OMTM, Volume *23*

Supplemental information

Monitoring CAR T cell generation with a CD8-targeted lentiviral vector by single-cell transcriptomics

Filippos T. Charitidis, Elham Adabi, Frederic B. Thalheimer, Colin Clarke, and Christian J. Buchholz

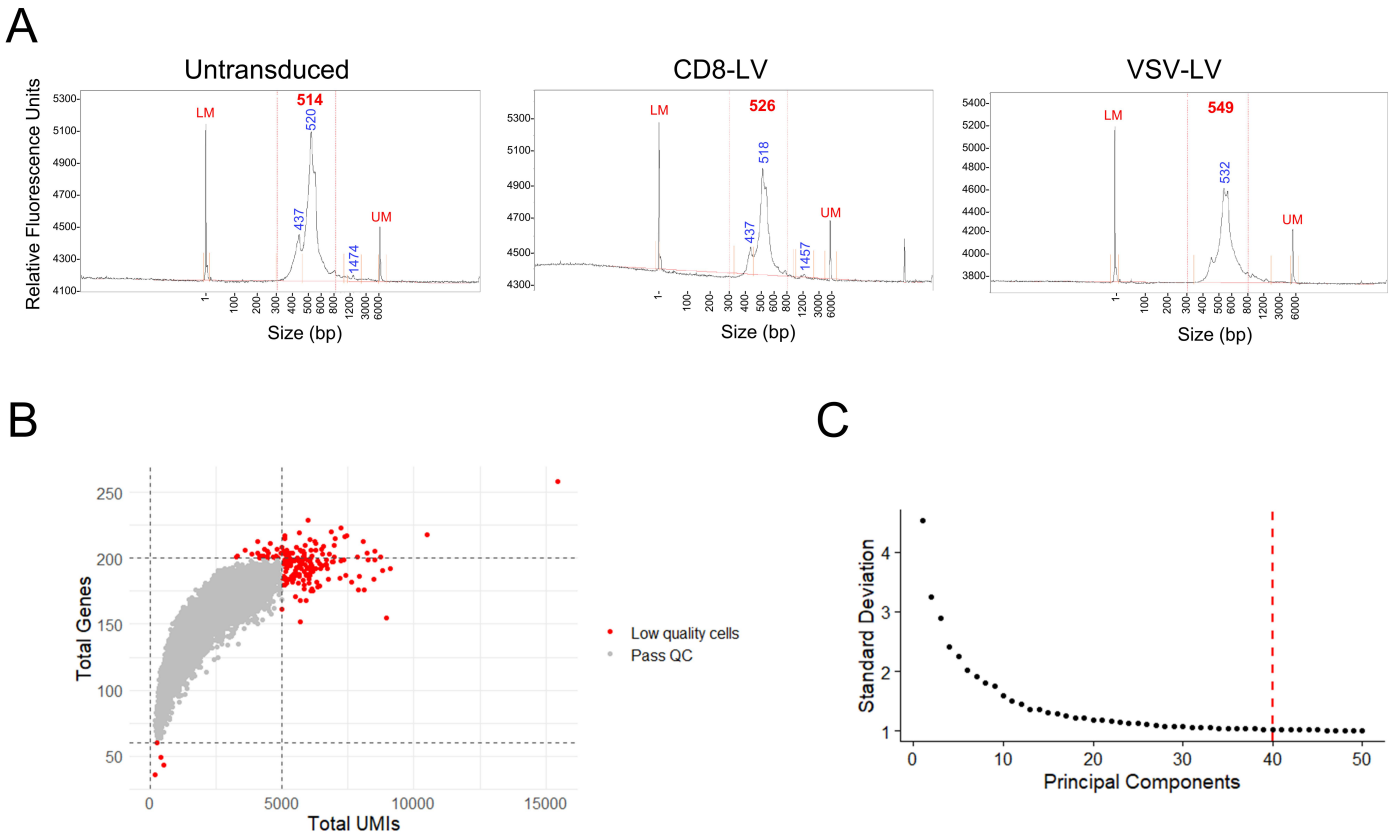

Figure S1: Quality control and pre-processing.
(**A**) Quality control of final sequencing libraries by Fragment Analyzer. The average sizes of the libraries were measured within the indicated area of 300-800 base-pairs (bp) and are shown in red numbers. Lower marker (LM) and upper marker (UM) fragments are indicated. (**B**) Overlay of cells from all groups are shown. Low quality cells (red) were filtered out. (**C**) Elbowplot selecting the top 40 principal components for UMAP analysis.

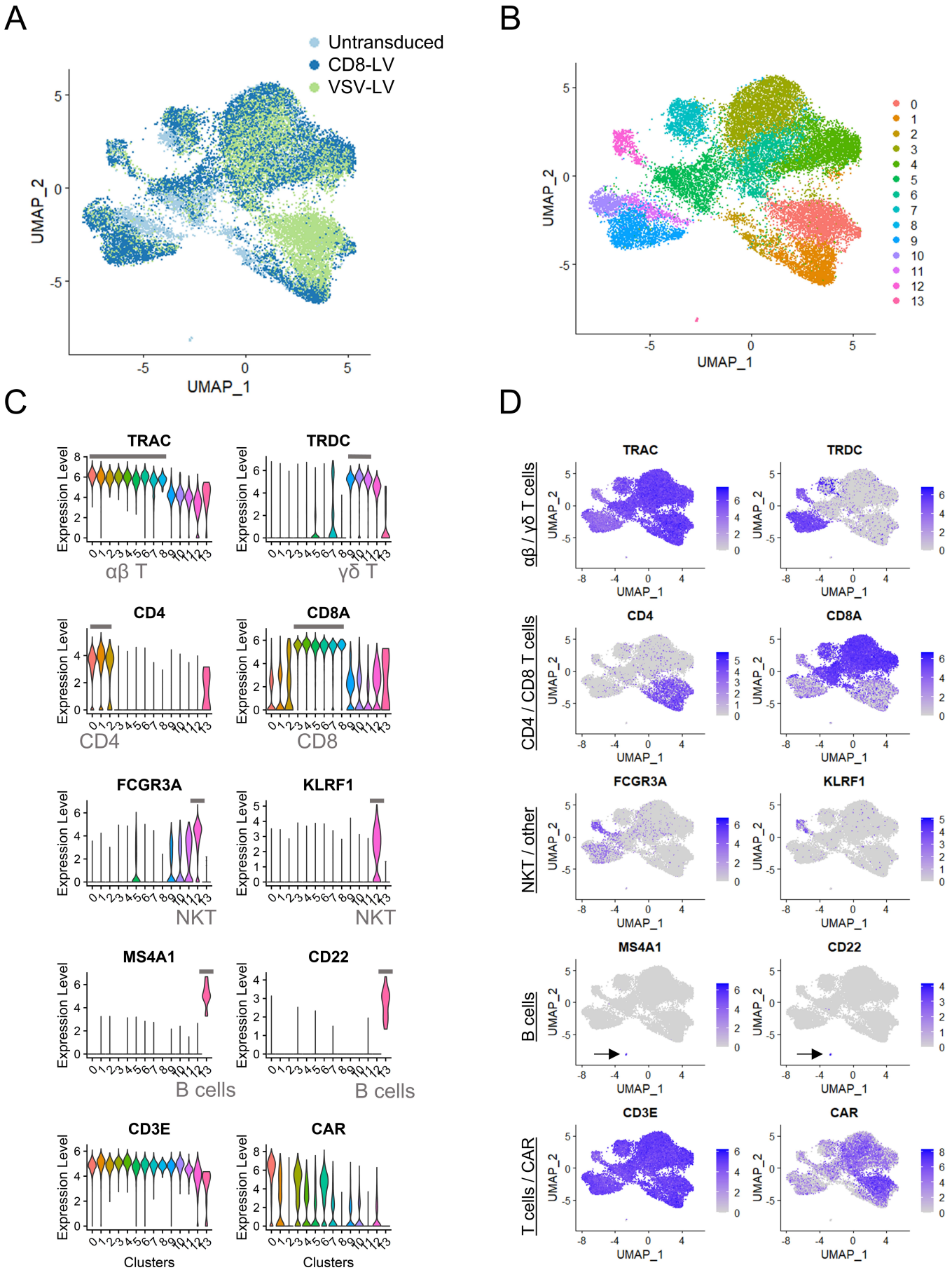

Figure S2: Clustering and identification of major immune subtypes in UMAP plots. (**A**) Overlapping merged samples in the UMAP plot. (**B**) Unsupervised clustering analysis and projection of clusters on UMAP plot, identified by the functions of *FindNeightbors* and *FindClusters.* (**C**) Expression of major T cell marker genes across the clusters. (**D**) UMAP plots of all samples merged, visualizing the localization of major immune cell types. Arrows indicate the location of B cells into the plots.

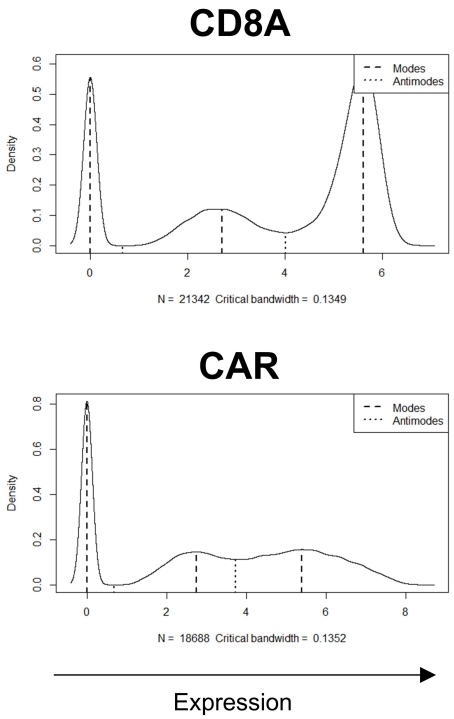

Figure S3: Multimodal analysis.
Multimodal analysis on density of data points across the normalized expression of *CD8A* or *CAR* genes, on total samples or transduced samples, respectively, calculated by the package *multimode*. Modes are shown in dashed lines, while antimodes in dotted lines.

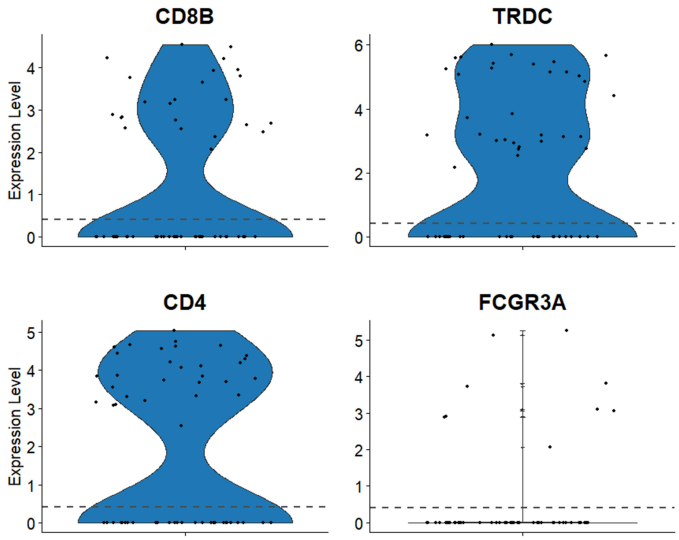

Figure S4: Violin plots of marker gene expression of potential off-target cells in CD8-LV sample.
Identification of cell types present in *CAR^{high} CD8A$^{neg/low}$* cells generated with CD8-LV, based on the expression of *CD8B*, *TRDC*, *CD4* and *FCGR3A* (referring to 1.26% of cells shown in Fig. 2D). Dashed lines indicate the lower cutoff for determining the cells positive for the plotted marker. Cell number for each cell type is shown in Table 2.

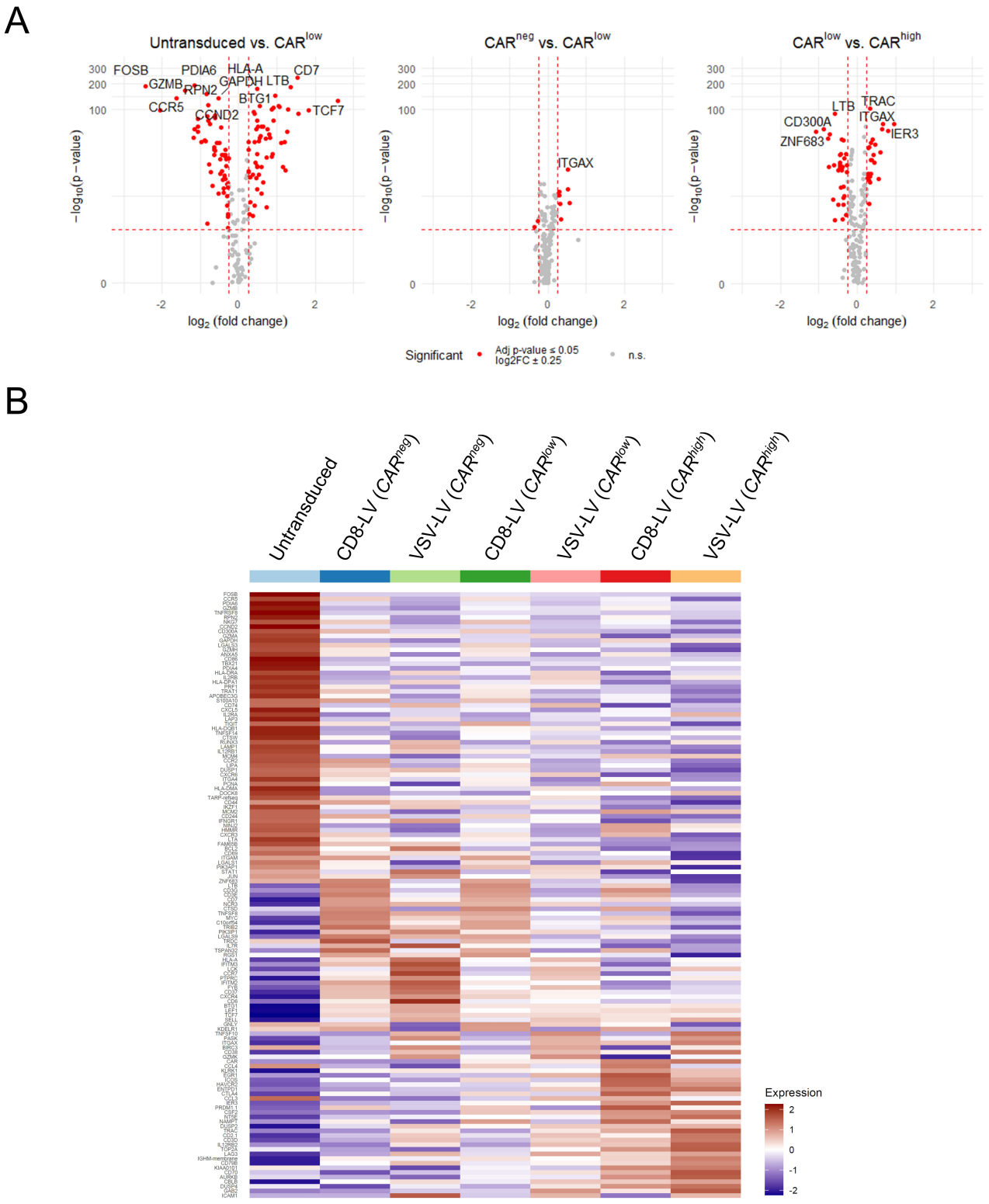

Figure S5. Comparisons of *CAR^{neg}*, *CAR^{low}* and *CAR^{high}* cells.
(**A**) Volcano plots of CD8+ CAR^{neg} versus CAR^{low} cells (left) and CD8+ *CAR^{low}* versus *CAR^{high}* cells (right). (**B**) Heatmap plot of upregulated genes between untransduced, *CAR^{neg}*, *CAR^{low}* and *CAR^{high}* identified by *FindAllMarkers* function in Seurat (log2FC ± 0.25, adj.p-value <0.05).

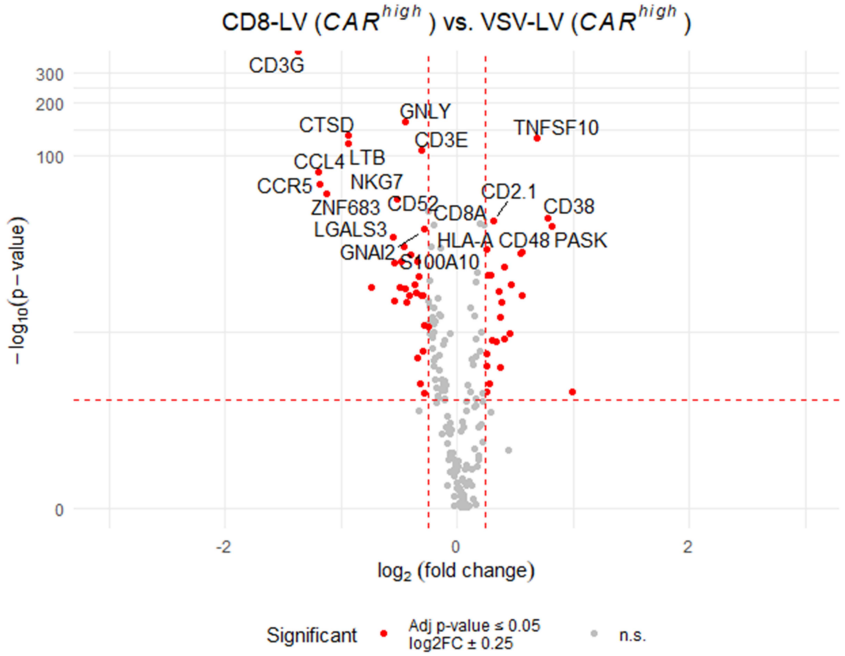

Figure S6. Volcano plot of comparing *CAR^{high}* cells generated by the two vectors.
Volcano plots of CD8+ *CAR^{high}* cells generated by either CD8-LV or VSV-LV showing the differentially expressed genes identified by *FindMarkers* function in Seurat.

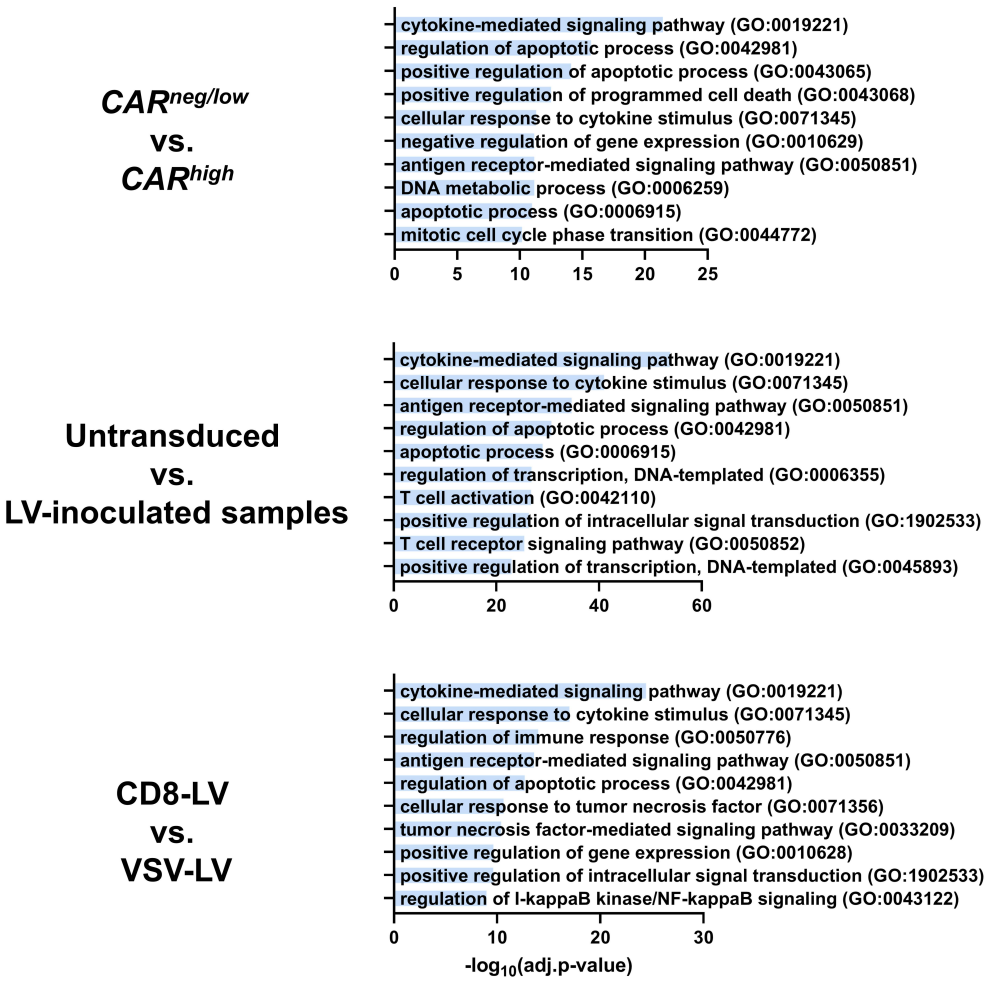

Figure S7. Gene set enrichment analysis.
Gene set enrichment analysis of differentially expressed genes identified in the indicated comparisons with the gene sets of Gene Ontology (GO) Biological Process database. Top 10 results in each category are plotted. LV-inoculated samples: concatenated CD8-LV and VSV-LV samples.

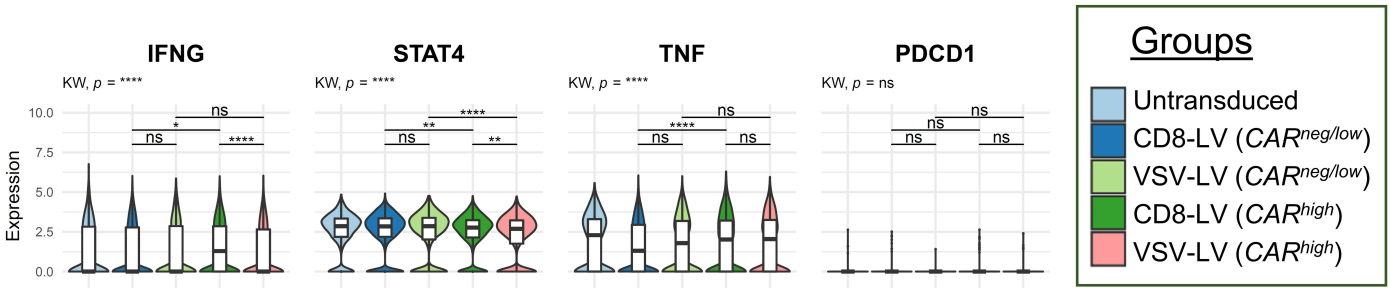

Figure S8. Supplemental violin plots.
Unaltered expression of particular genes of interest (KW: Kruskal-Wallis multiple comparison test. Wilcoxon Rank Sum test performed for paired-wise comparisons, p-values adjusted based on Bonferroni. ns: non-significant, * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001).

S1 Table. Metrics of cells pre- and post-filtering in scRNA-seq analysis.

	Untransduced	CD8-LV	VSV-LV
Unfiltered cells	2711	9793	9023
Post-filtered cells	2654	9713	8975
Total removed cells	57	80	48
Proportion of removed cells	2.10 %	0.82 %	0.53 %