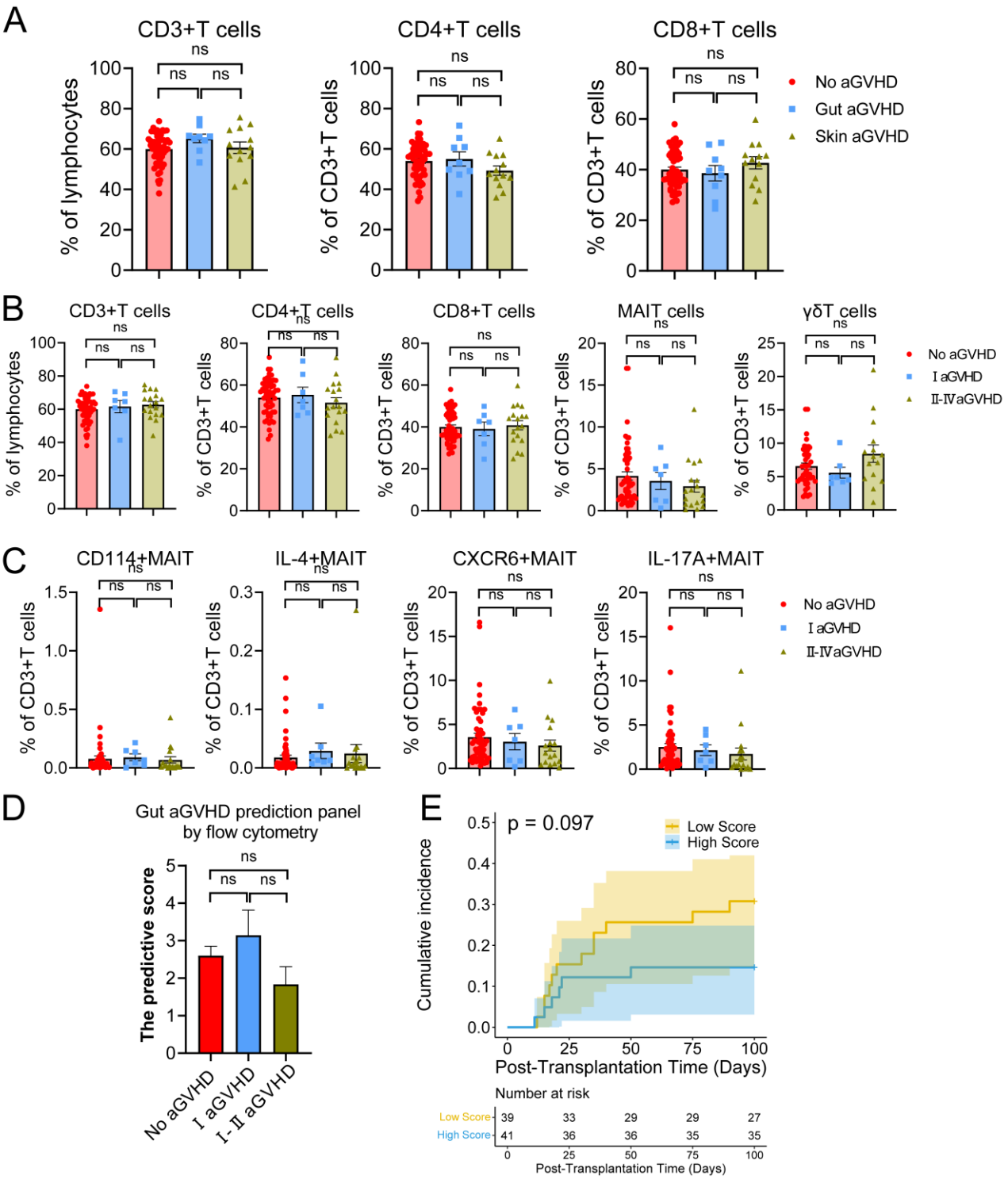
Supplementary Figure 3. The effect of G-CSF on MAIT cells within Cluster 12 expressing CSF3R. (A) EnhancedVolcano showing the DEGs of MAIT cells between unstimulated MAIT and G-CSF mobilized MAIT cells. DEGs were detected by differential expression analysis (two-sided Student's t-test). Each dot represents a single gene. **(B)** The proportion of MAIT cell clusters before and after G-CSF mobilization. **(C)** Dot plot showing functional signatures distribution in MAIT cells clusters. **(D)** Gene set enrichment analysis (GSEA) showing the major enriched terms of DEGs before and after G-CSF administration.



Supplementary Figure 4. The characteristics of tissue repair and chemotactic migration of MAIT cells may be involved in its anti-intestinal GVHD effect. (A) Dot plot showing the average expression of cell tissue repair-related genes on G-MAITs between groups of gut aGVHD (n=1) and no gut aGVHD (n=3). **(B)** Dot plot showing the expression of chemokine receptor genes on MAIT clusters. **(C)** Flow cytometric analysis showing chemokine receptor expression in G-MAIT cells between groups of gut aGVHD (n=9), skin aGVHD (n=13) and no aGVHD (n=55).



Supplementary Figure 5. Distribution of MAIT cells in G-grafts within NSG mouse transplantation models. (A) Single-marker immunofluorescence and histological score (H-score). **(B)** Immunofluorescence showing the ratios and densities of CD3+CD161+ and CD3+CD161- cell populations in the intestinal tissues of control and anti-CXCL16 groups. Density = Number of positive cells / Tissue area (mm^2). **(C)** Flow cytometric analysis showing CD4 and CD8 expression of MAIT cells in PB, BM and recipient tissues post- transplantation. **(D)** Immunofluorescence showing population densities in the intestinal tissues of control and anti-CXCL16 groups.



Supplementary Figure 6. The relationship between predictive signatures in MAIT cells and aGVHD severity. (A) Flow cytometry analysis showing the frequency of T cells and subsets among groups of gut aGVHD (n=9), skin aGVHD (n=13) and no aGVHD (n=55). **(B)** Flow cytometry analysis showing the frequency of T cells and subsets among groups stratified by aGVHD grade (I aGVHD: n=7; II–IV aGVHD: n=18). **(C)** GVHD–related functional gene of MAIT cells were compared between patients stratified by aGVHD grade. **(D)** The GVHD scores were compared between patients stratified by aGVHD grade. **(E)** Kaplan-Meier analysis showing

the incidence of II–IV aGVHD in high and low score from predictive scores.