Supporting Information

Single cell multi-omics profiling of resting and activated pre-infusion CAR T products generated from healthy donor and ALL patients

Zhiliang Bai[1,2], Stefan Lundh[6], Dongjoo Kim[1], Steven Woodhouse[8], David M. Barrett[9,10], Regina M. Myers[9], Stephan A. Grupp[9,10], Marcela V. Maus[11,12], Carl H. June[5,6,7], Pablo G. Camara[8], J. Joseph Melenhorst[5,6,7]*, Rong Fan[1,3,4]*

[1] Department of Biomedical Engineering, Yale University, New Haven, Connecticut 06511, USA

[2] State Key Laboratory of Precision Measurement Technology and Instrument, Tianjin University, Tianjin 300072, China

[3] Yale Stem Cell Center and Yale Cancer Center, Yale School of Medicine, New Haven, Connecticut 06511, USA

[4] Human and Translational Immunology, Yale School of Medicine, New Haven, Connecticut 06511, USA

[5] Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[6] Center for Cellular Immunotherapies, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[7] Parker Institute for Cancer Immunotherapy at University of Pennsylvania, Philadelphia, PA 19104, USA

[8] Department of Genetics and Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA

[9] Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[10] Departments of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[11] Cellular Immunotherapy Program and Cancer Center, Massachusetts General Hospital, Charlestown, MA 02129, USA

[12] Harvard Medical School, Boston, MA 02115, USA

* These authors jointly directed this work

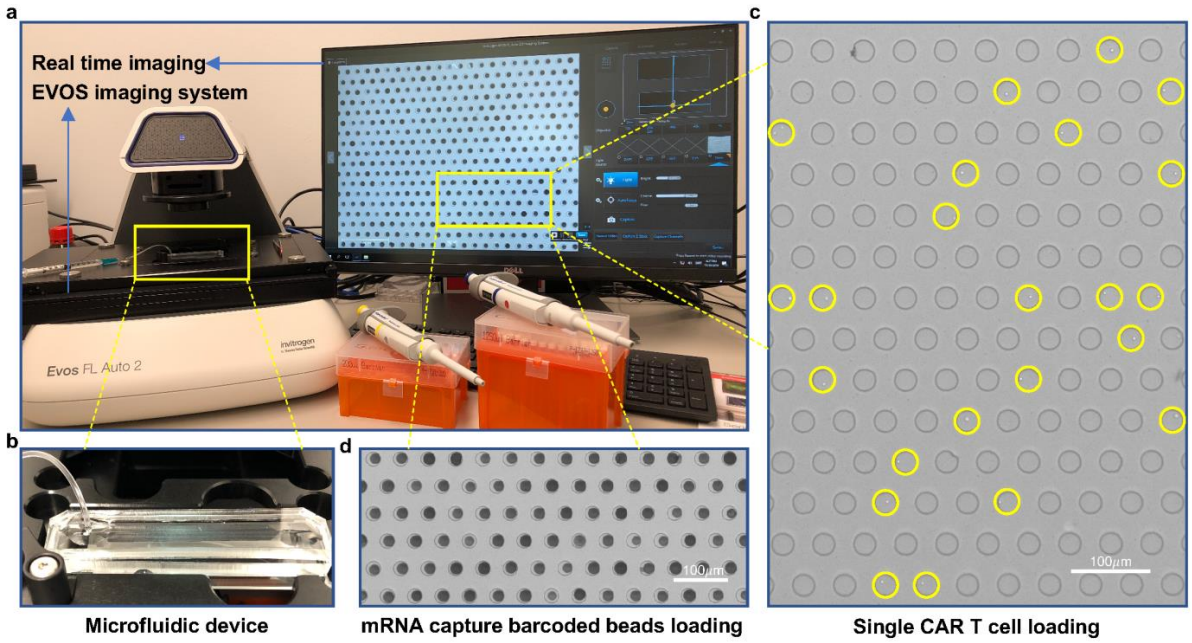
Correspondence to:

J.J.M. (mej@pennmedicine.upenn.edu)

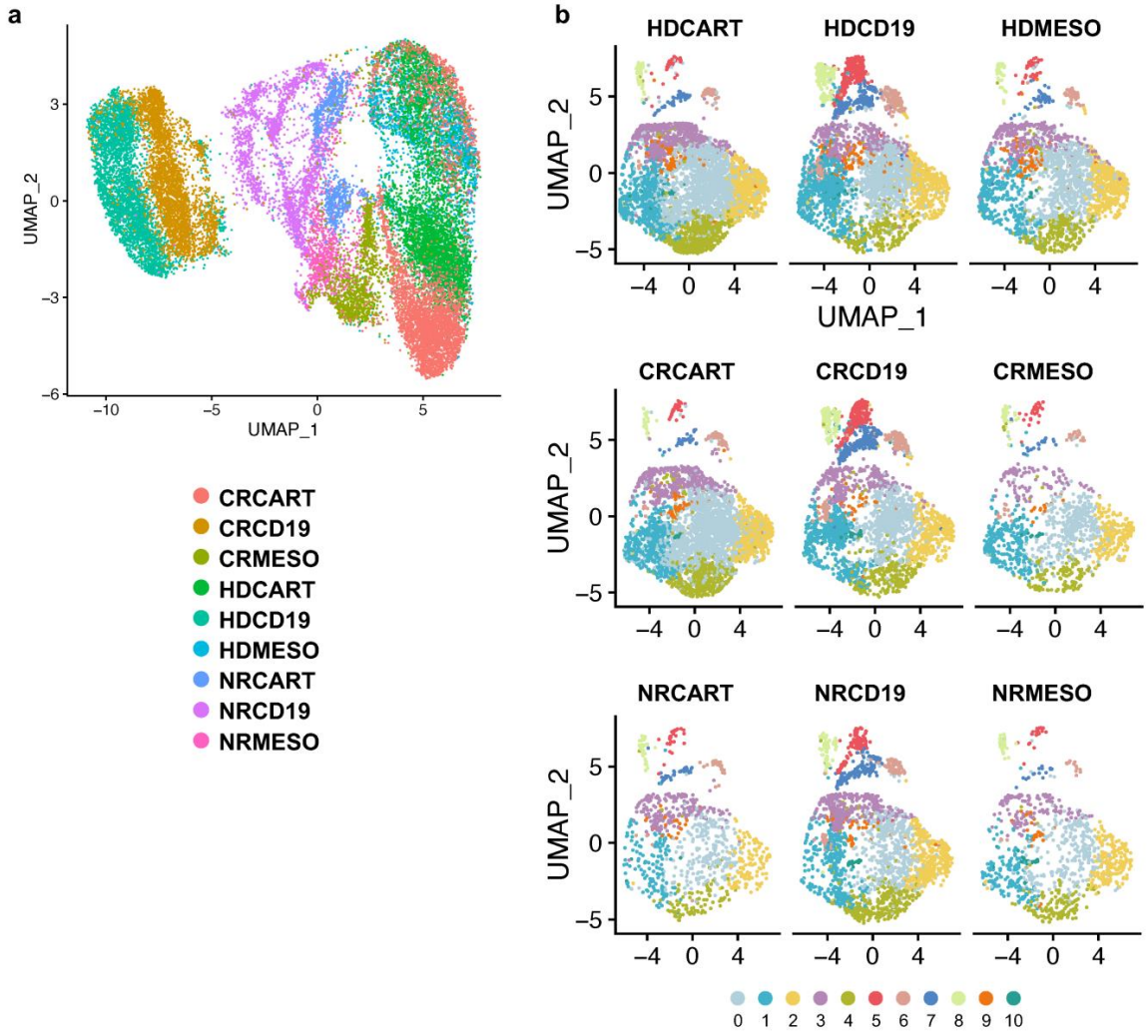
R.F. (rong.fan@yale.edu)

Supplementary Table 1. Patient demographics and clinical data

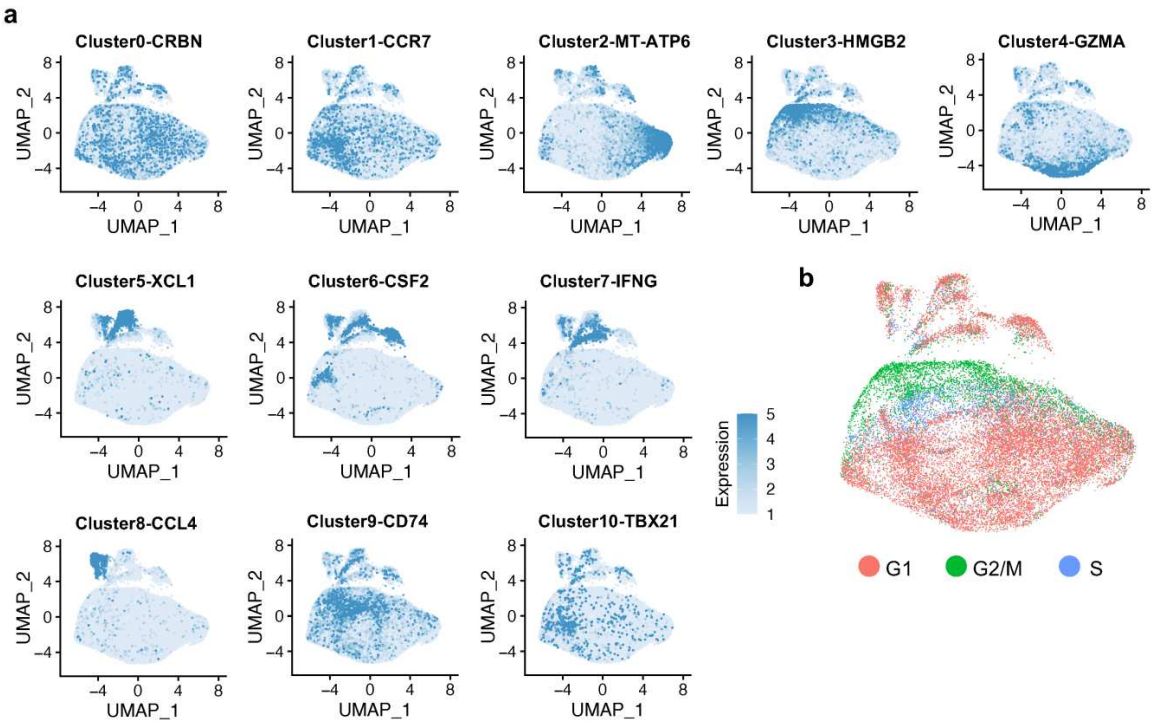
ID	Infusion date	Age	Sex	Response to CART19	CRS grade	B cell aplasia (BCA)	BCA duration	Relapse	Last contact days
CR	4/16/13	9.6	F	MRD negative CR	2	Yes	72 months	No	2171
NR	5/9/13	16.2	M	No response	2	N/A	N/A	N/A	79
HD	N/A	41	M	N/A	N/A	N/A	N/A	N/A	N/A



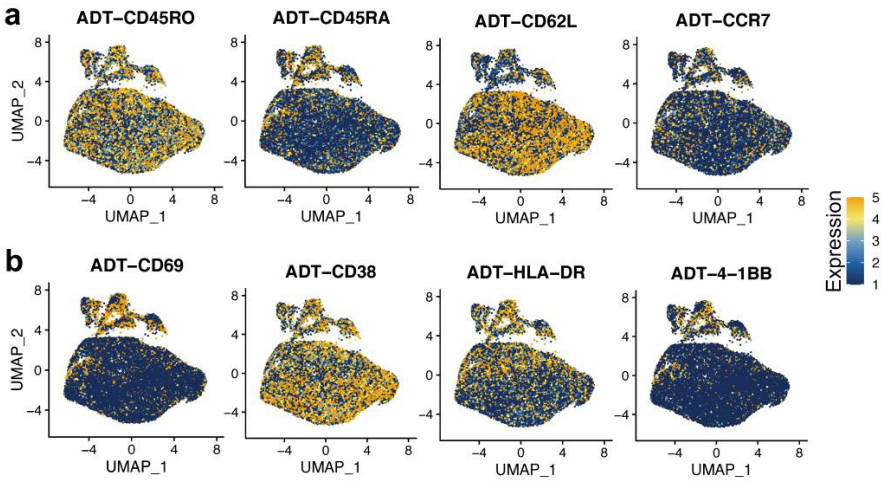
Supplementary Figure 1. Experimental setup and microfluidic device for single cell isolation and mRNA capture using scFTD-seq. (**a**) The setup to operate scFTD-seq includes a 10" tubing connected with a BD syringe, pipettes and a EVOS FL Auto system for real time imaging during cell and bead loading. (**b**) PDMS-fabricated microfluidic device. (**c**) Optical image of cell loading and capture. Yellow circles indicate microwells that successfully capture single cell. (**d**) Optical image of barcoded beads loading and capture.



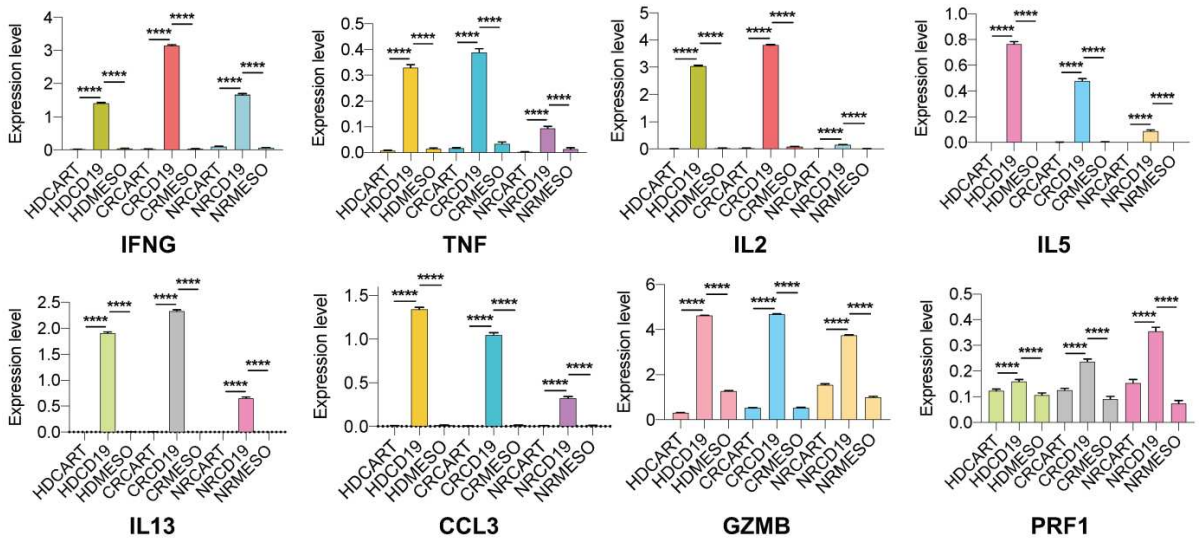
Supplementary Figure 2. Sample- and condition-related batch effects are corrected using canonical correlation analysis. (**a**) Integrated UMAP before the correction. Clusters were evidently aggregated by individual sample. (**b**) UMAP plot split by sample after batch correction. Single cells were clustered by their biological characteristics, and all the samples were broadly distributed in every identified clusters.



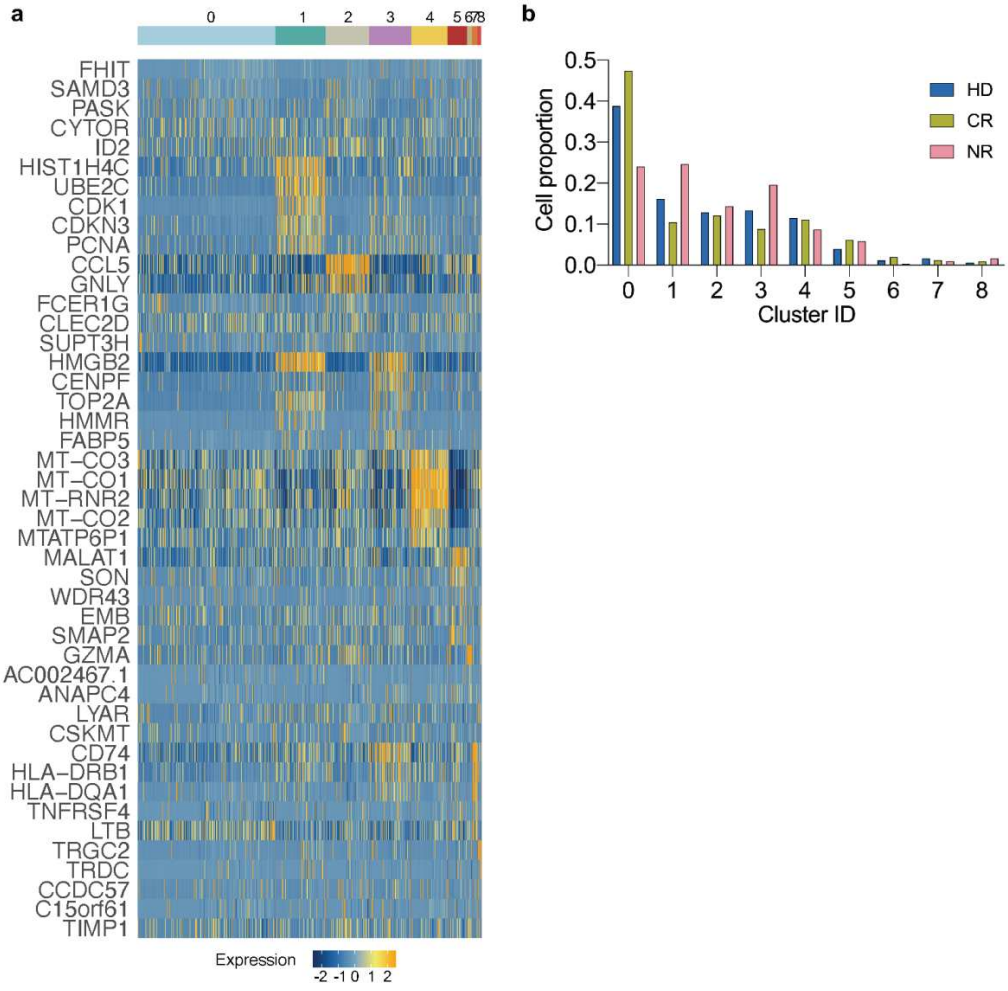
Supplementary Figure 3. Gene marker of identified clusters and cell cycle signatures. (**a**) The distribution of marker gene of each identified cluster. Color bar represents normalized expression level. (**b**) Cell cycle expression pattern.



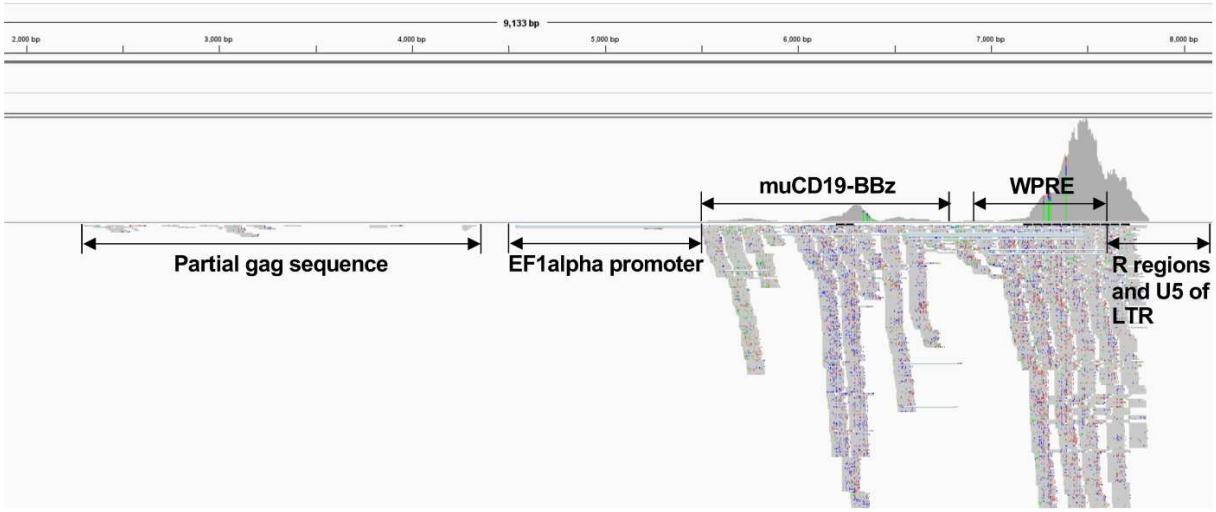
Supplementary Figure 4. The distribution of surface protein expression revealed by CITE-seq. (**a**) Memory markers (**b**) Activation markers. Color bar represents normalized expression level.



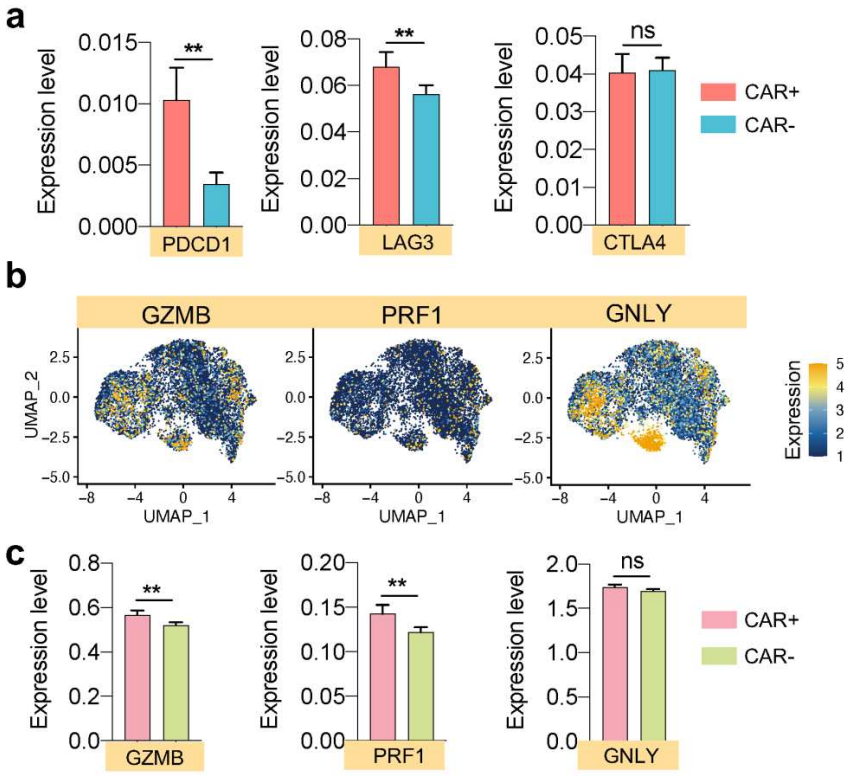
Supplementary Figure 5. The comparison of functional gene expression between each condition, including *IFNG*, *TNF*, *IL2*, *IL5*, *IL13*, *CCL3*, *GZMB* and *PRF1*. Bar plots show mean±s.e.m. The *P* values were calculated with two-tailed Mann-Whitney test. **** $P < 0.0001$.



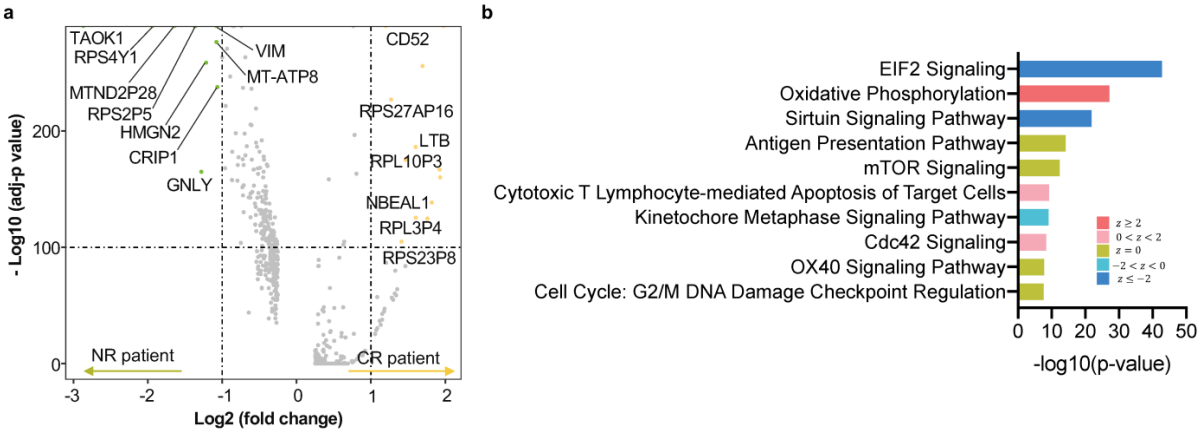
Supplementary Figure 6. DEGs and cell proportion of each identified cluster in unstimulated CAR T cells. (**a**) Heatmap of top5 DEGs identified in each cluster. (**b**) Cell proportion in each cluster grouped by sample resources.



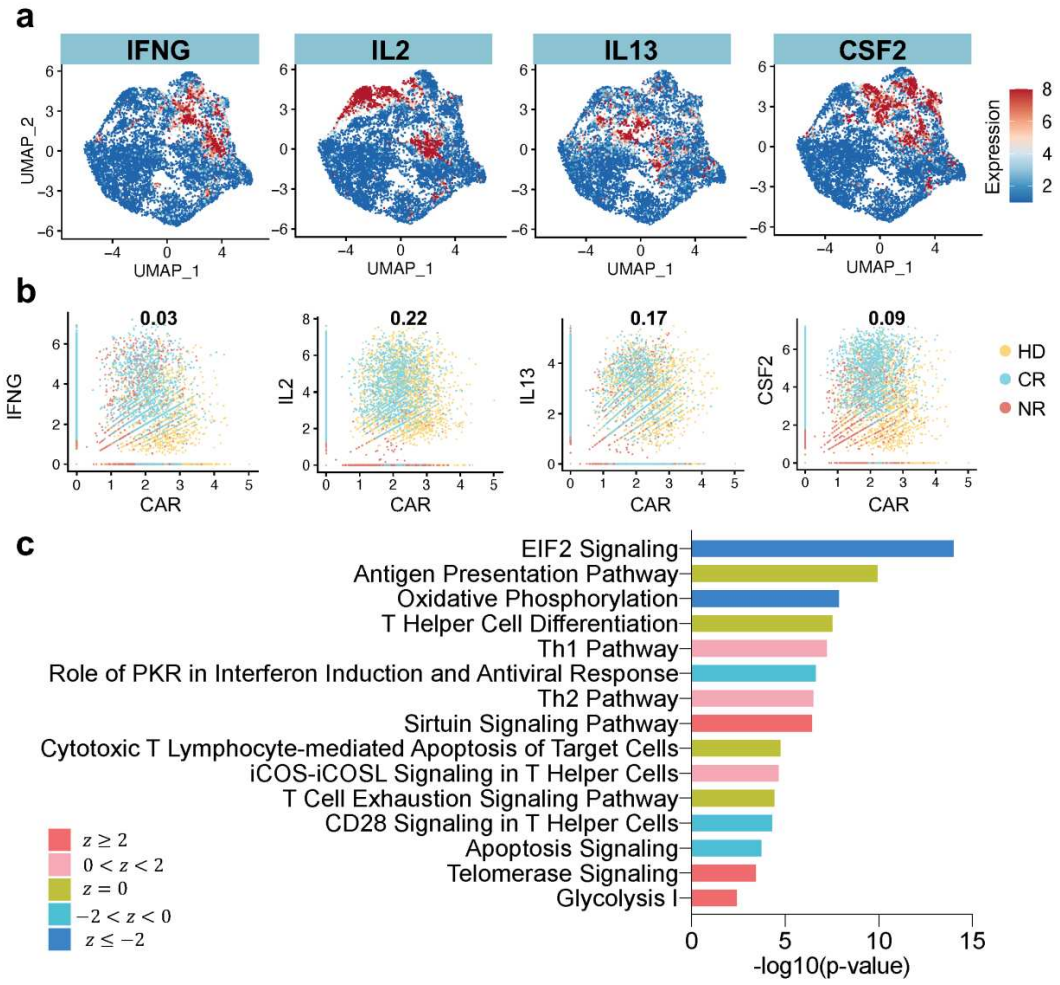
Supplementary Figure 7. Identify the expression of the lentiviral vector elements in the CAR construct using Integrative genomics viewer (IGV) tool. WPRE represents Woodchuck Hepatitis Virus (WHP) post-transcriptional regulatory element. LTR represents long terminal repeat viral region.



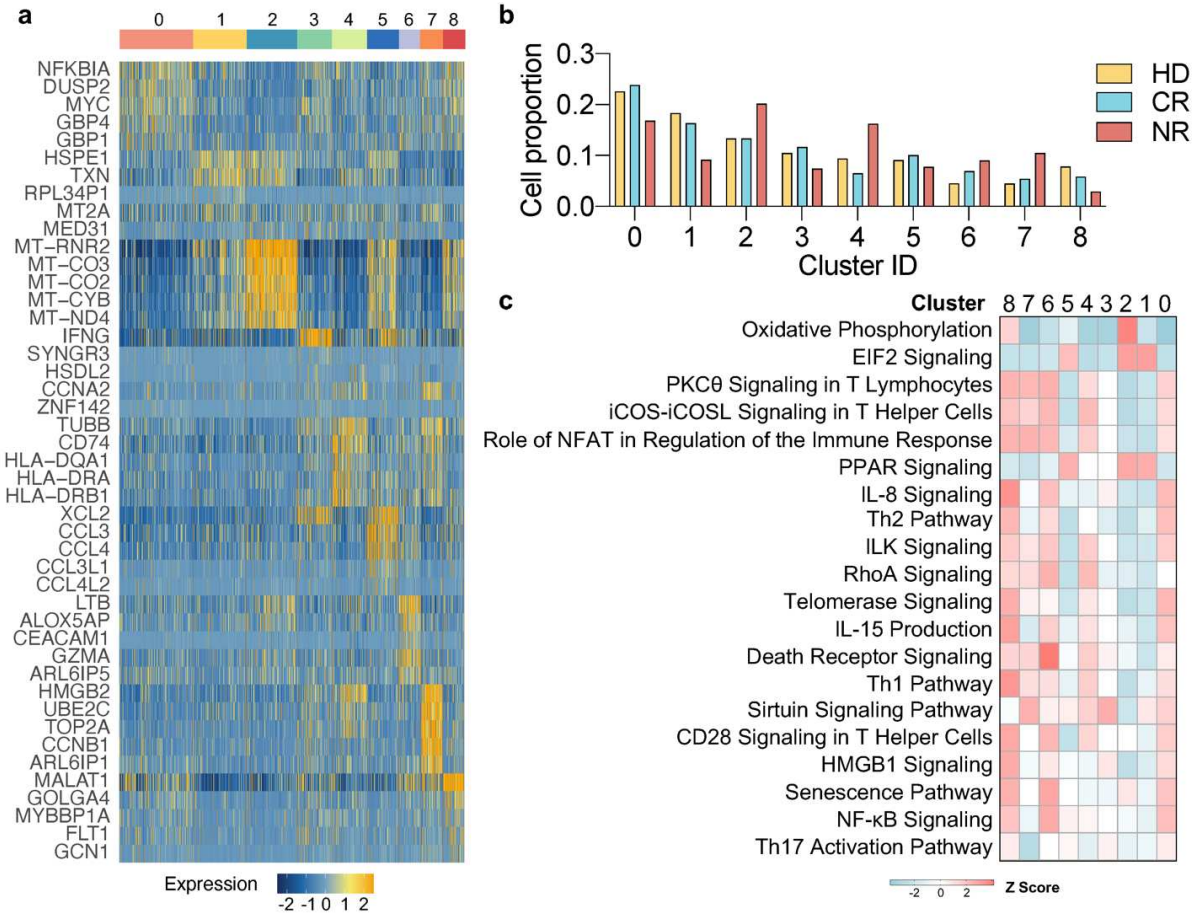
Supplementary Figure 8. Dysfunctional and cytotoxic gene expression in CAR+ and CAR- unstimulated products. (**a**) Comparison of well-known T cell dysfunction signature expression between CAR+ and CAR- cells. (**b**) The distribution of cytotoxic gene markers. Color bar represents normalized expression level. (**c**) Comparison of cytotoxic gene signature expression level between CAR+ and CAR- cells. Bar plots show mean±s.e.m. The *P* values were calculated with two-tailed Mann-Whitney test. ** $P < 0.01$, ns, not significant.



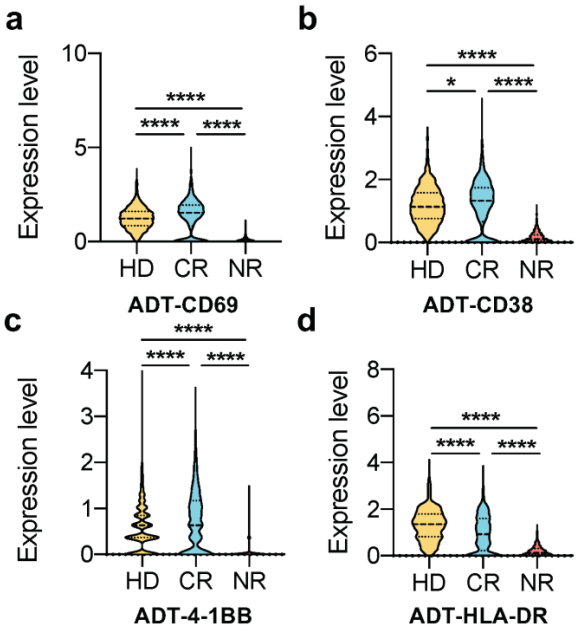
Supplementary Figure 9. DEGs and their corresponding signaling pathways upregulated in unstimulated CR versus NR sample. (**a**) Volcano plot of DEGs when comparing CR with NR sample. (**b**) Signaling pathways regulated by DEGs identified in (**a**). z score reflects the predicted activation level ($z < 0$, inhibited; $z > 0$, activated; $z \geq 2$ or $z \leq -2$ can be considered significant).



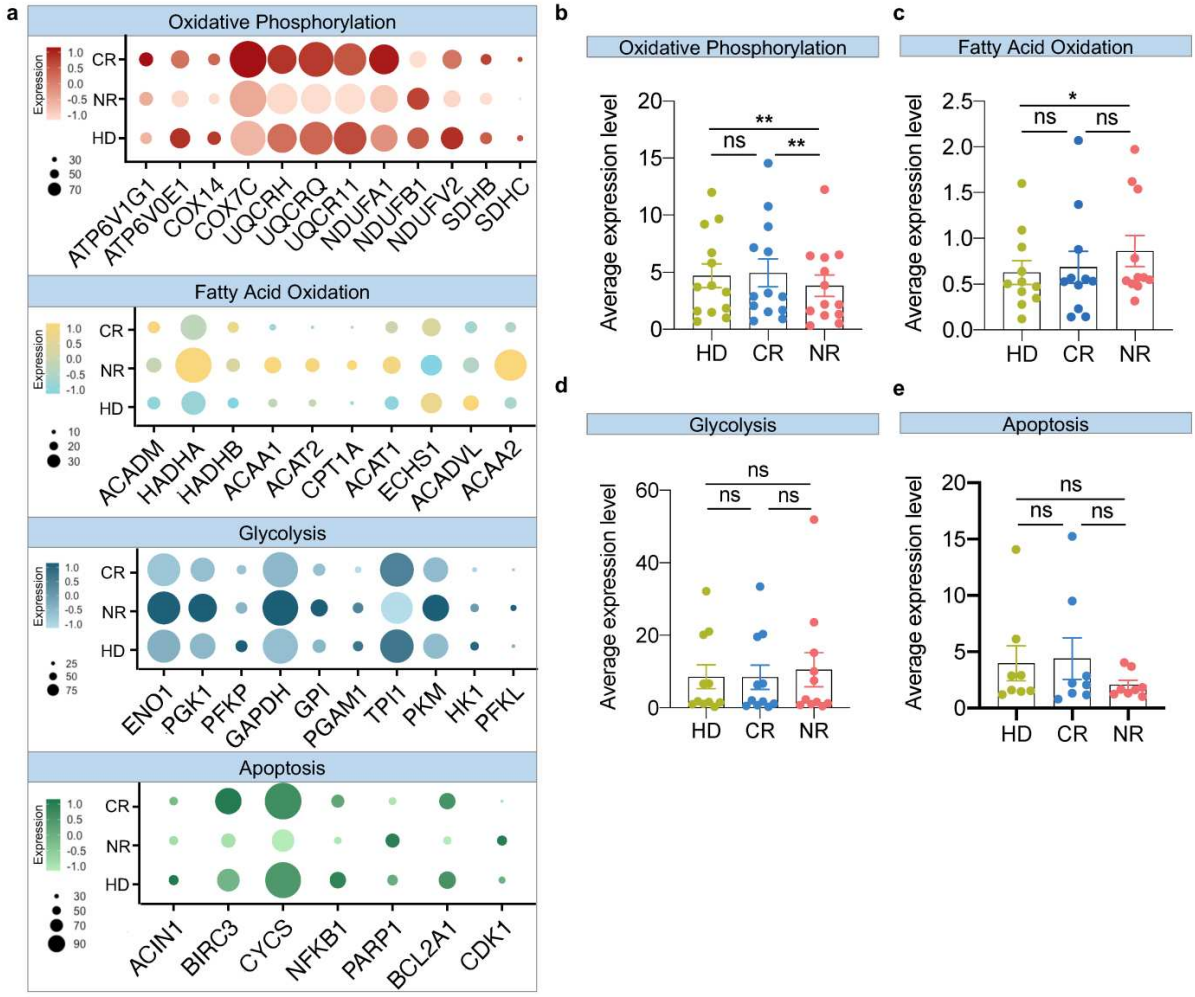
Supplementary Figure 10. Intrinsic characteristics of CAR+ cells upon CD19-specific stimulation. (**a**) Distribution of representative functional genes. Color bar represents normalized expression level. (**b**) The correlation between CAR expression and other functional genes. (**c**) The canonical pathways underlined in cluster 0 that has enriched CAR expression. *z* score reflects the predicted activation level ($z < 0$, inhibited; $z > 0$, activated; $z \geq 2$ or $z \leq -2$ can be considered significant).



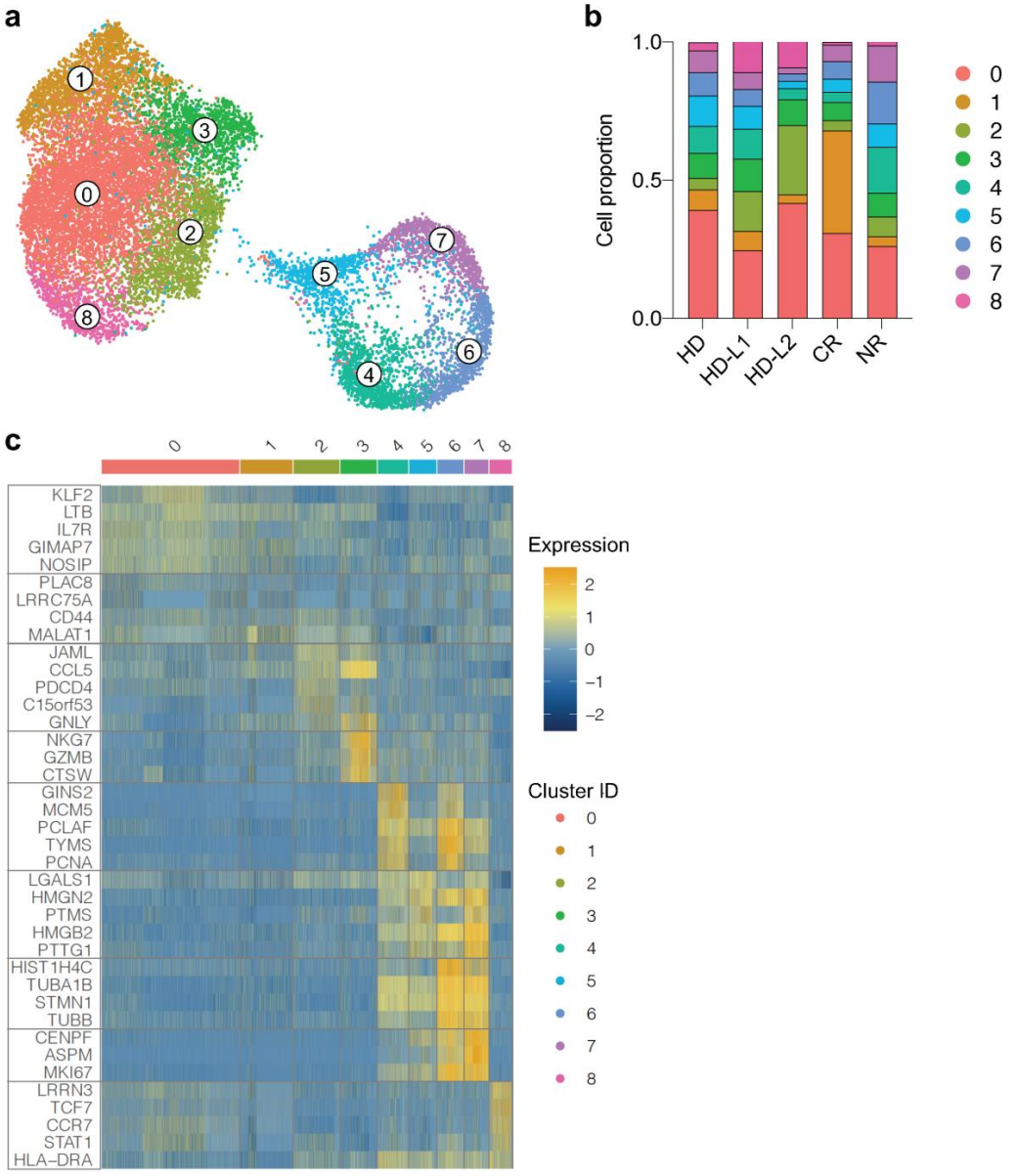
Supplementary Figure 11. DEGs and their corresponding signaling pathways of identified clusters in CAR+ cells upon CD19-specific stimulation. (**a**) Heatmap of top5 DEGs identified in each cluster in only activated CAR+ cells. (**b**) Cell proportion in each cluster separated by each group. (**c**) Comparisons of canonical pathways revealed in each cluster. z score reflects the predicted activation level ($z < 0$, inhibited; $z > 0$, activated; $z \geq 2$ or $z \leq -2$ can be considered significant).



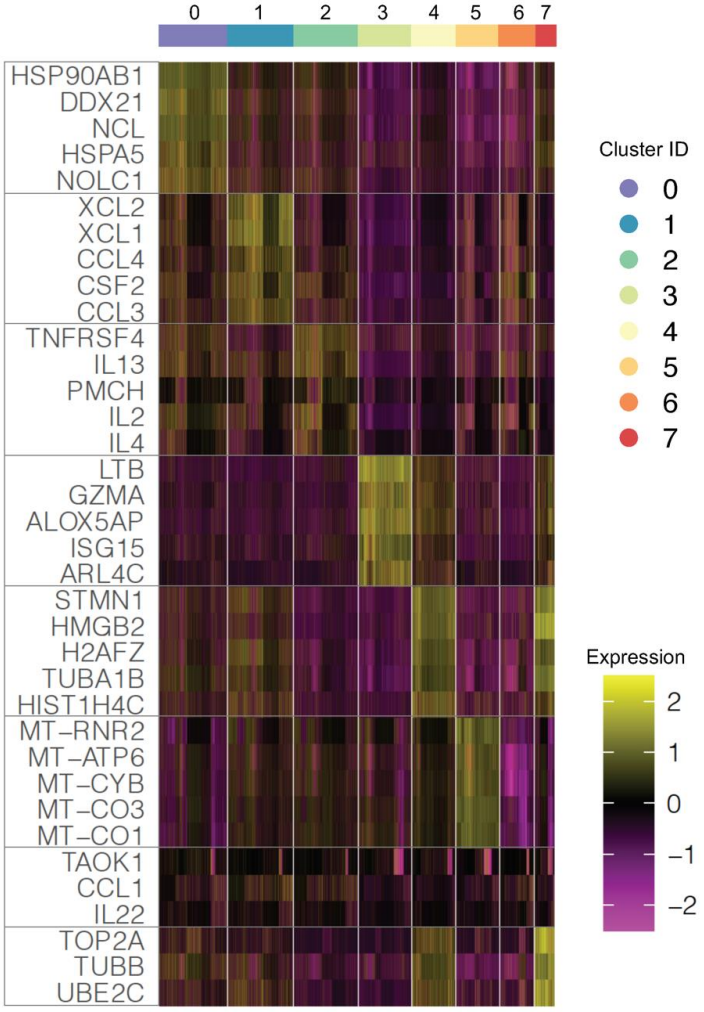
Supplementary Figure 12. Comparison of activation related surface protein level between resource groups. (**a**) The comparison of ADT-CD69 expression. (**b**) The comparison of ADT-CD38 expression. (**c**) The comparison of ADT-4-1BB expression. (**d**) The comparison of ADT-HLA-DR expression. The *P* values were calculated with two-tailed Mann-Whitney test. ∗*P*<0.05, ∗∗∗∗*P*<0.0001, ns, not significant.



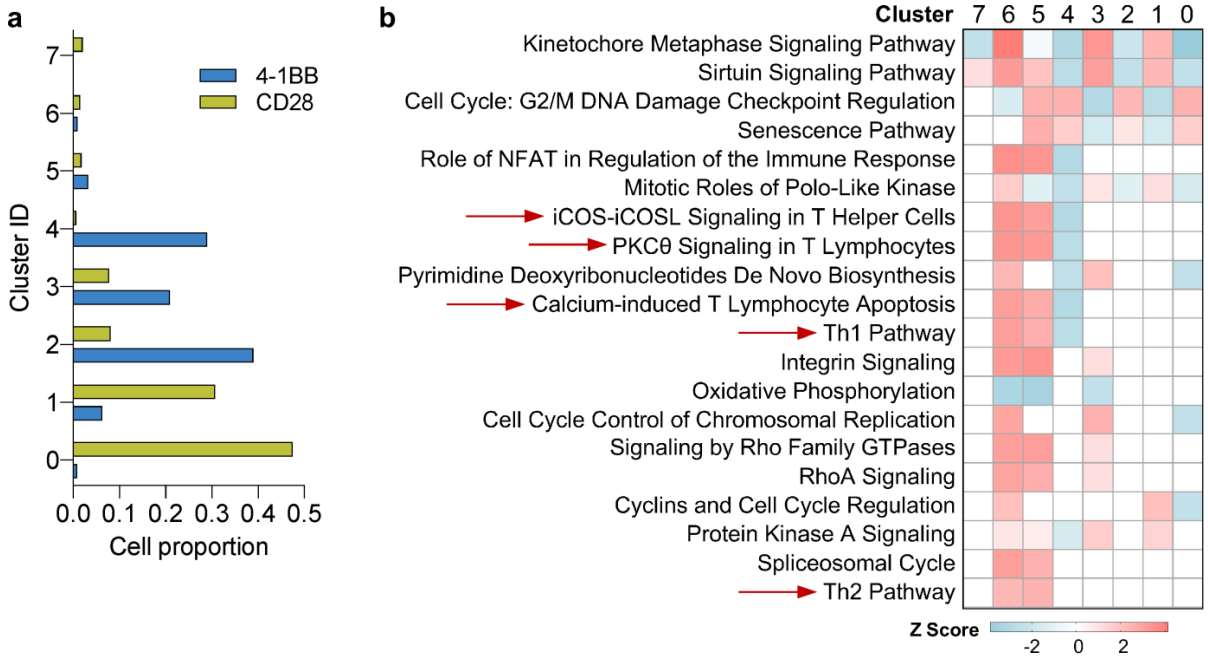
Supplementary Figure 13. Comparison of metabolic gene sets expression between activated HD, CR and NR group. (**a**) Dot plot of metabolic gene sets expression. The size of circle represents proportion of single cells expressing the gene, and the color shade indicates normalized expression level. (**b**) Comparison of oxidative phosphorylation of each group. (**c**) Comparison of fatty acid oxidation of each group. (**d**) Comparison of glycolysis of each group. (**e**) Comparison of apoptosis of each group. Scatter plot with bar shows mean±s.e.m. Symbols represent average expression value of each gene across all single cells. The *P* values were calculated with Wilcoxon signed-rank test. $* P < 0.05$, $** P < 0.01$, ns, not significant.



Supplementary Figure 14. Integrated clustering analysis of two more literature-derived 4-1BB donor CAR T scRNA-seq datasets at basal condition. (**a**) UMAP plot of pooled basal CAR T cells. (**b**) Cell proportion in each cluster. HD-L1 and HD-L2 represent literature-derived samples. (**c**) Heatmap of top5 DEGs identified in each cluster, from which we found that cells mainly clustered based on their inherent biological activities.



Supplementary Figure 15. Heatmap of top5 DEGs in each cluster that identified in Figure 6a.



Supplementary Figure 16. Clustering and pathway analysis of the integrated donor-derived 4-1BB CAR T and CD28 CAR T cells at basal state. (a) Cell proportion in each cluster identified in Figure 7a. (b) Comparisons of canonical pathways revealed in each cluster. The red arrows indicated pathways related to T cell activation and immune functions. z score reflects the predicted activation level (z<0, inhibited; z>0, activated; z≥2 or z≤-2 can be considered significant).