Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Single-cell genomics links CAR architectures to in vivo effector chromatin states in T cells responding to antigen-low leukemia

Kole DeGolier
National Jewish Health
Stephanie DeVoe
National Jewish Health https://orcid.org/0000-0003-2626-8366
Marwan Abushawish
National Jewish Health
Angela Prete
National Jewish Health
Terry Fry
University of Colorado Anschutz Medical Campus
James Scott-Browne
scottbrownej@njhealth.org

National Jewish Health

Article

Keywords: CAR, T CELLS, CD22-CARs,

Posted Date: June 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9859689/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Yes there is potential Competing Interest. T.J.F. is the holder of a patent related to the affinity matured CD22-specific monoclonal antibody and uses thereof (US11939377B2) filed by the National Institutes of Health. T.J.F. consults for Sana Biotechnology.

Single-cell genomics links CAR architectures to *in vivo* effector chromatin states in T cells responding to antigen-low leukemia

Kole R. DeGolier[1], Stephanie DeVoe[1,2], Marwan. A. Abushawish[1,2], Angela Prete[1], Terry J. Fry[3,4], James P. Scott-Browne[1,4]*

[1]Department of Immunology and Genomic Medicine, National Jewish Health, Denver, CO, USA, 80206
[2]Immunology Graduate Program, University of Colorado Anschutz Medical Campus, Aurora, CO, USA, 80045
[3]Department of Pediatrics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA, 80045
[4]Department of Immunology & Microbiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA, 80045

*Correspondence should be addressed to J.P.S-B. (scottbrownej@njhealth.org).

James P. Scott-Browne, Ph.D.
National Jewish Health
1400 Jackson St.
Denver, CO, 80220
Phone: 303.398.1228
Email: scottbrownej@njhealth.org

Nature Communications

Abstract Word Count: 199 (200)
Main Text (Intro, Results, Discussion) Word Count: 3460 (5000)
Display Items (Figures and Tables) Count: 4 Figures, No Tables (10)
Extended Data: 5 Figures, 2 Tables (10)
References: 52 (70)

ABSTRACT

Chimeric antigen receptor (CAR) T cells produce extraordinary remission rates in some hematologic malignancies, but relapses occur, often through antigen-low tumor escape. Clinical efficacy against antigen-low cancers is enhanced by modifying CAR architecture, with a drastic effect observed in CD22-CAR T cells. However, efforts to define biology driving CAR T cell responses to antigen-low tumors have largely relied on functional readouts or transcriptome, limiting resolution of transcriptional programs and long-term T cell fates. We previously generated a panel of CD22-CARs in which differential binding affinity and tonic signaling produce graded therapeutic efficacies against antigen-low leukemia. Here, we develop a genetic barcoding strategy that captures CAR abundance and single-cell chromatin states using scATAC-seq. In an orthogonal murine CAR T cell model, this approach resolved IL-2-dependent chromatin signatures. We then directly benchmarked multiple CD22-CAR architectures against a robust CD19-CAR anti-leukemia response within the same host. We report CD22-CAR binding affinity and tonic signaling regulate *in vivo* T cell expansion and effector differentiation, uncovering programs that improve antigen-low leukemia control. These data establish genetic barcoding to tie functional manipulations to single CAR T cell chromatin profiles, providing a scalable platform with potential to inform generalizable rules to engineer antigen-sensitive CAR T cells.

Cancer immunotherapy via adoptive transfer of chimeric antigen receptor (CAR) T cells can mediate remarkable patient outcomes[1-4]. While CARs targeting CD19 (CD19-CAR) have been highly successful in treating B-lineage hematologic tumors, complete loss of CD19 characterizes many relapsed tumors, highlighting the need for robust CAR targeting of other antigens[5]. CD22-targeted CARs (CD22-CAR) mediate high remission rates in B-lineage acute lymphoblastic leukemias (B-ALL) and lymphomas that are resistant to CD19-CAR therapy, but patients can relapse with low tumoral CD22 density ($CD22^{Lo}$)[2,6]. CD22-CARs containing a short linker between the heavy and light chain in the antigen binding domain drive antigen-independent 'tonic' signaling and have the greatest clinical efficacy, partly due to enhanced antigen sensitivity[2, 7]. However, this trait has been linked to T cell dysfunction and a severe toxicity known as Immune Effector Cell-Associated Hemophagocytic Lymphohistiocytosis-Like Syndrome (IEC-HS)[7-10], with a putative mechanism of tonic CAR signaling and low-antigen leukemia resulting in sustained T cell activation with reduced effector function. While CD28 can drive enhanced efficacy in other CAR formats by improving effector function and antigen sensitivity, 4-1BB-costimulated CD22-CARs remain the most efficacious, likely due to increased T cell persistence[11, 12]. Although the clinical experience demonstrates the functional impact of a tonically signaling CD22-CAR[7], the influence of specific CAR designs on long term T cell programming remains to be elucidated.

Indeed, modifying CAR structure has improved therapeutic efficacy across multiple CARs in preclinical studies and clinical trials[2, 7, 11, 13-15], but tying pre-transfer manipulations to long-term T cell fates which contribute to patient outcomes has been challenging. We and others have shown that chromatin accessibility reflects activity of transcription factors (TFs) controlling gene expression programs driving T cell development, activation, and function[16-19]. Single-cell chromatin accessibility assays identify heterogeneity in cell fates via projection of genome-wide TF activity and assessment of linked regulatory elements to determine how transcriptional programs may be controlled. We recently demonstrated that modifications to CD22-CAR binding affinity and tonic signaling produce a set of CAR architectures with graded therapeutic efficacies against antigen-low leukemia[13]. Here, we sought to understand the biology of CD22-CARs with distinct binding affinities and tonic signaling levels responding to antigen-low leukemia by directly comparing to the CD19-CAR robustly responding to high antigen levels on the same tumor within the same host. We hypothesized that an *in vivo* chromatin accessibility readout would uncover T cell regulatory programs conferred by CAR architecture which drive CD22-CAR T cell responses to antigen-low leukemia, with barcode abundance simultaneously revealing competitive fitness advantages.

Spear-ATAC (single-cell perturbations with an accessibility read-out using assay for transposase accessible chromatin with sequencing) was previously used to determine the effects of CRISPR/Cas9 perturbations on chromatin profiles of pooled cell lines *in vitro* at single cell resolution[20]. We build on this method, establishing a strategy which uses genetic barcodes to track the *in vivo* chromatin signatures of human T cells expressing CD22-CARs with different antigen binding domains. After validation of our barcoding strategy using pooled cell lines from different species and then primary pooled murine CAR T cells, we use this approach to profile human CD22-CAR variants *in vitro* and following transfer into mice bearing antigen-low leukemia to mimic clinical relapse patterns. We report antigen-independent 4-1BB signaling, widely considered essential for clinical activity

of CD22-CARs[2, 7], alters *in vivo* T cell expansion and confers distinct effector programming in both CD4 and CD8 CAR T cells relative to the CD19-CAR, which may both reduce anti-tumor efficacy and contribute to IEC-HS toxicities. Overall, we establish an approach for pooled fate tracking with scATAC-seq, using genetic barcodes linked to CD22-CAR architecture, and decode chromatin-level determinants of *in vivo* CAR T cell function against antigen-low leukemia.

RESULTS

Species-associated barcoding guides optimal filtering parameters to balance fidelity and cell capture efficiency

To generate a Spear-ATAC-compatible vector, we modified our human-CD19-reactive murine CAR construct[21] to embed a DNA barcode flanked by Nextera sequencing adapters downstream of the mouse Thy1.1 (mThy1.1) reporter (Figure 1A). To validate barcode detection and assignment, we transduced murine EL4 and human Jurkat T cell lines with three distinct genetic barcodes each before sorting for mThy1.1+ cells using fluorescence-activated cell sorting (FACS) and pooling for scATAC-seq (Figure 1B, S1A). Barcode transduction did not alter expected scATAC-seq fragment sizes or standard QC metrics (Figure S1B-C). Barcodes were amplified by spike-in of a barcode-amplification oligonucleotide during capture, followed by nested PCR to enrich and index barcode-containing fragments from the final scATAC library (Figure S1D-E). We filtered for cells meeting QC thresholds before visualization by t-Distributed Stochastic Neighbor Embedding (t-SNE). These comparisons revealed distinct populations of human, mouse and doublet cells determined by the numbers of genomic reads in each cell mapping to mouse and human genomes (Figure 1C). We assigned barcode identity to each cell based on 'barcode specificity' (proportion of barcode reads matching barcode with most reads) and number of barcode reads for the barcode with most reads. We compared classification of cell species based on assigned barcode versus the determination based on numbers of reads mapping to each genome (Figure 1D). We found that increasing the specificity threshold reduced proportions of incorrect species assignments, while increasing the reads threshold reduced multiplets (Figure S1F). A sensitivity threshold of 0.95 with at least 100 reads maximized cell inclusion and minimized multiplets or incorrect species assignments across several barcode constructs (Figure 1D-E, S1F-G). Importantly, with equal pooling of distinctly barcoded cells we recovered similar proportions of each barcode, indicating barcode sequence did not affect recovery rate (Figure 1F-G). These analyses validated the approach and empirically determined *in silico* thresholds to ensure data quality and method fidelity for downstream experiments.

Orthogonal *in vitro* validation using primary murine CAR T cells with culture condition-associated barcodes reveals stable and transient transcriptional programs dictated by IL-2 concentration

The T cell growth factor IL-2 plays a central role in protocols for expanding tumor-specific T cells for adoptive cell therapy, with varying concentrations used during T cell culture[22, 23]. However, IL-2 concentration during culture can vary, with few specific comparisons to characterize long-term effects on T cell differentiation[24]. To assess whether CAR barcoding with scATAC-seq could detect meaningful biological differences in primary T

cells, we used Spear-ATAC to compare chromatin accessibility of murine CAR T cells cultured in three IL-2 supplementation conditions: 100IU/mL ('Hi'), 10IU/mL ('Lo'), or a 'Swap' condition which started at 100IU/mL before switching to 10IU/mL (Figure 2A). We transduced T cells with two distinctly barcoded CAR constructs for each culture condition (Hi/Lo/Swap) before sorting and pooling for scATAC-seq (Figure 2A, S2A). We anticipated that Swap could emulate IL-2 deprivation experienced by T cells post-transfer (a driver of reduced persistence) and reveal stable versus transient transcriptional programs[25]. After filtering and QC (Figure S2B-D), we observed distinct accessibility profiles for each condition, replicable across each set of two barcodes (Figure 2B, S2E-F). To infer global TF activity and gene expression, we computed ChromVAR[26] z-scores and gene scores. T cells with high STAT5 motif-associated z-scores (Hi) versus high *Il2* gene scores (Lo and Swap) were discrete, consistent with exclusivity of populations sending or receiving IL-2 priming signals[27] (Figure 2C-D). Unbiased clustering separated cells by culture condition, with overlap between Lo and Swap supporting plasticity of some IL-2-dependent programs (Figure 2E-F). Ranked pairwise comparisons revealed strong condition-dependent differences in TF activity (Figure 2G). STAT5 activity was amplified in Hi and some Lo cells, while low activity in Swap indicated uniform IL-2 dependance. FOSL/JUND activity was bimodal in Lo, with increased activity in Hi dropping after Swap. RUNX activity between Hi and Lo was similar, with increased activity in Swap. TCF/LEF activity dropped in Hi compared to Lo, recovering in Swap. Finally, BATF/IRF only minorly retracted upon Swap (Figure 2H). Globally ranked TF activity revealed FOSL/JUND as having the greatest variability between conditions by a large margin (Figure S2G). Plotting this motif in tandem with others revealed unique patterns of coregulation (BATF/IRF, STAT5) or opposition (TCF/LEF) (Figure 2I-K). Thus, we confirm our multiplexed approach can detect meaningful biological differences *in vitro*, revealing that IL-2 concentration during CAR T cell transduction and expansion has a prominent impact on TF activity, with heterogeneity in stable versus transient programs.

In vitro and *ex vivo* coupling of CD22-CAR architecture-associated barcodes with chromatin accessibility in human T cells exposes temporal patterns associated with adoptive transfer

CD22-CAR T cells have shown promise in the clinic, but patients often relapse with diminished leukemic CD22 expression compared to pre-treatment[2, 8, 9]. This highlights an unmet clinical need for a better understanding of the biology underlying CAR T cell responses to antigen-low tumors to bolster CAR therapeutic efficacy. Prior work demonstrates that transcriptional perturbations including overexpression of RUNX2, T-bet, or c-JUN can increase CAR functionality against antigen-low leukemias[16, 28, 29] but programs necessary for robust responses to antigen-low tumors remain unclear.

We previously modified the NALM6 B-ALL cell line to express low levels of CD22 (NALM6-CD22^{Lo})[2] and verified high endogenous CD19 expression and absence of CD33 (Figure 3A, S3A). We also previously demonstrated that CD22-CAR antigen binding domain architecture determines graded anti-tumor efficacies of CD22-CAR T cells against NALM6-CD22^{Lo} with the greatest activity observed using a high affinity single chain fragment variable (scFv) with a long linker between heavy and light chain, reducing tonic signaling[13]. CD22-CARs tested incorporated standard (SA) or high affinity (HA) binders with short (SL) or long linkers (LL), totaling four CD22-

CAR variants (Figure 3B, C). We confirmed expression of each barcoded CAR variant on T cells (Figure S3B-D).

To understand how CD22-CAR affinity and tonic signaling modulate T cell fate, we employed Spear-ATAC to track pooled T cells expressing different CD22-CARs during response to NALM6-CD22^{Lo}, including internal positive (CD19-CAR) and negative (CD33-CAR) controls (Figure 3C-D, S3E-J). Relative CD22-CAR abundance was replicable across donors and in accordance with previously reported anti-tumor efficacies of the CAR variants[13] (Figure 3E). We combined two donors and two timepoints (n = 25,236 total cells) to perform global chromatin accessibility analyses. We assigned CD4 or CD8 coreceptor subset labels based on cutoffs for *CD4*, *CD8A* and *CD8B* gene scores (Figure 3F). All CARs had a greater proportion of CD4 T cells after transduction ('Pre-Transfer'), while transfer into leukemia-bearing hosts ('Post-Transfer') resulted in accumulation of CD8 cells, regardless of target antigen density[13] (Figure 3G, S4A-B). Uniform Manifold Approximation and Projection (UMAP) analysis revealed firm separation by timepoint, even with the CD33-CAR, indicating that distinct environmental cues between *in vitro* culture and *in vivo* xenograft drive critical antigen-independent programs in adoptively transferred T cells (Figure 3H, S4C).

Interspecies mapping of IL-2-dependent murine genomic regions to human T cells supports decreased IL-2 availability as a key determinant of altered chromatin accessibility upon xenogeneic transfer

The distinctions between Pre-Transfer and Post-Transfer in all CARs (Figure 3G-H, S4A-C) suggests adoptive transfer as a major determinant of chromatin accessibility, independent of antigen exposure. We hypothesized that reduced IL-2 concentration experienced by T cells upon transfer contributed to distinct chromatin profiles. Cells with the highest STAT5 activity (Pre-Transfer, CD8), were distinct from those with the highest *IL2* gene scores (Post-Transfer, CD4) (Figure S4D-E). To further interrogate this, we created reference sets of human peaks with high sequence similarity to regions of the mouse genome that were differentially accessible between Hi and Lo IL-2 conditions. We scored our human T cell data on the reference peaks which were more accessible in Hi than Lo ('Hi > Lo') or vice versa ('Hi < Lo'). Pre-Transfer cells mirrored the STAT5 activity data, with the CD8 compartment showing the highest average z-score for Hi > Lo peaks, indicative of IL-2 signaling. Conversely, Post-Transfer cells mirrored *IL2* gene scores, exhibiting high z-scores for peaks enriched in Hi < Lo. Slightly lower z-scores in CD4 cells were consistent with a role in IL-2 production (Figure S4F-I). The central role of IL-2 was also apparent from decreased accessibility across the *IL2RA* locus after transfer (Figure S4J). Together, these data pinpoint the drop in IL-2 concentration upon adoptive transfer as a major determinant of chromatin accessibility and support a helper role for CD4 CAR T cells, producing IL-2 *in vivo*.

Efficacy-associated chromatin signatures revealed by global profiling of human T cells

To define global distinctions in chromatin signatures between CD22-CAR architectures that could be linked to anti-tumor efficacy, we clustered cells from both donors at both timepoints (Figure 3I). Strong efficacy-delineated CAR enrichment patterns were observed within the Post-Transfer timepoint (Cluster 2-5) (Figure 3J). HALL shared enrichment patterns with the strongly responding CD19-CAR (Cluster 4), and SALL clustered with the

bystander CD33-CAR and SASL (Clusters 2,5). The two SL CD22-CARs coincided across timepoints, signifying linker-driven programming (Cluster 3,7 in Figure 3I-J). We also observed shared enrichment of SL CARs in CD4-associated Cluster 3, and CD19 and HALL in CD8-associated Cluster 4 (Figure 3J). We next identified differentially accessible peaks in each cluster compared to all cells and compared TF motif enrichment within these groups of regions to infer TF activities. Peaks linked to Cluster 4 were enriched for a T-box motif and a zinc-finger (ZNF) motif with a similar sequence, indicative of robust effector differentiation. In contrast, peaks linked to the CD33 and SALL CAR-enriched Cluster 2 were associated with the TCF7L/LEF motif, suggesting a resting or memory-like population. Finally, ZNF, GATA, RUNX and NFκB motifs were enriched in the SL CAR-associated Cluster 3, indicating restricted or altered effector differentiation programs (Figure 3K). Limited CD8 effector programming in SL CAR-enriched Cluster 3 was validated by patterns of increased (*LEF1)* or decreased (*PRF1*) chromatin accessibility relative to Cluster 4 (Figure 3L-M). In summary, high anti-tumor potential is linked to strong CD8 effector programming seen in CD19 and HALL CARs, and lower anti-tumor potential with altered effector programs (SL CARs) or resting/memory-like programs (SALL/CD33-CARs).

CD22-CAR architectures drive distinct *in vivo* effector transcriptional programs in CD4 and CD8 CAR T cells

To further elucidate transcriptional programs driven by differences in CD22-CAR architecture, we separated CD4 and CD8 CAR T cells and plotted differences in mean ChromVAR z-scores between CD22-CARs after grouping by affinity or linker. Activity of several TF families showed affinity (NFκB, T-box) or linker (NFκB, MZF, BATF/IRF) biases. We also observed CAR-specific TF activity preferences with HALL (T-box), HASL (Sox), SALL (IRF/STAT), and SASL (RORC/RORA, RUNX) (Figure 4A-D). Interestingly, we also found the magnitude of linker-driven differences was greater in CD8 than CD4 T cells Pre-Transfer, with the opposite pattern Post-Transfer, indicating tonic 4-1BB signal may robustly program CD4 T cells after transfer, while pre-programming CD8 T cells (Figure 4A-D, S5A-D). We observed several consistent and robust differences in TF activity in CD4 and CD8 T cells, with T-box activity associated with efficacy, NFκB activity showing a SL bias and nuclear receptor (RORC/RORA) activity showing SASL bias in CD4 T cells, highlighting the substantial influence of CD22-CAR architecture on effector and helper-lineage differentiation (Figure 4A-H, S5A-L). To interrogate potential connections between TF activities, we made pairwise comparisons stratified by CAR and coreceptor assignment (Figure 4I-N). The SASL architecture drove a high proportion of CD4 cells with RORC/RORA activity and non-canonical RORC/RORA^{Hi}/TBX/EOMES^{Hi} cells (Figure 4I,K,L). T-box activity showed an efficacy-associated gradient, highest in HALL and CD19 CARs for both CD4 and CD8 subsets (Figure 4E,G,I,J). Finally, we observed that SL CD22-CARs blunted proportions of RORC/RORA^{Lo}/TBX/EOMES^{Hi} (CD4) and NFKB/RELA^{Lo}/TBX/EOMES^{Hi} or TCF7L/LEF^{Lo}/TBX/EOMES^{Hi} (CD8) T cells, populations highly enriched in both efficacious HALL and CD19 CARs (Figure 4L,N, S5H,K,L). Together, these data demonstrate that CD22-CAR architectures drive distinct NFκB, RORC/RORA, TBX/EOMES and TCF/LEF activities, with SL CARs promoting altered helper-lineage and effector transcriptional programs in CD4 and CD8 CAR T cells, respectively.

DISCUSSION

CD22-CAR T cells have shown promise in the clinic, but durable outcomes have been limited by the emergence of antigen-low relapse[2, 8, 30]. Lack of antigen sensitivity has been shown in many CARs targeting antigens expressed on hematologic and solid tumors, highlighting the need for a better understanding of the biology of CAR T cells responding to antigen-low malignancies[11, 31-33]. We set out to characterize CD22-CAR responses to antigen-low leukemia, profiling architecturally distinct CD22-CARs in which modifications to antigen binding affinity and tonic signaling have been shown to drive drastic changes in anti-tumor efficacy in multiple previous reports[2, 13, 34]. Recent strategies employing pooled *in vivo* assays allow for high throughput profiling in a homogenous environment and facilitate competition between T cells with mixed perturbations with a direct readout of abundance, aiming to identify manipulations which augment T cell function. Pooled analyses also reduce batch effects and use of sequencing reagents and research animals. However, these methods often rely solely on abundance-based, transcriptomic, or standard functional readouts and may be limited in predicting long-term cellular fate[14, 35-38]. To this end, we developed a barcoded scATAC-seq strategy building on the previously reported Spear-ATAC method[20] to link CD22-CAR architectures to effects on TF activities and chromatin states in T cells, predicting a chromatin-level readout may reveal biology to enhance long-term anti-tumor efficacy.

To validate our strategy, we transduced cells from different species with distinct barcodes to determine optimal parameters for barcode detection and assignment. To further evaluate barcoding fidelity in primary T cells in a well-controlled system, we characterized the chromatin landscape of murine T cells programmed by fixed IL-2 concentrations or a stepwise concentration swap to mimic adoptive transfer. We observed IL-2 concentration dictated inferred IL-2 production versus signaling, with ranging stability of TF activities (notably bZIP, RUNX, TCF/LEF). Overall, these data suggest strategies to tune T cell transcriptional fate to synergize with IL-2 driven programs and validate our approach while supporting broader utility of barcoded scATAC-seq for linking defined manipulations to T cell chromatin states across experimental contexts.

We then applied our platform to understand the biology driving anti-tumor responses of T cells expressing different CD22-CAR variants. The adoptive transfer process emerged as a key determinant of chromatin accessibility, even without antigen engagement. Mapping accessibility changes in our murine IL-2 data to conserved human genomic elements revealed scarcity of IL-2 as a major factor. We also note qualitative and quantitative differences in tonic signal-driven transcriptional signatures in CD4 and CD8 CAR T cells, with greater magnitude Pre-Transfer and Post-Transfer, respectively. This suggests that engineering CD4 and CD8 populations with distinct CAR constructs or perturbations could maximize combined anti-tumor efficacy. Our data further identify early CD8 T cell expansion and CD4/CD8 effector differentiation in T cells expressing CARs which more effectively control $CD22^{Lo}$ leukemia, with the HALL CD22-CAR closely emulating a CD19-CAR response to high antigen levels.

Prior studies described positive effects of tonic 4-1BB signaling toward rescuing exhaustion[10] or inducing the TF *BACH2*, which could be overexpressed in CD28-costimulated CAR T cells to restrain terminal differentiation[34]. Conversely, we report that tonic 4-1BB signaling in short linker CD22-CARs alters effector differentiation in both

CD4 and CD8 CAR T cells and observe that long linker CARs with lower tonic signal confer more physiologic effector programs. Specifically, we find that short linker CARs skew CD8 T cells away from NFKB/RELA^{Lo}/TBX/EOMES^{Hi} and TCF7L/LEF^{Lo}/TBX/EOMES^{Hi} effector populations, which are present in the efficacious CD19 and HALL CARs. Additionally, we observe the short linker CARs modify CD4 helper-lineage differentiation, exhibiting diminished proportions of RORC/RORA^{Lo}/TBX/EOMES^{Hi} T cells. Strong effector responses are likely integral to robust targeting of CD22^{Lo} leukemia, a key clinical relapse modality and a requirement for CD22-CARs. We recently demonstrated overexpression of T-bet in human CD19-CAR T cells with 4-1BB signaling boosted early responses to antigen-low leukemia[28]. Refining T-bet mediated enhancements or modulating NFκB to synergize with specific CD22-CAR transcriptional programs could prove fruitful. Future studies should also examine contributions of tonic signaling from other CARs to T cell differentiation states, and CD4/CD8-specific perturbations.

Our data also provide insights into potential mechanisms of IEC-HS, a severe clinical toxicity associated with the SASL CD22-CAR. As classical Hemophagocytic Lymphohistiocytosis (HLH) is associated with mutations in effector genes like *PRF1*[39, 40], we propose that altered effector differentiation driven by short linker CD22-CARs may contribute to IEC-HS and could be remedied by the non-tonic signaling HALL CD22-CAR. Tracking T cell fates will also be crucial to optimizing therapeutic efficacy and reducing toxicities for promising allogeneic donor-derived and *in vivo* transduction-based therapies with less defined safety profiles.

In summary, we establish a barcoded scATAC-seq approach to track *in vivo* transcriptional programs in CD22-CAR T cell pools during response to antigen-low leukemia, and report CAR antigen binding domain architecture as a central determinant of T cell effector chromatin signatures. While tonic 4-1BB signaling driven by a short linker has been considered necessary for CD22-CAR efficacy, we show that short linker CARs alter effector differentiation, limiting TBX/EOMES activity while disproportionately driving RORC/RORA, NFKB/RELA and TCF/LEF activity during the effector response. In contrast, T cells expressing an affinity-matured long linker CAR (HALL) show robust effector differentiation, most comparable to highly efficacious CD19-CAR T cells responding to high leukemic CD19 levels. We expect these data and methodologies will inform strategies to tune long-term functional fitness of engineered immune cells, improving outcomes across CARs and disease settings.

MATERIALS AND METHODS

Culture of cell lines

EL4 and Jurkat E6-1 cells were obtained from the American Type Culture Collection (ATCC). NALM6-CD22^{Lo} were generated in the Fry laboratory and were originally obtained from the Leibniz Institute DSMZ German Collection of Microorganisms and Cell Cultures GmbH (https://www.dsmz.de/collection/catalogue/details/culture/ACC-128). All lymphocyte cell lines were cultured in Complete RPMI (cRPMI), consisting of RPMI-1640 (Gibco, 11875-093) with 10% (vol/vol) Fetal Bovine Serum (FBS) (Innovative Bioscience, Cat No. 11-01-500), 55μM ß-mercaptoethanol (Sigma-Aldrich, Cat No. M3148-25ML), 2mM L-Glutamine (Gibco, 25030-081) and 500μg/ml (1% vol/vol) of Penicillin/Streptomycin (Gibco, 15140-122). For viral production, the Lenti-X cell line was obtained from Takara Bio (Cat No. 632180) and the Platinum-E cell line was obtained from Cell BioLabs (Cat. No. RV-101). Both viral production cell lines were cultured in Complete DMEM (cDMEM), consisting of DMEM (Gibco, 11965-092) with 10% (vol/vol) FBS, 55μM ß-mercaptoethanol, 2mM L-Glutamine and 500μg/ml (1% vol/vol) of Penicillin/Streptomycin. Viral production cell lines were passaged at least once prior to transfection with viral plasmids. The EL4 cell line stained positive for mouse PD1 and mouse Thy1.2 and negative for human Thy1 and mouse Thy1.1. The Jurkat E6-1 cell line stained positive for human Thy1 and negative for mouse Thy1.2 and Thy1.1. Lenti-X and Platinum-E cell lines were validated morphologically. All cell lines tested negative for *Mycoplasma* contamination using a Universal Mycoplasma Detection Kit (ATCC, Cat No. 30-1012K).

Preparation of human donor T cells

Healthy human donor blood was obtained from National Jewish Health Blood Donation Core. Blood was processed using Ficoll-Paque Premium (GE Healthcare, Cat No. 17-5442-02) followed by red blood cell lysis in ACK Lysing Buffer (Gibco, Cat No A10492-01) to isolate PBMCs. CD3+ cells were isolated using the MojoSort Human CD3 Selection Kit (BioLegend, Cat No. 480134) and cryopreserved in Freeze Media consisting of 90% FBS and 10% DMSO (Fisher Bioreagents, Cat No. BP231-100) for later use.

Generation of CAR constructs

Basic murine CAR construction was described previously[21], with a hCD8A signal peptide, Myc-Tag, anti-human CD19 scFv with $(G_4S)_{x3}$ linker, mCD28 hinge/transmembrane/costimulatory domain, and mCD3ζ in the MSCV-puro (Clontech) murine retroviral vector, with a porcine teschovirus-1 2A peptide (P2A) followed by mThy1.1 reporter. This original construct was modified to insert a barcode cassette consisting of a 16bp DNA barcode flanked by Nextera sequencing adapters downstream of the mThy1.1 reporter. Basic human CAR construction was described previously, with all CARs consisting of a CD8A signal peptide, respective scFv, CD8A hinge/transmembrane domain, and 4-1BB and CD3ζ signaling domains in the pLenti lentiviral vector[13]. scFvs consisted of: anti-hCD33 (clone huM195, $(G_4S)_{x3}$ linker), anti-hCD22 (clone m971 (SA) or m971-L7 (HA) with either $(G_4S)_{x1}$ (SL) or $(G_4S)_{x3}$ (LL) linker), or anti-hCD19 (clone FMC63, $(G_4S)_{x3}$ linker). All human CARs contained

the P2A, mThy1.1 reporter and a similar barcode cassette as described for the murine CAR. Both DNA constructs contained a WPRE3 sequence[41] to enhance transgene expression and provide a filler DNA sequence for the nested PCR barcode amplification.

Barcode generation

The original barcode cassette was generated by assembly of an oligonucleotide pool containing random 16nt sequence (specified as 'NNNNNNNNNNNNNNNN') into the plasmid containing murine CAR. Individual colonies were picked and barcode sequence was identified from whole plasmid sequencing (Quintara Biosciences). We selected 6 barcodes with low similarity to other barcodes and low numbers of nucleotide repeats. For the human CAR constructs, a list of indel-correcting 16bp DNA barcodes from the *FreeBarcodes* resource[42] was filtered using the *create.dnabarcodes()* function[43] for those with a minimum hamming distance of 13.

Production of gammaretrovirus and lentivirus

For gammaretrovirus production, $3.5x10^5$ Platinum-E cells were plated in a 6 well plate in cDMEM (see above for recipe) the evening before transfections. The following day, media was replaced with 2.5mL fresh cDMEM, followed by transfection with 1.8µg transfer plasmid and either 0.6µg pCL-Eco (for ecotropic virus used on primary mouse T cells) or 0.6µg pCL-10A1 (for dual pseudotyped amphotropic/ecotropic virus used in barnyard experiment) using TransIT-LT1 Transfection Reagent in 250µL OptiMEM (Gibco, 31985-062). For lentivirus production, $1x10^7$ Lenti-X cells were plated in a T-75 flask in cDMEM the evening before transfections. The following day, media was replaced with 16mL fresh OptiMEM+, consisting of OptiMEM (Gibco, 31985-062) with 5% (vol/vol) FBS, 2mM L-Glutamine and 1mM Sodium Pyruvate (Gibco, 11360-070), followed by transfection with 11.25µg transfer plasmid and pMDLg/pRRE (7.5µg), pMD2.G (3.75µg) and pRSV-Rev (7.5µg) using TransIT-LT1 Transfection Reagent (Mirus Bio, Cat No. 2300) in 2mL OptiMEM (Gibco, 31985-062). Virus-containing supernatant was collected 48hrs later and spun for 5min at 2000xg to remove dead cell particulate before use. All packaging plasmids were obtained from Addgene, with the exception of pCL-10A1 (Novus Biologicals, Cat #NBP2-29542).

Barnyard experiment transduction

EL4 murine T cell lymphoma and Jurkat E6-1 human T cell leukemia cell lines were cultured in cRPMI as described above. Cell lines were thawed and split once before one round of transduction at 2000xg for 90 minutes with fresh gammaretrovirus and 10µg/mL polybrene, followed by removal of gammaretrovirus and polybrene after the spin and replacement with fresh cRPMI. Cells were transduced with dual pseudotyped amphotropic/ecotropic gammaretrovirus (adding pCL-10A1 envelope instead of pCL-Eco). Transduction efficiency was measured by FACS and then cells were FACS sorted (Lymphocytes>Singlets>Live>mThy1.1+) and captured by scATAC-seq.

Generation of murine CAR T cells

Mouse CAR T cells were cultured in cRPMI as described above. CD8+ T cells were isolated from 10-16 week old C57BL/6 mouse splenocytes using the MojoSort Mouse CD8 T Cell Isolation Kit (BioLegend, Cat No. 480035). On day 0, T cells were activated with 1µg/mL anti-CD3 and 1µg/mL anti-CD28 overnight. Cells were transduced on day 1 by centrifugation at 2000xg for 90 minutes at 37°C with fresh gammaretrovirus and 10ug/mL polybrene, followed by removal of retrovirus and polybrene after the spin and replacement with fresh media containing the specified amount of recombinant human IL-2, either 10IU/mL ('Hi') or 100IU/mL ('Lo') (National Cancer Institute). Transduction was repeated on the following day (day 2), and then cells were removed from anti-CD3/anti-CD28 stimulation and plated at $1x10^6$ cells/mL in fresh media containing IL-2. On day 3 and 4, cells were replated at $1x10^6$ cells/mL in fresh media containing IL-2. For the 'Swap' condition, IL-2 concentration was changed from 100IU/mL to 10IU/mL. Cells were FACS sorted (Lymphocytes > Singlets > Live > CD8+ > mThy1.1+) and captured for scATAC-seq on day 5, 16 hours after the IL-2 swap (or media replacement for the Hi and Lo conditions).

Generation of human CAR T cells

CD3+ human T cells were cultured in Human T Cell Expansion Media (hTCEM), consisting of AIM-V (Gibco, Cat No 12055-091) supplemented with 5% FBS and 10mM HEPES (Gibco, 15630-080). On day 0, T cells were thawed and activated with Dynabeads Human T cell Expander CD3/CD28 (Gibco, Cat No. 11141D) at a 3:1 bead:cell ratio at $1x10^6$ cells/mL, in hTCEM supplemented with 40IU/mL IL-2 and cultured for 48 hours. On day 2, T cells were counted and transduced at $2x10^6$ cells per well in a 6 well plate, with 10µg/mL Protamine Sulfate (Millipore-Sigma, Cat No. P4020-1G), spinning at 1000xg for 2.5 hours at 32°C. On day 3, T cells were transduced again with fresh virus. On day 4, beads were removed, and cells were plated in fresh hTCEM supplemented with 100IU/mL recombinant human IL-2 at $0.5x10^6$ cells/mL. On day 6, cells were replated at $0.5x10^6$ cells/mL with fresh hTCEM and 100IU/mL IL-2, and a small aliquot was taken to check transduction efficiency. On day 8, cells were counted with a small aliquot taken for FACS sorting (Lymphocytes > Singlets > Live > CD4+ or CD8+ > mThy1.1+) and scATAC-seq capture (Pre-Transfer timepoint), with the rest cryopreserved in Freeze Media as described above.

Nalm6 xenograft *in vivo* studies

NOD-Scid-Gamma ('NSG', NOD.Cg-*Prkdc*^{scid}*Il2rgtm1Wjl*/SzJ; Strain #: 005557, Jackson Laboratories) were obtained commercially and then bred in-house. For xenograft studies, between 4 and 9 male or female NSG mice (up to 20 weeks of age, same gender in a single experiment) received $1x10^6$ NALM6-CD22^{Lo} leukemia on day -3, followed by $1x10^6$ total CAR+ cells ($1.67x10^5$ of each CAR, pooled equally) on day 0. Leukemia burden was monitored before CAR injection (day -1), on day 4, and on day 7 before marrow harvest, using the Xenogen In Vivo Imaging System (IVIS) Spectrum (Caliper Life Sciences) after intraperitoneal injection with D-Luciferin Potassium Salt (Revvity, Cat No. 122799A-5). Total flux (photons/s) was measured for each mouse using Living Image software (Revvity, v4.8.2). On day 7, bone marrow was harvested, filtered through a 70µm strainer, and red blood cells were lysed using ACK Lysing Buffer. Cells were then filtered again, and human cells were

enriched using MojoSort Mouse CD45 Nanobead (BioLegend, Cat No. 480028) depletion of mouse bone marrow. Cells were stained and FACS sorted (Lymphocytes > Singlets > Live > CD45+/GFP- > CD4+ or CD8+ > mThy1.1+) for scATAC-seq capture (Post-Transfer timepoint). Xenograft animal studies were conducted at University of Colorado Anschutz Medical Campus in accordance with an animal protocol approved by the Institutional Animal Care and Use Committee (IACUC, Protocol 751).

Flow cytometry and FACS sorting

Flow cytometric analysis was performed on an A5 Flow Cytometer (BD Biosciences). For the barnyard experiment in cell lines and the IL-2 experiment in primary mouse T cells, transduction was detected by Myc-Tag+(CAR+)/mThy1.1+ double positive cells. For the human CAR T cell experiments, transduction was detected by Protein-L+/mThy1.1+ double positive cells (to compare geometric mean fluorescence intensity (GMFI) across all CAR constructs) and by Protein-Fc+/mThy1.1+ double positive cells (using either CD19, CD22, or CD33 Protein-Fc as a direct readout of antigen binding for each CAR). Both readouts showed similar detection; Protein-Fc/mThy1.1+ cells was used as the final readout of CAR+ cells for downstream experiments. Cells were stained and sorted in FACS Buffer consisting of PBS (Fisher Bioreagents, Cat No. BP2438-4) with 2mM EDTA (Fisher Bioreagents, Cat No. BP2482-100) and 2% FBS. Sorted cells were collected into 1.5mL tubes containing Sort Buffer consisting of PBS with 0.04% UltraPure BSA (Invitrogen, Cat. No. AM2618). Dead cells were excluded using eBioscience Fixable Viability Dye (Invitrogen, Cat. No. 65-0865-18). NALM6 leukemia was identified by GFP. Flow cytometry data was analyzed and plotted using FlowJo (Becton Dickinson & Company, v10.10.0). Some graphs quantifying flow cytometry data were generated using Prism (GraphPad, v10.4.1).

scATAC-seq capture and library preparation

All steps prior to 10X scATAC-seq capture were performed on ice or in a microcentrifuge cooled to 4°C. For in vivo assays, cells were pooled before transfer and sorted into one tube upon harvest. For *in vitro* assays, cells from each condition were sorted, analyzed for live cell counts via 0.4% Trypan Blue staining on the Countess 3 (ThermoFisher), and pooled at equal numbers. After pooling, cells were washed once in cold PBS and then lysed in Lysis Buffer (Nuclease-Free H_2O with 10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM MgCl2, 0.1% NP40, 0.1% Tween20, and 0.01% Digitonin) for the following times: 5 mins (cell lines), 3 mins (*in vitro* primary cells), 3 mins 45 seconds (Post-Transfer primary cells). Cells were then resuspended in 1X Nuclei Buffer and analyzed for nuclei count and lysis efficiency via 0.4% Trypan Blue staining on the Countess 3 (ThermoFisher), with a typical lysis efficiency of >95%. Cells were then diluted to ≤$2.4x10^6$ cells/mL for capture of ~8000 nuclei in a single 10X capture well, following manufacturer protocols for Chromium Next GEM Single Cell ATAC Reagent Kits v2 (10X Genomics, User Guide CG000496). Two modifications were required to this protocol for barcode enrichment, as described previously[20]: first, 1.32µL of 50µM spike-in barcode amplification oligonucleotide (Oligo 1032, see Supplementary Table 1) was added to the GEM reaction master mix (step 2.1a in 10X protocol); second, the PCR following GEM generation was increased from 12 to 15 cycles (step 2.5a in 10X protocol). Libraries were

prepared per manufacturer specifications and analyzed using the TapeStation (Agilent) and Qubit Fluorometer (ThermoFisher).

Barcode enrichment and library preparation

For specific sequences of all primers used, see Supplementary Table 1. Barcodes were enriched from the final scATAC library with the following 'Barcode Enrichment' PCR:

PCR Mix		*Thermocycling*	
100uL	NEB Q5 2x MasterMix	98°C 1min	
2uL	Oligo 908	98°C 10sec	}
2uL	Oligo 1033	67°C 15sec	} X cycles (see *Note* below)
1.5uL	Diluted scATAC Library (5ng total)	72°C 1min	}
94.5uL	H_2O	72°C 1min	
		10°C ∞	

Note: The following cycle numbers were optimized and used for the barcode enrichment PCR: 16 cycles (retroviral transductions), 25 cycles (lentiviral transductions). We predict that our lentiviral protocol, optimized for lower vector insertion copy numbers per cell, required a greater number of cycles for amplification of the barcode fragments from tagmented chromatin.

The Barcode Enrichment PCR product was cleaned up using the DNA Clean and Concentrator-5 Kit (Zymo Research, D4004). Purified DNA was then run on 2% E-Gel SizeSelect II Agarose Gels (Invitrogen, Cat No. G661012) to select the enriched barcode DNA. Unique i7 indices were added to each individual barcode library using custom indexing primers by running the following 'Barcode Indexing' PCR:

PCR Mix		*Thermocycling*	
100uL	NEB Q5 2x MasterMix	98°C 1min	
2uL	Oligo 908	98°C 10sec	}
2uL	Oligo 92X*	67°C 15sec	} 13 cycles
1.5uL	Diluted E-Gel Enrichment PCR (12 pg)	72°C 1min	}
94.5uL	H_2O	72°C 1min	
		10°C ∞	

*Oligo 92X indicates any oligonucleotide from series 925-928, which are identical except for the 8bp P7 index which allows pooling of indexed samples.

The Barcode Indexing PCR product was taken directly into a 1.1x left side size selection with SPRIselect Beads (Beckman Coulter, Cat No. B23318) using manufacturer protocols for isolation of final barcode library. DNA quality and size distribution was monitored by Tapestation after each step, and by Tapestation and Qubit analysis after the final SPRI selection.

Sequencing

Final scATAC and barcode libraries were pooled to 10nM concentration for sequencing. Final concentrations for individual libraries in the pool were calculated to achieve a target sequencing depth of 40,000 reads/nuclei (scATAC library) and 1,000 reads/nuclei (barcode library), and pooled libraries were sequenced on an Illumina NextSeq 2000 at the National Jewish Health Genomics Facility (Research Resource Identifier – RRID: SCR_023051).

Alignment

Raw sequencing data was processed into fastq files using *cellranger-atac* (10X Genomics, v2.1.0), using the *mkfastq* function, and reads were aligned to the Genome Reference Consortium human hg38 and/or murine mm10 reference genomes using the *count* function.

Barnyard analysis

In addition to the barcode-transduction experiment described above, scATAC was performed on a mixed pool of untransduced EL4 and Jurkat cells to verify barcode distribution did not change fragment size distribution (Figure S1A). Barnyard analysis is n=1 experiment each for the untransduced and barcode-transduced pools. After alignment, barcodes were mapped in R using custom functions and functions adapted from the Spear-ATAC manuscript[20]. Reads were mapped to a combination mm10/hg38 reference genome. Quality control values, number of reads mapping to mouse and human genomes, and species assignments were taken directly from cellranger outputs.

Spear-ATAC Analysis: General

The ArchR R package (https://www.archrproject.com, v1.0.3)[44] was used for all other downstream processing and analysis of scATAC-seq data in primary murine and human T cells. General analysis parameters were followed for both murine and human datasets, with several modifications for the human data as noted, which were either generally more appropriate or accounted for the increased experimental complexity associated with additional replicates and timepoints. We preprocessed the data, reading in barcodes using *getValidBarcodes()* followed by creation of Arrow files using *createArrowFiles()* to read accessible genomic fragments, filter to include TSS enrichment >10 and >10,000 fragments for murine cells (TSS enrichment >8 and >10,000 fragments for human cells), creating a genome-wide TileMatrix using 500bp bins, and creating a GeneScoreMatrix using the appropriate reference genome. An ArchR project was then created using and *ArchRProject()*. We then filtered to include cells with assigned barcodes, reading in all fragments with valid cell barcodes and viral barcodes before assigning each cell a viral barcode using cutoffs of 0.95 for 'barcode specificity' (proportion of barcode reads matching barcode with most reads) and 100 reads for $\log_{10}$(# of barcode reads) for barcode with most reads, with the final barcode assignment being the barcode with the most reads. Dimensionality reduction was performed using ArchR's native Iterative Latent Semantic Indexing implementation[45-47] via *addIterativeLSI()*, with *LSImethod* = 3 (logtf-logidf), *dimsToUse* = 1:13, *useMatrix* = TileMatrix, with *varFeatures* = 25,000 for the murine

data (increased to 50,000 for human data). Harmony batch correction was run on human data to correct for pooling biological donors[48], using the *addHarmony()* function with *groupBy* = Replicate. Clusters were called using Seurat[49], and UMAP added using *addUMAP()* with the IterativeLSI-reduced dimensions (Harmonized IterativeLSI for human data). Finally, peaks were called on Seurat clusters, using *addGroupCoverages()* (minReplicates increased to 3 for human), and *addReproduciblePeakSet()* with *method* = q, *cutoff* = 0.005 (0.001 for human), *peaksPerCell* = 500, *maxPeaks* = 75,000 (65,000 for human), extendSummits = 250, with MACS2[50] before creating a PeakMatrix using *addPeakMatrix()*. Motif annotations were added using *addMotifAnnotations()* with the vierstra archetype motif set[51], and motif deviations archetypeMatrix was added using *addDeviationsMatrix()*.

Spear-ATAC Analysis: Murine CAR T cell specific analyses

Multidimensional Scaling (MDS) dimensionality reduction was used to analyze global similarity in chromatin accessibility profile between pseudobulked cells from each barcode replicate and by condition. Murine IL-2 culture condition analysis is n = 1 experiment with n = 2 independent barcoded cultures per culture condition.

Spear-ATAC Analysis: Human CAR T cell specific analyses

Coreceptor assignment cutoffs were designated by visualizing *CD4*, *CDA*, *CD8B* gene scores, and timepoint/coreceptor subsets were created using *subsetArchRProject()*. To create the motif enrichment heatmap in Figure 3, we defined differentially accessible peaks for each cluster using *getMarkerFeatures()* with *testMethod* = "wilcoxon" and *getMarkers()* with false discovery rate ≤ 0.01 and the absolute value of $\log_2$(fold-change) ≥ 1 before plotting with *plotMarkerHeatmap()*. Motifs enriched within differential accessible peaks were then defined using *peakAnnoEnrichment()* with false discovery rate ≤ 0.01 and the absolute value of $\log_2$(fold-change) ≥ 0.5, before plotting using plotEnrichHeatmap() with n = 6 maximum motifs per cluster and *cutoff* = 20 for minimum adjusted p-value to be included in the heatmap. Comparisons of accessibility for genetic loci tracks were made using either the previously defined differentially accessible peaks for each cluster or by defining differentially accessible peaks between all Pre-Transfer and Post-Transfer cells, both with false discovery rate ≤ 0.1 and $\log_2$(fold-change) ≥ 0.5. Coordinates for the TF activity summary plots to visualize motifs associated with either CD22-CAR affinity or linker were calculated as the difference between ChromVAR[26] motif-associated chromatin accessibility z-scores for the comparisons as indicated on each axis. Points closer to the y = x diagonal line are more directly biased by affinity or linker as indicated, while points away from the diagonal exhibit preferential bias by either the x-axis or y-axis comparison. Cutoff value of >0.5 in any direction was applied for Post-Transfer data (Figure 4) and >0.3 for pre-transfer data (Figure S4). Human CAR T cell analysis is n = 2 experiments with independent biological human donors.

Interspecies genomic mapping

The University of California Santa Cruz Genome Browser 'Liftover' tool[52] was used to compare the relative chromatin accessibility of human T cell populations at homologous regions between mm10 and hg38 reference

genomes. The regions used were the sets of peaks that were differentially accessible between mouse T cells cultured in 'Hi' and 'Lo' IL-2, mapping between mm10 and hg38 reference genomes. Default settings were used, with the exception of the "minimum ratio of bases that must remap" setting, which was changed to 0.5 to increase stringency. Liftover for peaks down in Hi compared to Lo resulted in 3998 converted peaks and 2894 failing to meet cutoffs, while up in Hi compared to Lo resulted in 4871 converted peaks and 3606 failing to meet cutoffs. Liftover comparisons were exported as bed files and peak annotations were added to the ArchR project using *addDeviationsMatrix()* before UMAP visualization.

DATA AVAILABILITY STATEMENT

All data is available from the corresponding author upon reasonable request. Genomics data is uploaded to the Gene Expression Omnibus (GEO) under accession number GSE331391.

ACKNOWLEDGEMENTS

This work was funded by National Institute of Allergy and Infectious Diseases R01AI151021-04 (awarded to J.P.S-B.) and by the Cancer League of Colorado 2025 Research Grant (awarded jointly to K.R.D., T.J.F. and J.P.S-B.). We acknowledge all members of the Scott-Browne Lab for useful input on the manuscript. We acknowledge Trey Farmer and Ralen Johnson at the National Jewish Health Genomics Facility (Research Resource Identifier – RRID: SCR_023051) for assistance with sequencing and Josh Loomis at the National Jewish Health Cytometry Core for assistance with flow cytometry and sorting. We thank Jack Congel at the National Jewish Health Blood Donor Core, and anonymous human blood donors for samples. We additionally acknowledge the Biological Resource Center at National Jewish Health and the Office of Laboratory Animal Resources at University of Colorado for care, husbandry and assistance related to laboratory animals.

AUTHOR CONTRIBUTIONS

Conceptualization: K.R.D., T.J.F., J.P.S-B., Data curation: K.R.D, J.P.S-B., M.A.A., Formal analysis: K.R.D., S.D., J.P.S-B., Funding acquisition: K.R.D., T.J.F., J.P.S-B., Investigation: K.R.D, Methodology: K.R.D., M.A.A., T.J.F., J.P.S-B., Project administration: J.P.S-B., Resources: K.R.D, A.P., T.J.F, J.P.S-B., Software: K.R.D., J.P.S-B., Supervision: K.R.D., T.J.F., J.P.S-B., Validation: K.R.D., J.P.S-B., Visualization: K.R.D., S.D., J.P.S-B., Writing – original draft: K.R.D., J.P.S-B., Writing – review & editing: All authors.

DECLARATION OF INTERESTS STATEMENT

T.J.F. is the holder of a patent related to the affinity matured CD22-specific monoclonal antibody and uses thereof (US11939377B2) filed by the National Institutes of Health. T.J.F. consults for Sana Biotechnology.

KEYWORDS

Leukemia, CAR T cells, tonic signaling, affinity, barcoding, chromatin, ATAC-seq

REFERENCES

1. Brentjens, R.J. et al. CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Sci Transl Med* **5**, 177ra138 (2013).
2. Fry, T.J. et al. CD22-targeted CAR T cells induce remission in B-ALL that is naive or resistant to CD19-targeted CAR immunotherapy. *Nat Med* **24**, 20-28 (2018).
3. Grupp, S.A. et al. Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *N Engl J Med* **368**, 1509-1518 (2013).
4. Turtle, C.J. et al. CD19 CAR-T cells of defined CD4+:CD8+ composition in adult B cell ALL patients. *J Clin Invest* **126**, 2123-2138 (2016).
5. Majzner, R.G. & Mackall, C.L. Clinical lessons learned from the first leg of the CAR T cell journey. *Nat Med* **25**, 1341-1355 (2019).
6. Frank, M.J. et al. CD22-directed CAR T-cell therapy for large B-cell lymphomas progressing after CD19-directed CAR T-cell therapy: a dose-finding phase 1 study. *The Lancet* **404**, 353-363 (2024).
7. Singh, N. et al. Antigen-independent activation enhances the efficacy of 4-1BB-costimulated CD22 CAR T cells. *Nat Med* **27**, 842-850 (2021).
8. Dreyzin, A. et al. Ten-year experience of CD22 CAR T cells for children and young adults with B-cell acute lymphoblastic leukemia. *Blood Adv* **10**, 1700-1712 (2026).
9. Shah, N.N. et al. CD4/CD8 T-Cell Selection Affects Chimeric Antigen Receptor (CAR) T-Cell Potency and Toxicity: Updated Results From a Phase I Anti-CD22 CAR T-Cell Trial. *Journal of Clinical Oncology* **38**, 1938-1950 (2020).
10. Long, A.H. et al. 4-1BB costimulation ameliorates T cell exhaustion induced by tonic signaling of chimeric antigen receptors. *Nat Med* **21**, 581-590 (2015).
11. Majzner, R.G. et al. Tuning the Antigen Density Requirement for CAR T-cell Activity. *Cancer Discov* **10**, 702-723 (2020).
12. Haso, W. et al. Anti-CD22-chimeric antigen receptors targeting B-cell precursor acute lymphoblastic leukemia. *Blood* **121**, 1165-1174 (2013).
13. DeGolier, K.R. et al. Rational redesign of antigen binding domain improves in vivo efficacy of the CD22-CAR. *Mol Ther* **34**, 2175-2188 (2026).
14. Goodman, D.B. et al. Pooled screening of CAR T cells identifies diverse immune signaling domains for next-generation immunotherapies. *Sci Transl Med* **14**, eabm1463 (2022).
15. Hudecek, M. et al. The nonsignaling extracellular spacer domain of chimeric antigen receptors is decisive for in vivo antitumor activity. *Cancer Immunol Res* **3**, 125-135 (2015).
16. DeGolier, K.R. et al. Antigen experience history directs distinct functional states of CD8(+) CAR T cells during the antileukemia response. *Nat Immunol* **26**, 68-81 (2025).
17. Philip, M. et al. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* **545**, 452-456 (2017).

18. Scott-Browne, J.P. et al. Dynamic Changes in Chromatin Accessibility Occur in CD8(+) T Cells Responding to Viral Infection. *Immunity* **45**, 1327-1340 (2016).
19. Wang, W. et al. TCF-1 promotes chromatin interactions across topologically associating domains in T cell progenitors. *Nature Immunology* **23**, 1052-1062 (2022).
20. Pierce, S.E., Granja, J.M. & Greenleaf, W.J. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nature Communications* **12**, 2969 (2021).
21. Chen, J. et al. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530-534 (2019).
22. Ghaffari, S. et al. Optimizing interleukin-2 concentration, seeding density and bead-to-cell ratio of T-cell expansion for adoptive immunotherapy. *BMC Immunol* **22**, 43 (2021).
23. Kaartinen, T. et al. Low interleukin-2 concentration favors generation of early memory T cells over effector phenotypes during chimeric antigen receptor T-cell expansion. *Cytotherapy* **19**, 689-702 (2017).
24. Pipkin, M.E. et al. Interleukin-2 and inflammation induce distinct transcriptional programs that promote the differentiation of effector cytolytic T cells. *Immunity* **32**, 79-90 (2010).
25. Hughes, D.P. et al. Fate and function of anti-CD3/CD28-activated T cells following adoptive transfer: IL-2 promotes development of anti-tumor memory T cells in vivo. *Cytotherapy* **7**, 396-407 (2005).
26. Schep, A.N., Wu, B., Buenrostro, J.D. & Greenleaf, W.J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nature Methods* **14**, 975-978 (2017).
27. Kahan, S.M. et al. Intrinsic IL-2 production by effector CD8 T cells affects IL-2 signaling and promotes fate decisions, stemness, and protection. *Sci Immunol* **7**, eabl6322 (2022).
28. Cimons, J.M. et al. T-bet overexpression enhances CAR T cell effector functions and antigen sensitivity. *J Immunother Cancer* **13** (2025).
29. Lynn, R.C. et al. c-Jun overexpression in CAR T cells induces exhaustion resistance. *Nature* **576**, 293-300 (2019).
30. Shah, N.N. et al. CD4/CD8 T-Cell Selection Affects Chimeric Antigen Receptor (CAR) T-Cell Potency and Toxicity: Updated Results From a Phase I Anti-CD22 CAR T-Cell Trial. *J Clin Oncol* **38**, 1938-1950 (2020).
31. Heitzeneder, S. et al. GPC2-CAR T cells tuned for low antigen density mediate potent activity against neuroblastoma without toxicity. *Cancer Cell* **40**, 53-69.e59 (2022).
32. Walker, A.J. et al. Tumor Antigen and Receptor Densities Regulate Efficacy of a Chimeric Antigen Receptor Targeting Anaplastic Lymphoma Kinase. *Mol Ther* **25**, 2189-2201 (2017).
33. Watanabe, K. et al. Target antigen density governs the efficacy of anti-CD20-CD28-CD3 ζ chimeric antigen receptor-modified effector CD8+ T cells. *J Immunol* **194**, 911-920 (2015).
34. Chang, T.C. et al. BACH2 regulates T cell lineage state to enhance CAR T cell function. *Nat Immunol* **27**, 413-424 (2026).
35. Daniels, K.G. et al. Decoding CAR T cell phenotype using combinatorial signaling motif libraries and machine learning. *Science* **378**, 1194-1200 (2022).

36. Datlinger, P. et al. Systematic discovery of CRISPR-boosted CAR T cell immunotherapies. *Nature* **646**, 963-972 (2025).
37. Knudsen, N.H. et al. In vivo CRISPR screens identify modifiers of CAR T cell function in myeloma. *Nature* **646**, 953-962 (2025).
38. Perez, C.R. et al. Library-based single-cell analysis of CAR signaling reveals drivers of in vivo persistence. *Cell Systems* **16**, 101260 (2025).
39. Göransdotter Ericson, K. et al. Spectrum of Perforin Gene Mutations in Familial Hemophagocytic Lymphohistiocytosis. *The American Journal of Human Genetics* **68**, 590-597 (2001).
40. Stepp, S.E. et al. Perforin Gene Defects in Familial Hemophagocytic Lymphohistiocytosis. *Science* **286**, 1957-1959 (1999).
41. Choi, J.H. et al. Optimization of AAV expression cassettes to improve packaging capacity and transgene expression in neurons. *Mol Brain* **7**, 17 (2014).
42. Hawkins, J.A., Jones, S.K., Finkelstein, I.J. & Press, W.H. Indel-correcting DNA barcodes for high-throughput sequencing. *Proceedings of the National Academy of Sciences* **115**, E6217-E6226 (2018).
43. Buschmann, T. & Bystrykh, L.V. Levenshtein error-correcting barcodes for multiplexed DNA sequencing. *BMC Bioinformatics* **14**, 272 (2013).
44. Granja, J.M. et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature Genetics* **53**, 403-411 (2021).
45. Cusanovich, D.A. et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910-914 (2015).
46. Granja, J.M. et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol* **37**, 1458-1465 (2019).
47. Satpathy, A.T. et al. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat Biotechnol* **37**, 925-936 (2019).
48. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods* **16**, 1289-1296 (2019).
49. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology* **42**, 293-304 (2024).
50. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
51. Vierstra, J. et al. Global reference mapping of human transcription factor footprints. *Nature* **583**, 729-736 (2020).
52. Casper, J. et al. The UCSC Genome Browser database: 2026 update. *Nucleic Acids Res* **54**, D1331-d1335 (2026).

FIGURES

Figure 1

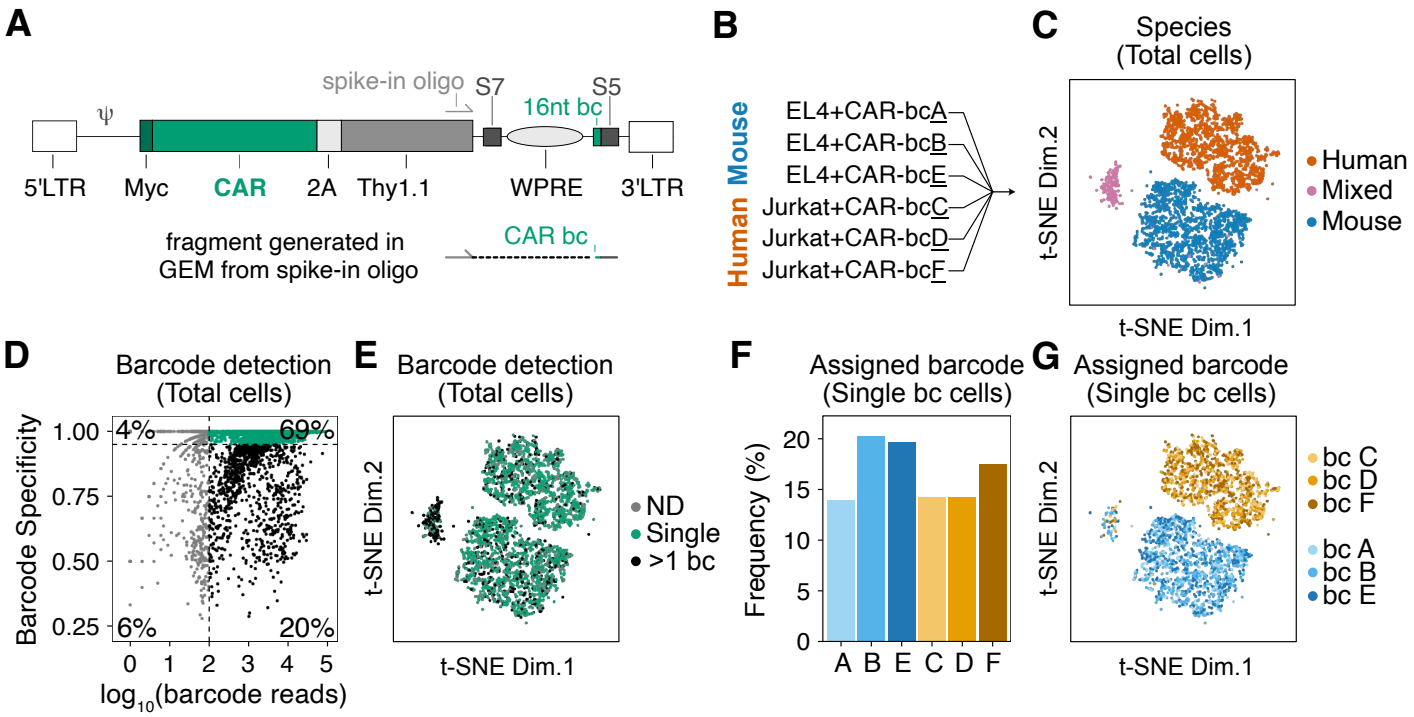


Figure 1: Species-associated barcoding to optimize *in silico* filtering parameters for genetically-barcoded scATAC-seq

1A: Diagram of gammaretroviral murine CAR construct with barcode cassette. Constructs contain 5'UTR, Myc-tagged CAR transgene, P2A, murine Thy1.1 reporter, barcode cassette containing WPRE3 and 16bp DNA barcode flanked by Nextera sequencing adapters, and self-inactivating (SIN) 3'LTR. A spike-in oligo primes amplification of the barcode fragment during scATAC capture to facilitate later enrichment of barcode library. **1B:** Experimental schematic: Three distinctly-barcoded CAR constructs per species were transduced into either murine EL4 T cell lymphoma or human Jurkat T cell leukemia lines before pooling and capture for scATAC-seq. **1C:** t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization of all pooled cells, colored by species. **1D:** Barcode detection in all pooled cells by scATAC-seq, with cell filtering thresholds indicated for both 'barcode specificity' (proportion of barcode reads matching barcode with most reads) and $\log_{10}$(# barcode reads per cell). Cells in top right quadrant (green) meet inclusion thresholds. **1E:** t-SNE visualization of cells that met QC cutoffs, colored by inclusion criteria ('Single' have been confidently assigned a barcode, '>1 bc' are barcode multiplets, and 'ND' are not designated). **1F:** Frequency of cells meeting inclusion criteria (Single) which were assigned to each barcode. **1G:** t-SNE visualization of cells meeting inclusion criteria (Single), colored by barcode.

Figure 2

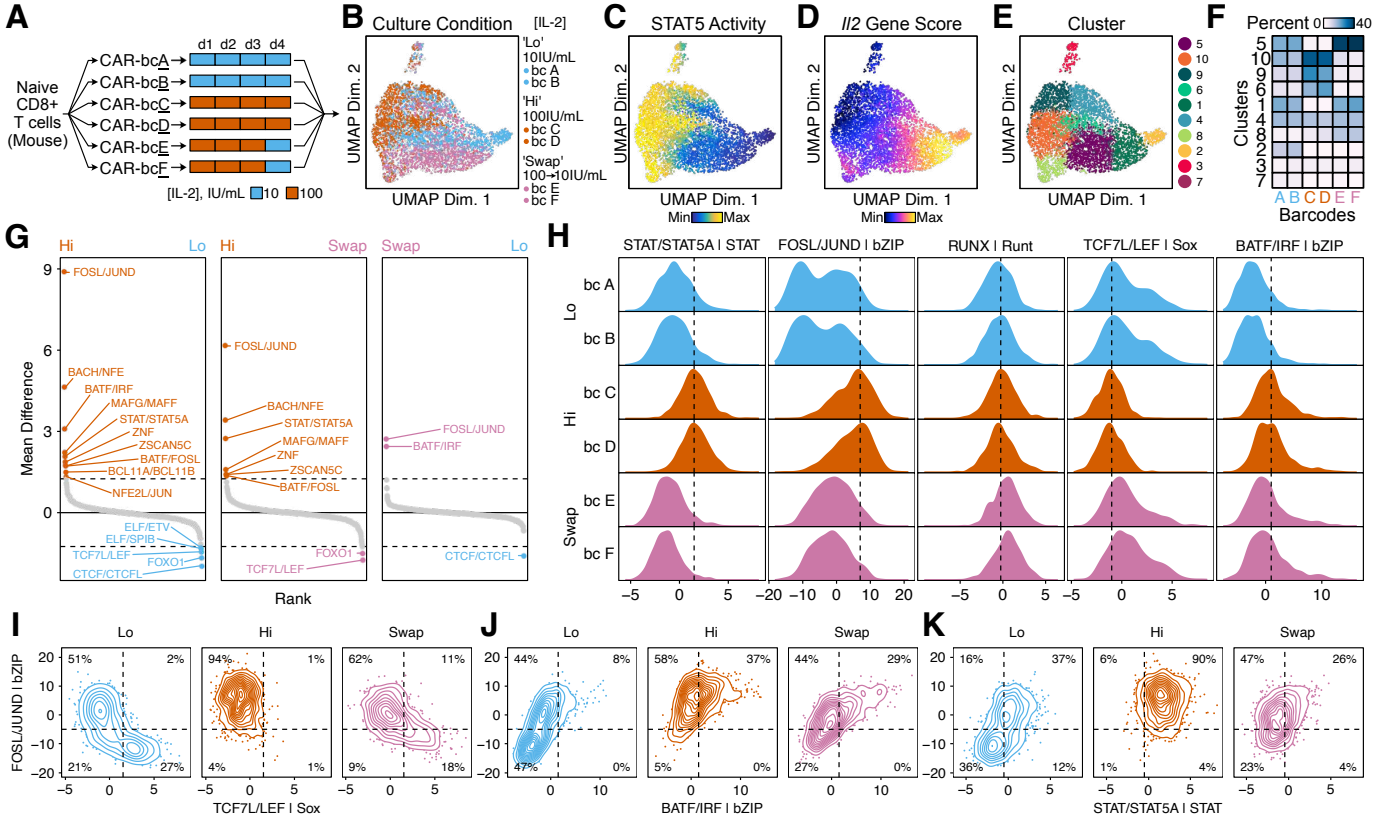

A
Naive CD8+ T cells (Mouse)
CAR-bcA
CAR-bcB
CAR-bcC
CAR-bcD
CAR-bcE
CAR-bcF
d1 d2 d3 d4
[IL-2], IU/mL 10 100
B
Culture Condition
[IL-2]
'Lo' 10IU/mL
bc A
bc B
'Hi' 100IU/mL
bc C
bc D
'Swap' 100→10IU/mL
bc E
bc F
UMAP Dim. 1
UMAP Dim. 2
C
STAT5 Activity
D
Il2 Gene Score
E
Cluster
F
Percent 0 40
Clusters
Barcodes
A B C D E F
G
Hi Lo
Hi Swap
Swap Lo
Mean Difference
FOSL/JUND
BACH/NFE
BATF/IRF
MAFG/MAFF
STAT/STAT5A
ZNF
ZSCAN5C
BATF/FOSL
BCL11A/BCL11B
NFE2L/JUN
ELF/ETV
ELF/SPIB
TCF7/LEF
FOXO1
CTCF/CTCFL
Rank
H
STAT/STAT5A | STAT
FOSL/JUND | bZIP
RUNX | Runt
TCF7L/LEF | Sox
BATF/IRF | bZIP
Lo
Hi
Swap
bc A
bc B
bc C
bc D
bc E
bc F
I
Lo
Hi
Swap
FOSL/JUND | bZIP
TCF7L/LEF | Sox
J
BATF/IRF | bZIP
K
STAT/STAT5A | STAT

Figure 2: Orthogonal *In vitro* validation in primary murine CAR T cells using culture condition-associated barcoding reveals stable and transient transcriptional programs dictated by IL-2 concentration

2A: Experimental schematic: Two distinctly barcoded CAR constructs per IL-2 condition ('Lo,' 10IU/mL; 'Hi,' 100IU/mL; 'Swap,' 100IU/mL changed to 10IU/mL 16 hours prior to harvest) were transduced into murine CD8 T cells before pooling and capture for scATAC-seq. **2B-E:** Uniform Manifold Approximation and Projection (UMAP) visualization of all pooled cells, colored by IL-2 condition **(B)**, ChromVAR-computed STAT5 motif-associated chromatin accessibility (z-score) **(C)**, *Il2* gene score **(D)**, or Seurat-computed cluster **(E)**. **2F:** Proportion of cells assigned each barcode per cluster (percentages computed by column). **2G:** Difference in mean (deviation) of ChromVAR z-scores across the three pairwise comparisons. **2H-K:** ChromVAR-computed motif-associated chromatin accessibility z-scores for all cells assigned each barcode, for the indicated transcription factor motifs **(H)**, or the indicated pairwise motif comparisons **(I-K)**.

Figure 3

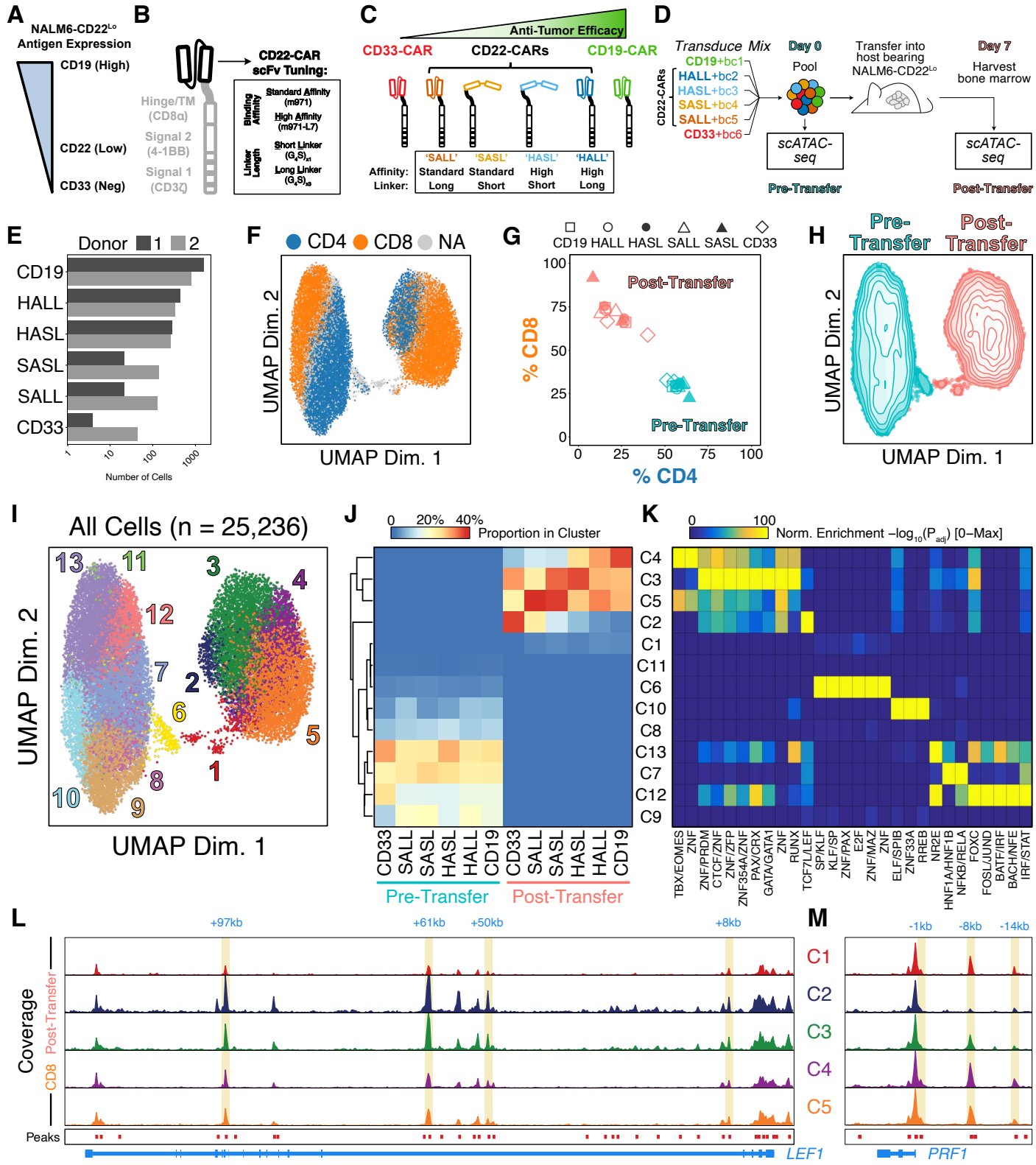

A
NALM6-CD22Lo
Antigen Expression
CD19 (High)
CD22 (Low)
CD33 (Neg)
B
Hinge/TM (CD8α)
Signal 2 (4-1BB)
Signal 1 (CD3ζ)
CD22-CAR scFv Tuning:
Binding Affinity
Standard Affinity (m971)
High Affinity (m971-L7)
Linker Length
Short Linker (G4S)x1
Long Linker (G4S)x3
C
Anti-Tumor Efficacy
CD33-CAR
CD22-CARs
CD19-CAR
'SALL' 'SASL' 'HASL' 'HALL'
Affinity: Standard Standard High High
Linker: Long Short Short Long
D
Transduce Mix
CD19+bc1
HALL+bc2
HASL+bc3
SASL+bc4
SALL+bc5
CD33+bc6
CD22-CARs
Day 0
Pool
scATAC-seq
Pre-Transfer
Transfer into host bearing NALM6-CD22Lo
Day 7
Harvest bone marrow
scATAC-seq
Post-Transfer
E
Donor 1 2
CD19
HALL
HASL
SASL
SALL
CD33
Number of Cells
F
CD4 CD8 NA
UMAP Dim. 2
UMAP Dim. 1
G
CD19 HALL HASL SALL SASL CD33
Post-Transfer
Pre-Transfer
% CD8
% CD4
H
Pre-Transfer
Post-Transfer
UMAP Dim. 2
UMAP Dim. 1
I
All Cells (n = 25,236)
UMAP Dim. 2
UMAP Dim. 1
J
Proportion in Cluster
CD33 SALL SASL HASL HALL CD19 CD33 SALL SASL HASL HALL CD19
Pre-Transfer
Post-Transfer
K
Norm. Enrichment −log10(Padj) [0–Max]
C4 C3 C5 C2 C1 C11 C6 C10 C8 C13 C7 C12 C9
TBX/EOMES ZNF ZNF/PRDM CTCF/ZNF ZNF/ZFP ZNF354A/ZNF PAX/CRX GATA/GATA1 ZNF RUNX TCF7L/LEF SP/KLF KLF/SP ZNF/PAX E2F ZNF/MAZ ZNF ELF/SPIB ZNF33A RREB NR2E HNF1A/HNF1B NFKB/RELA FOXC FOSL/JUND BATF/IRF BACH/NFE IRF/STAT
L
Coverage
Post-Transfer
CD8
Peaks
+97kb
+61kb
+50kb
+8kb
LEF1
M
C1
C2
C3
C4
C5
-1kb
-8kb
-14kb
PRF1

Figure 3: *In vitro and ex vivo* coupling of CD22-CAR architecture-associated barcodes with chromatin accessibility in human T cells exposes temporal and efficacy-linked chromatin accessibility signatures **3A-C:** Schematics depicting: Relative antigen expression on NALM6-CD22^{Lo} B-lineage acute lymphoblastic leukemia. **(A)**. Conserved CAR architecture and modifications to scFv binding affinity or tonic-signaling profile (linker length) for CD22-CARs. **(B)**. Previously observed efficacy against NALM6-CD22^{Lo} for CD22-CARs, benchmarked against the CD19-CAR (positive control) and CD33-CAR (negative control) **(C)**. **3D:** Experimental schematic: Distinctly barcoded CARs with the architectures described in 3C were transduced into primary human T cells and pooled at equal ratios for scATAC-seq profiling ('Pre-Transfer) or adoptively transferred (1e6 CAR+ cells/mouse) into mice which had received 1e6 NALM6-CD22^{Lo} 3 days prior. 7 days after CAR transfer, bone marrow was harvested and pooled for scATAC-seq profiling ('Post-Transfer'). Data in **3E-3K** are from pooled T cells from both donors and both timepoints (n = 25,236 cells from n = 2 biological donor replicates, profiled pre- and post-transfer) **3E:** T cell abundance for each CAR in the scATAC-seq pool, separated by donor. **3F:** Uniform Manifold Approximation and Projection (UMAP) visualization of coreceptor assignment (CD4, CD8 or not assigned (NA)) **3G:** Proportion of cells with each coreceptor assignment, stratified by CAR barcode assignment and timepoint. **3H-I:** UMAP colored by timepoint **(H)** or Seurat-computed cluster **(I)**. **3J:** Proportion of cells assigned each barcode across clusters (percentages computed by column). **3K:** Normalized enrichment by column ($-\log10(P_{adj})$) for motif-associated chromatin accessibility (z-scores) within differentially accessible peaks for each cluster, for top 6 motifs which were differentially accessible within each cluster (can be shared across clusters). **3L-M:** Genomic chromatin accessibility tracks surrounding the indicated gene loci for CD8 CAR T cells, pseudobulked and stratified by cluster. Boxes indicate peaks which were differentially accessible between CD8 CAR T cells within Clusters 3 and 4 at putative or established regulatory elements surrounding each gene locus. Distance from TSS indicated above.

Figure 4

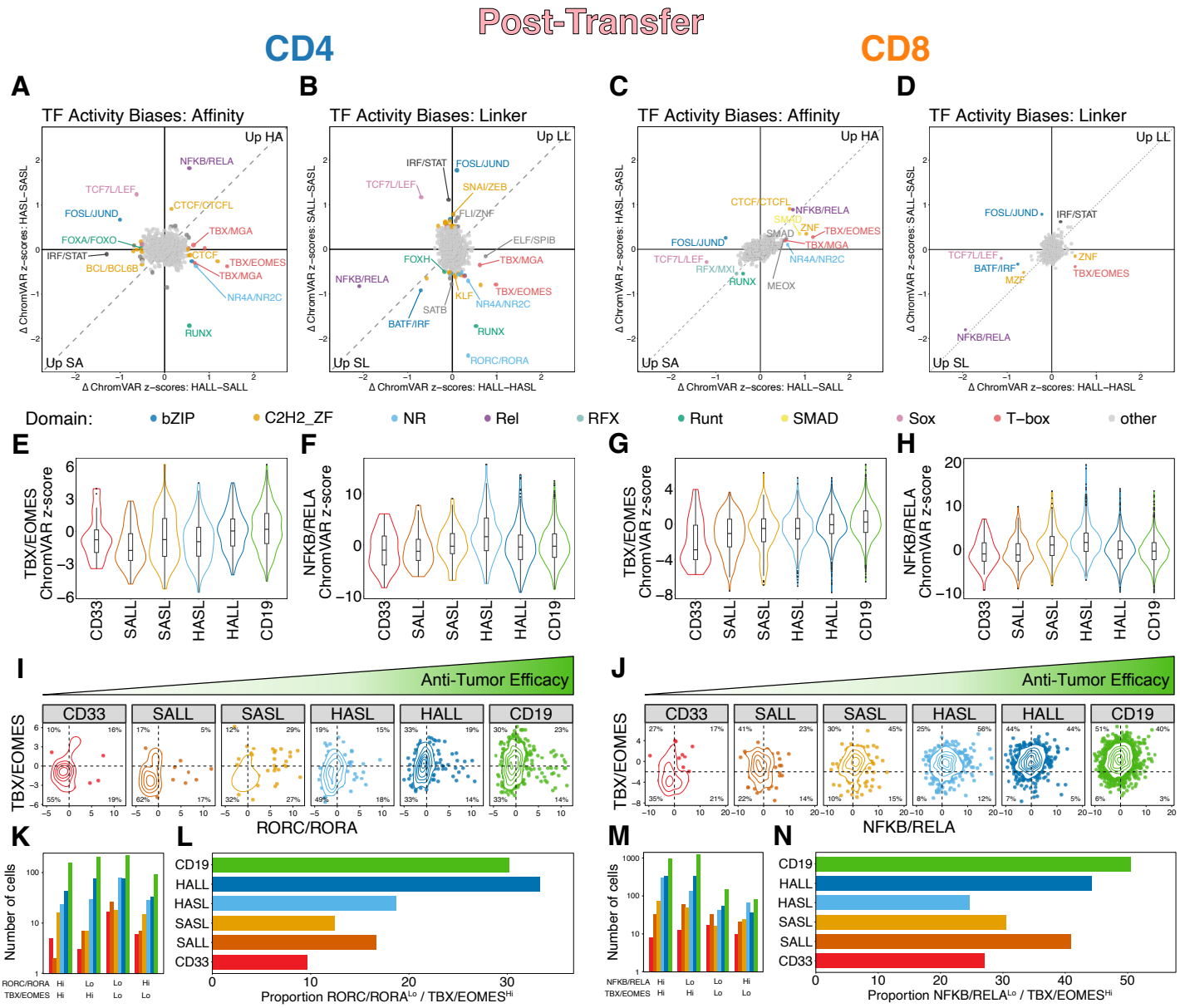

Post-Transfer
CD4
CD8
A
TF Activity Biases: Affinity
Up HA
Up SA
NFKB/RELA
TCF7L/LEF
CTCF/CTCFL
FOSL/JUND
TBX/MGA
FOXA/FOXO
CTCF
IRF/STAT
TBX/EOMES
BCL/BCL6B
TBX/MGA
NR4A/NR2C
RUNX
Δ ChromVAR z−scores: HASL−SASL
Δ ChromVAR z−scores: HALL−SALL
B
TF Activity Biases: Linker
Up LL
Up SL
IRF/STAT
FOSL/JUND
TCF7L/LEF
SNAI/ZEB
FLI/ZNF
ELF/SPIB
FOXH
TBX/MGA
NFKB/RELA
KLF
TBX/EOMES
NR4A/NR2C
SATB
BATF/IRF
RUNX
RORC/RORA
Δ ChromVAR z−scores: SALL−SASL
Δ ChromVAR z−scores: HALL−HASL
C
TF Activity Biases: Affinity
Up HA
Up SA
CTCF/CTCFL
NFKB/RELA
SMAD
ZNF
TBX/EOMES
SMAD
FOSL/JUND
TBX/MGA
TCF7L/LEF
NR4A/NR2C
RFX/MXL
RUNX
MEOX
Δ ChromVAR z−scores: HASL−SASL
Δ ChromVAR z−scores: HALL−SALL
D
TF Activity Biases: Linker
Up LL
Up SL
FOSL/JUND
IRF/STAT
TCF7L/LEF
ZNF
BATF/IRF
TBX/EOMES
MZF
NFKB/RELA
Δ ChromVAR z−scores: SALL−SASL
Δ ChromVAR z−scores: HALL−HASL
Domain:
bZIP
C2H2_ZF
NR
Rel
RFX
Runt
SMAD
Sox
T−box
other
E
TBX/EOMES ChromVAR z-score
CD33
SALL
SASL
HASL
HALL
CD19
F
NFKB/RELA ChromVAR z-score
G
TBX/EOMES ChromVAR z-score
H
NFKB/RELA ChromVAR z-score
I
Anti-Tumor Efficacy
TBX/EOMES
RORC/RORA
J
Anti-Tumor Efficacy
TBX/EOMES
NFKB/RELA
K
Number of cells
RORC/RORA
TBX/EOMES
L
Proportion RORC/RORA^Lo / TBX/EOMES^Hi
M
Number of cells
NFKB/RELA
TBX/EOMES
N
Proportion NFKB/RELA^Lo / TBX/EOMES^Hi

Figure 4: CD22-CAR architectures drive distinct *in vivo* NFκB and effector-associated transcriptional programs during the anti-leukemia response

Data on left side of page (panels A, B, E, F, I, K, L, O, P) are from CD4 CAR T cells and data on right side of page (panels C, D, G, H, J, M, N, Q, R) are from CD8 CAR T cells. **4A-D:** CD22-CAR architectural biases in transcriptional activity for transcription factors binding the indicated motifs, driven by either CD22-CAR affinity **(A, C)** or linker **(B, D)**. Coordinates are calculated as the difference between ChromVAR motif-associated chromatin accessibility z-scores for the comparisons as indicated on each axis. Points closer to the y = x diagonal line are more directly biased by affinity or linker as indicated, while points away from the diagonal exhibit preferential bias by either the x-axis or y-axis comparison. Cutoff value of >0.5 in any direction was applied. **4E-H:** Comparisons of ChromVAR motif-associated chromatin accessibility z-scores for all cells assigned each barcode for the indicated motifs. **4I-J**: Pairwise comparisons of ChromVAR motif-associated chromatin accessibility z-scores for all cells assigned each barcode for the indicated sets of motifs. **4K:** Quantifications of total cell numbers in each quadrant, stratified by CAR. **4L:** Proportions of each CAR in the indicated quadrant. **4M:** Quantifications of total cell numbers in each quadrant, stratified by CAR. **4N:** Proportions of each CAR in the indicated quadrant.

EXTENDED DATA

Figure S1

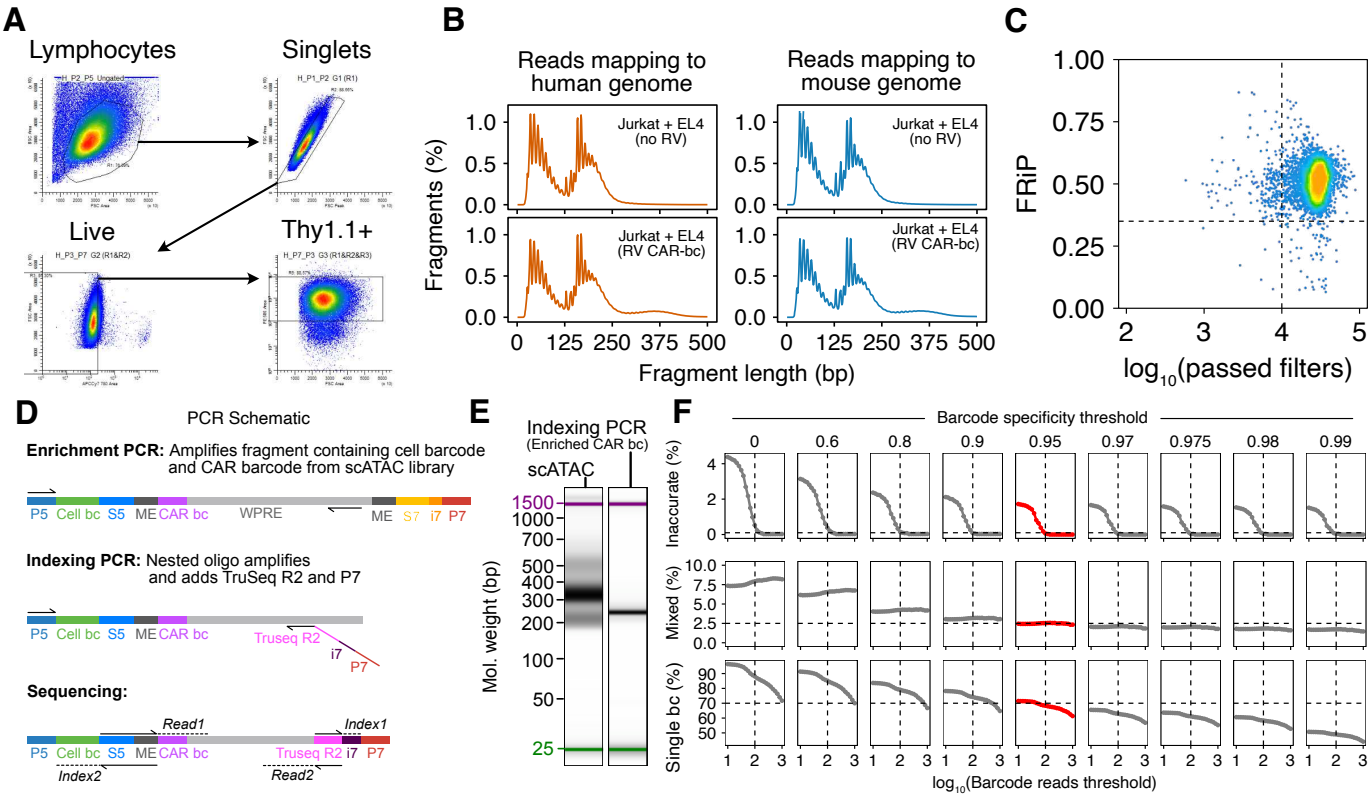
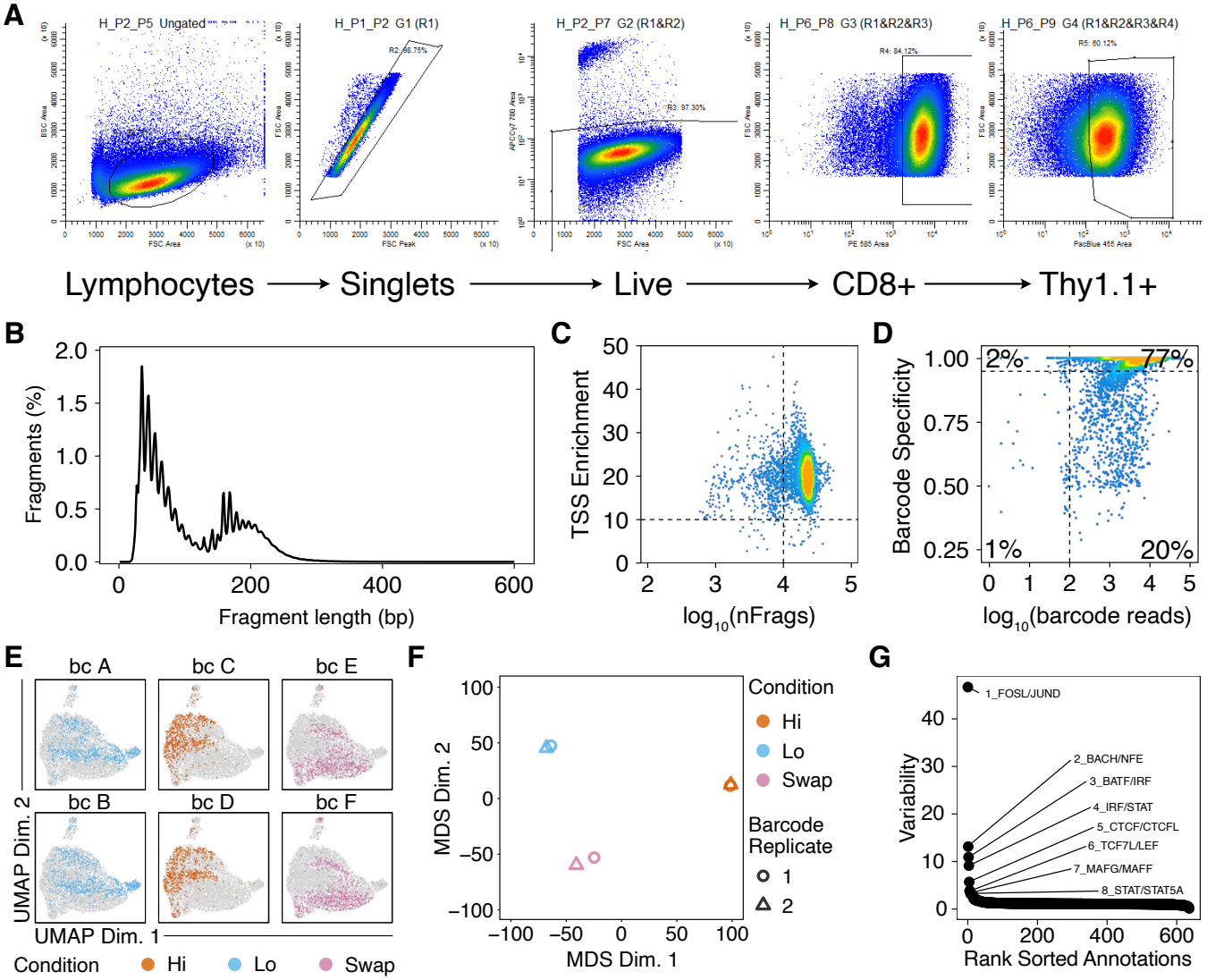


Figure S1: Further characterization of species-associated barcoding to optimize *in silico* filtering parameters

S1A: FACS gating schematic for the sorting of barcoded EL4 and Jurkat cell lines (Lymphocytes>Singlets>Live>mThy1.1+). **S1B:** Distribution of fragment sizes for reads mapping to human GRChg38 (left column) and mouse GRCmm10 (right column) reference genomes for untransduced EL4 and Jurkat pool (top row) and barcoded CAR-transduced EL4 and Jurkat pool (bottom row). All data for **S1C**-**S1G** is from the barcoded CAR-transduced EL4 and Jurkat pool. **S1C:** Fraction of Reads in Peaks (FRiP) by $\log_{10}$(# of reads passing filters). **S1D:** PCR Schematic: Two step nested PCR approach to 1) enrich for DNA fragments containing the introduced DNA barcode from the final scATAC library and 2) add a sample index and TruSeq adapter sequence to create the final 'barcode library' to be sequenced alongside the scATAC library. **S1E:** Agilent Tapestation analysis showing fragment size distribution for scATAC library (left) and final 'barcode library' after sample indexing PCR (right). **S1F:** Plots showing % of included cells with inaccurate species assignments (top row), containing mixed barcodes (middle row), and containing a single barcode (bottom row) on the y-axis, plotted by the barcode reads threshold ($\log_{10}$(# of barcode reads matching barcode with most reads)) for a range of 'barcode specificity' thresholds (proportion of barcode reads matching barcode with most reads) across each row. **S1G:** Plots of $\log_{10}$(# reads mapping to GRCmm10) by $\log_{10}$(# reads mapping to GRChg38). "All cells" plot contains all cells prior to application of 100 reads sensitivity and 0.95 specificity filtering thresholds, while ">1 Barcode" plot contains cells that had more than one barcode prior to filtering. Plots for each barcode contain cells which were assigned the indicated barcode after filtering cutoffs were applied.

Figure S2



851

Figure S2: Quality control and characterization of IL-2 culture condition-associated barcoding of primary murine CAR T cells

S2A: FACS gating schematic for the sorting of barcoded murine CAR T cells (Lymphocytes > Singlets > Live > CD8+ > mThy1.1+). **S2B:** Distribution of fragment sizes. **S2C:** $\log_{10}$(number of fragments) by TSS enrichment. **S2D:** $\log_{10}$(# of barcode reads matching barcode with most reads) by barcode specificity (proportion of barcode reads matching barcode with most reads), with filtering cutoffs indicated. **S2E:** UMAP visualization of cells assigned each barcode, colored by IL-2 condition. **S2F:** Multidimensional Scaling (MDS) dimensionality reduction indicating global similarity in chromatin accessibility profile between pseudobulked cells from each barcode replicate and by condition. Coloring indicates IL-2 condition. Shape indicates barcode replicate, where replicate 1 is barcodes A, C and E for Lo, Hi and Swap conditions, and replicate 2 is B, D and F for Lo, Hi and Swap conditions. **S2G:** Ranked list of most variable ChromVAR motif-associated chromatin accessibility z-scores between all conditions.

Figure S3

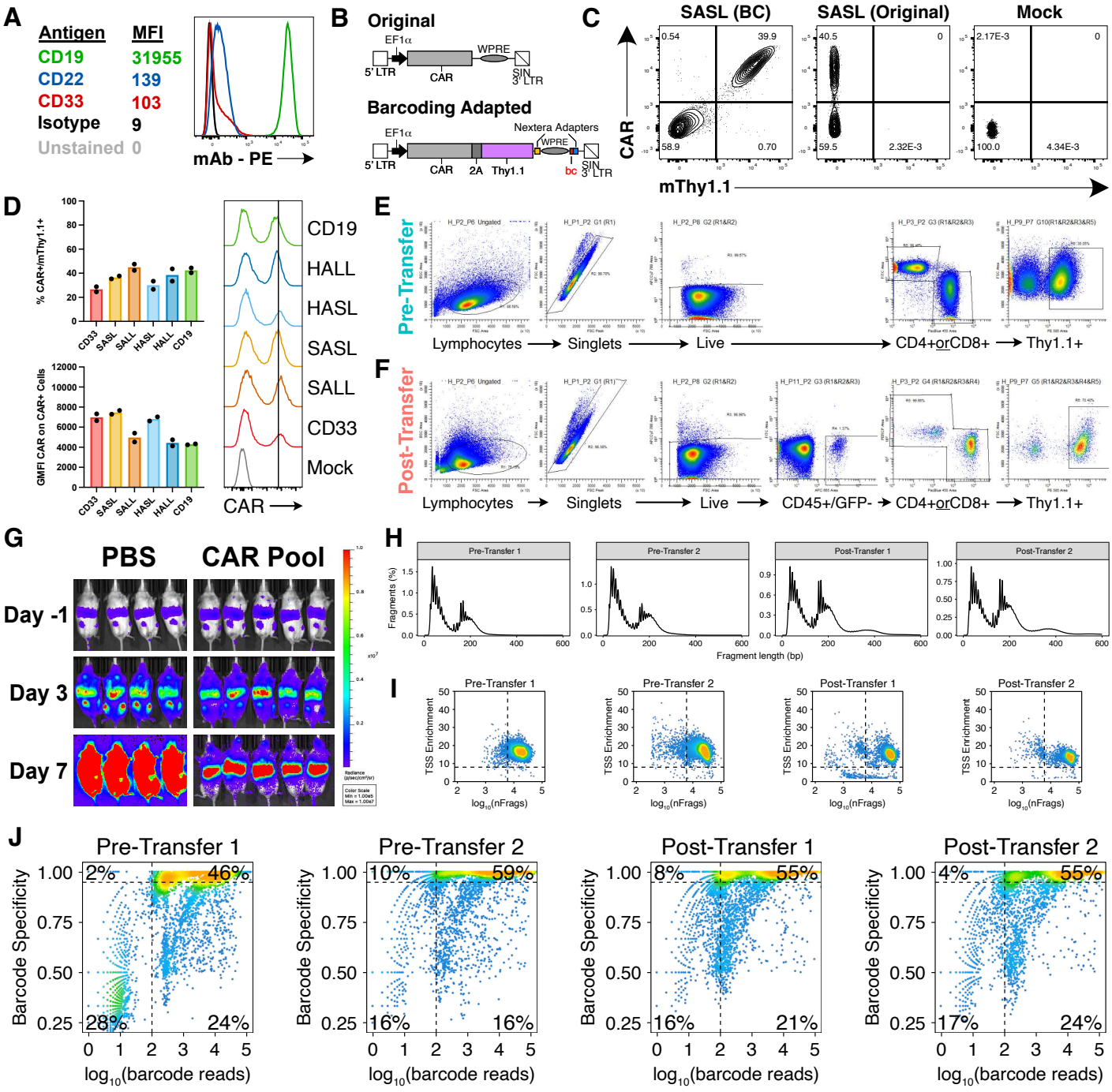

A
Antigen MFI
CD19 31955
CD22 139
CD33 103
Isotype 9
Unstained 0
mAb - PE
B
Original
EF1α
5' LTR
CAR
WPRE
SIN 3' LTR
Barcoding Adapted
Nextera Adapters
2A
Thy1.1
bc
C
SASL (BC)
SASL (Original)
Mock
CAR
mThy1.1
0.54
39.9
58.9
0.70
40.5
0
59.5
2.32E-3
2.17E-3
0
100.0
4.34E-3
D
% CAR+/mThy1.1+
GMFI CAR on CAR+ Cells
CD33
SASL
SALL
HASL
HALL
CD19
CD19
HALL
HASL
SASL
SALL
CD33
Mock
CAR
E
Pre-Transfer
Lymphocytes
Singlets
Live
CD4+orCD8+
Thy1.1+
F
Post-Transfer
Lymphocytes
Singlets
Live
CD45+/GFP-
CD4+orCD8+
Thy1.1+
G
PBS
CAR Pool
Day -1
Day 3
Day 7
H
Pre-Transfer 1
Pre-Transfer 2
Post-Transfer 1
Post-Transfer 2
Fragments (%)
Fragment length (bp)
I
Pre-Transfer 1
Pre-Transfer 2
Post-Transfer 1
Post-Transfer 2
TSS Enrichment
log10(nFrags)
J
Pre-Transfer 1
Pre-Transfer 2
Post-Transfer 1
Post-Transfer 2
Barcode Specificity
log10(barcode reads)
2%
46%
28%
24%
10%
59%
16%
16%
8%
55%
16%
21%
4%
55%
17%
24%

Figure S3: Quality control and characterization of CD22-CAR architecture-associated barcoding of primary human T cells

S3A: Flow cytometric analysis of CD19, CD22 and CD33 antigen expression on NALM6-CD22^{Lo} leukemia, compared to unstained and isotype-stained controls. All antigens were stained with PE-conjugated antibodies, and used the same PE-conjugated isotype control. Mean fluorescence intensity (MFI) for each antigen is indicated on the right. **S3B:** Schematic: Original CAR constructs (top) in pLenti backbone contain 5'UTR, EF1a promoter, CAR transgene, WPRE3 and self-inactivating (SIN) 3'LTR. Barcoding adapted constructs in pLenti backbone contain 5'UTR, EF1a promoter, CAR transgene, P2A, murine Thy1.1 reporter, barcode cassette containing WPRE3 and 16bp DNA barcode flanked by Nextera sequencing adapters, and self-inactivating (SIN) 3'LTR. **S3C:** Flow cytometry showing SASL CAR expression (CD22 Protein-Fc) by mThy1.1 reporter expression for the Barcoding Adapted CAR ("BC"), original CAR compared to Mock (untransduced) T cells. **S3D:** Transduction efficiency (top, % CAR+/mThy1.1+) and geometric mean fluorescence intensity (GMFI, bottom) on CAR+ T cells from both donors. Representative flow plots shown to the right. To facilitate equivalent CAR molecule GMFI detection independent of binding affinity or detection reagent, CAR T cells were stained with Protein L. **S3E-F:** FACS gating schematic for the sorting of barcoded human CAR T cells at Pre-Transfer (Lymphocytes > Singlets > Live > CD4+ or CD8+ > mThy1.1+) **(E)** and Pre-Transfer (Lymphocytes > Singlets > Live > CD45+/GFP- > CD4+ or CD8+ > mThy1.1+) **(F)** timepoints. **S3G:** Bioluminescence imaging showing leukemia burden over time in NSG mice given 1e6 NALM6-CD22^{Lo} on Day -3 before treatment at Day 0 with PBS (left) or 1e6 pooled CAR T cells (1.67e5 of each CAR, shown in Figure 3D). **S3H-J** show the following QC metrics for both timepoints (Pre-Transfer, Post-Transfer) and both donors (1, 2): **S3H:** Distribution of fragment sizes. **S3I:** $\log_{10}$(number of fragments) by TSS enrichment. **S3J:** $\log_{10}$(# of barcode reads matching barcode with most reads) by barcode specificity (proportion of barcode reads matching barcode with most reads), with filtering cutoffs indicated.

Figure S4

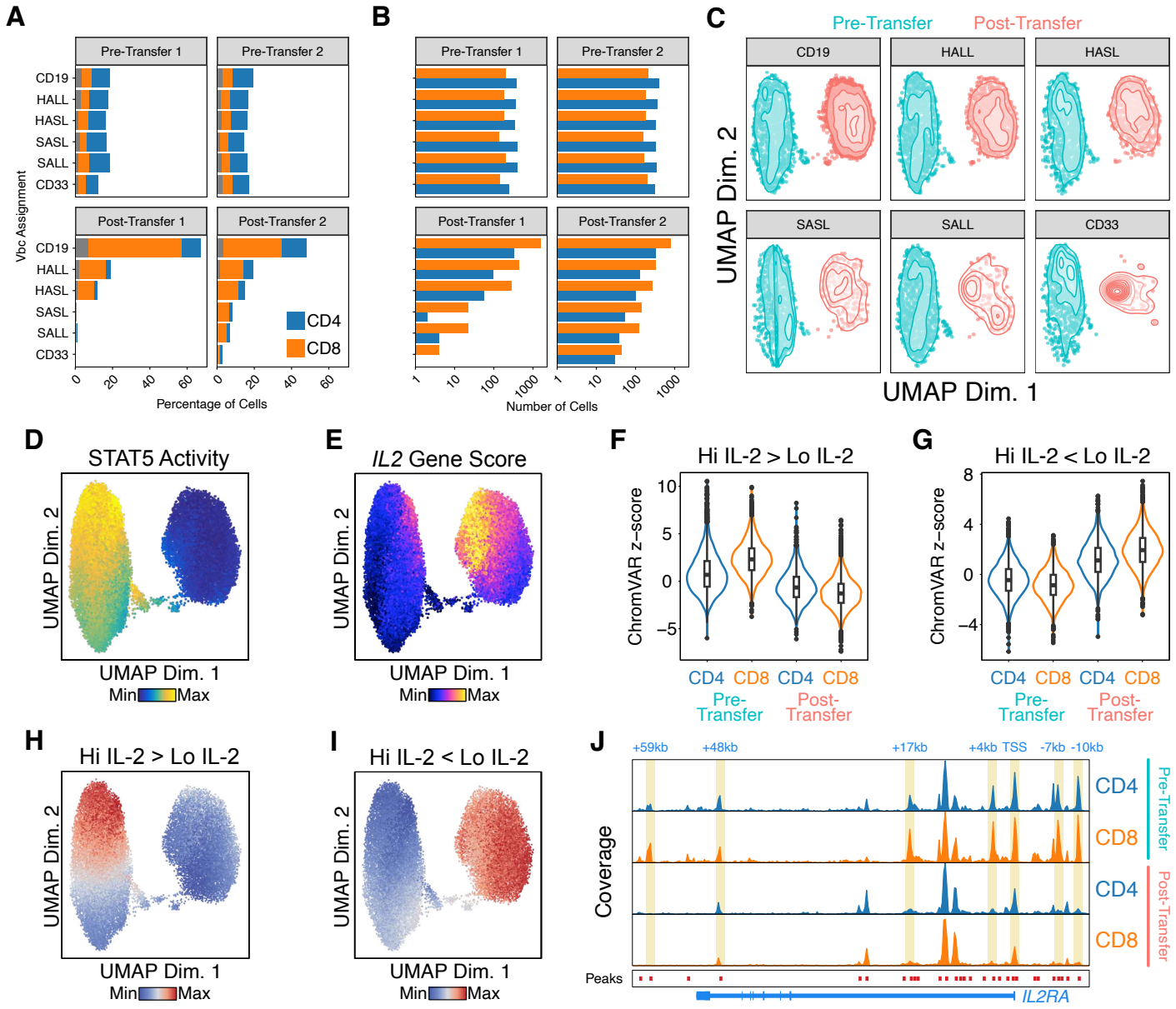

A
Pre-Transfer 1
Pre-Transfer 2
Post-Transfer 1
Post-Transfer 2
CD19
HALL
HASL
SASL
SALL
CD33
Vbc Assignment
CD4
CD8
Percentage of Cells
B
Number of Cells
C
Pre-Transfer
Post-Transfer
UMAP Dim. 2
UMAP Dim. 1
D
STAT5 Activity
UMAP Dim. 2
UMAP Dim. 1
Min
Max
E
IL2 Gene Score
F
Hi IL-2 > Lo IL-2
ChromVAR z−score
G
Hi IL-2 < Lo IL-2
H
I
J
+59kb
+48kb
+17kb
+4kb TSS
-7kb
-10kb
Coverage
Peaks
IL2RA

Figure S4: Temporal changes occurring from Pre- to Post-Transfer and murine-to-human IL-2 liftover analysis in primary human CAR T cells

S4A-B: Quantification of barcode and coreceptor assignments for pooled T cells at Pre-Transfer (top row) and Pre-Transfer (bottom row) timepoints, separated by donor replicates. Plots quantify proportions **(A)** and absolute numbers **(B)**. **S4C:** UMAP visualization of all cells, separated by CAR assignment and colored by timepoint. **S4D-E:** UMAP visualization of pooled cells from both timepoints, colored by STAT5 motif-associated chromatin accessibility (z-score) **(D)**, IL2 gene score **(E)**. **S4F-G:** Violin plots of ChromVAR z-scores quantifying enrichment of groups of peaks in human CAR T cell coreceptor/timepoint subgroups which were either more accessible in murine cells cultured in Hi compared to Lo IL-2 **(F)** or more accessible in murine cells cultured in Lo compared to Hi IL-2 **(G)**. **S4H-I:** UMAP visualization of the data in S4F (for panel **H**), or the data in S4G (for panel **I**). **S4J:** Genomic chromatin accessibility tracks surrounding the *IL2RA* locus for cells pseudobulked and stratified by timepoint (Pre-Transfer/Post-Transfer) and coreceptor (CD4/CD8). Boxes indicate peaks which were differentially accessible between timepoints at putative or established regulatory elements surrounding each gene locus. Distance from nearest TSS indicated.

Figure S5

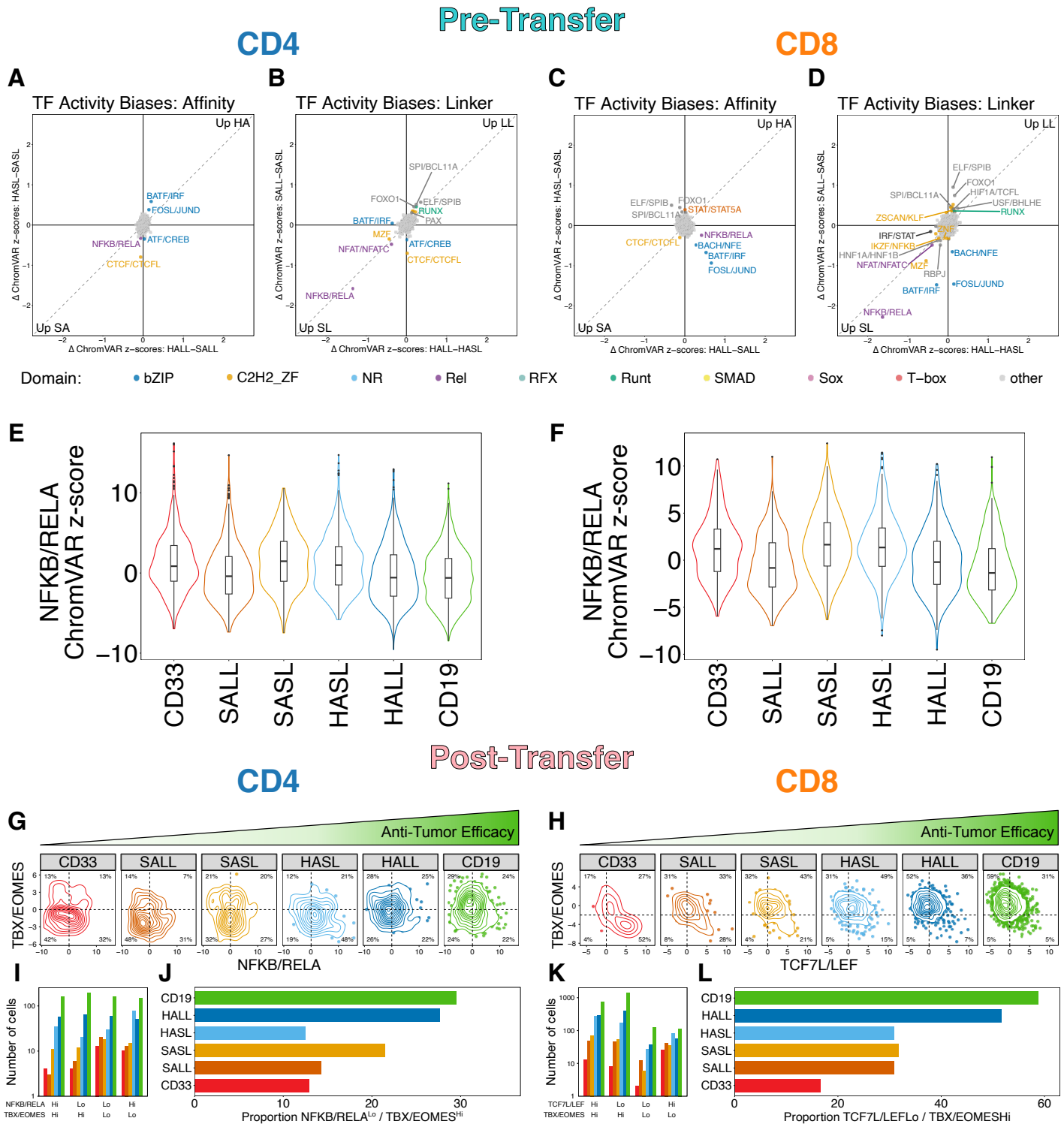

Pre-Transfer
CD4
CD8
A
TF Activity Biases: Affinity
Up HA
Up SA
BATF/IRF
FOSL/JUND
NFKB/RELA
ATF/CREB
CTCF/CTCFL
Δ ChromVAR z-scores: HASL−SASL
Δ ChromVAR z-scores: HALL−SALL
B
TF Activity Biases: Linker
Up LL
Up SL
SPI/BCL11A
FOXO1
ELF/SPIB
RUNX
BATF/IRF
PAX
MZF
NFAT/NFATC
ATF/CREB
CTCF/CTCFL
NFKB/RELA
Δ ChromVAR z-scores: SALL−SASL
Δ ChromVAR z-scores: HALL−HASL
C
TF Activity Biases: Affinity
Up HA
Up SA
ELF/SPIB
FOXO1
SPI/BCL11A
STAT/STAT5A
CTCF/CTCFL
NFKB/RELA
BACH/NFE
BATF/IRF
FOSL/JUND
Δ ChromVAR z-scores: HASL−SASL
Δ ChromVAR z-scores: HALL−SALL
D
TF Activity Biases: Linker
Up LL
Up SL
ELF/SPIB
FOXO1
SPI/BCL11A
HIF1A/TCFL
USF/BHLHE
RUNX
ZSCAN/KLF
ZNF
IRF/STAT
IKZF/NFKB
HNF1A/HNF1B
NFAT/NFATC
MZF
RBPJ
BACH/NFE
BATF/IRF
FOSL/JUND
NFKB/RELA
Δ ChromVAR z-scores: SALL−SASL
Δ ChromVAR z-scores: HALL−HASL
Domain:
bZIP
C2H2_ZF
NR
Rel
RFX
Runt
SMAD
Sox
T−box
other
E
NFKB/RELA ChromVAR z-score
CD33
SALL
SASL
HASL
HALL
CD19
F
NFKB/RELA ChromVAR z-score
CD33
SALL
SASL
HASL
HALL
CD19
Post-Transfer
CD4
CD8
G
Anti-Tumor Efficacy
CD33
SALL
SASL
HASL
HALL
CD19
TBX/EOMES
NFKB/RELA
H
Anti-Tumor Efficacy
CD33
SALL
SASL
HASL
HALL
CD19
TBX/EOMES
TCF7L/LEF
I
Number of cells
NFKB/RELA Hi Lo Lo Hi
TBX/EOMES Hi Hi Lo Lo
J
CD19
HALL
HASL
SASL
SALL
CD33
Proportion NFKB/RELALo / TBX/EOMESHi
K
Number of cells
TCF7L/LEF Hi Lo Lo Hi
TBX/EOMES Hi Hi Lo Lo
L
CD19
HALL
HASL
SASL
SALL
CD33
Proportion TCF7L/LEFLo / TBX/EOMESHi

Figure S5: CD22-CAR architectures drive distinct NFκB, TCF/LEF and effector-associated transcriptional programs prior to transfer and during the anti-leukemia response

Data on left side of page (panels A, B, E, G, I, J) are from CD4 CAR T cells and data on right side of page (panels C, D, F, H, K, L) are from CD8 CAR T cells. **S5A-D:** CD22-CAR architectural biases in transcriptional activity for transcription factors binding the indicated motifs, driven by either CD22-CAR affinity (**A**, **C**) or linker (**B**, **D**) at the Pre-Transfer timepoint. Coordinates are calculated as the difference between ChromVAR motif-associated chromatin accessibility z-scores for the comparisons as indicated on each axis. Points closer to the y = x diagonal line are more directly biased by affinity or linker as indicated, while points away from the diagonal exhibit preferential bias by either the x-axis or y-axis comparison. Cutoff value of >0.3 in any direction was applied. **S5E-F:** Comparisons of ChromVAR motif-associated chromatin accessibility z-scores for NFκB, for all cells assigned each barcode at the Pre-Transfer timepoint. **S5G-H:** Pairwise comparisons of ChromVAR motif-associated chromatin accessibility z-scores for all cells assigned each barcode for the indicated set of motifs. **S5I:** Quantifications of total cell numbers in each quadrant of S5G, stratified by CAR. **S5J:** Proportions of each CAR in the indicated quadrant of S5G. **S5K:** Quantifications of total cell numbers in each quadrant of S5H, stratified by CAR. **S5L:** Proportions of each CAR in the indicated quadrant of S5H.

Table S1

Oligo #	Name	Sequence	Length (bp)	Description
908	SI-PCRprimerB	AATGATACGGCGACCACCGAGA	22	Same as 10x ATAC SI-PCR primer b
925	PCR 2 primer 1 Spear ATAC	CAAGCAGAAGACGGCATACGAGAT ACGAGTAG GTGACTGGAGTTCAGACGTGTG	56	Spear ATAC - PCR 2 p7 indexed primer 1
926	PCR 2 primer 2 Spear ATAC	CAAGCAGAAGACGGCATACGAGAT ACCCGCAC GTGACTGGAGTTCAGACGTGTG	56	Spear ATAC - PCR 2 p7 indexed primer 2
927	PCR 2 primer 3 Spear ATAC	CAAGCAGAAGACGGCATACGAGAT AATTGAAC GTGACTGGAGTTCAGACGTGTG	56	Spear ATAC - PCR 2 p7 indexed primer 3
928	PCR 2 primer 4 Spear ATAC	CAAGCAGAAGACGGCATACGAGAT ACGCGGAA GTGACTGGAGTTCAGACGTGTG	56	Spear ATAC - PCR 2 p7 indexed primer 4
1032	Thy1.1 Spike-In Oligo	CCTGGACTTCATTTCTCTGTGATAA	25	Thy1.1-Specific Spike-In Oligo
1033	Enrich V2 JSB364 2	/5BiosG/GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTAGTTCTTGCCACGGCGGAAC	64	Barcode Enrichment Oligo. Specific to WPRE3.

*Note: Oligo # 1033 has 5' biotinylation to facilitate alternative enrichment strategies using Streptavidin pulldown as used in the original Spear-ATAC paper if necessary.
8nt sequence in middle of 92X primers is p7 index.

Table S1: List of oligonucleotides used in this study

Table S2

Barcode ID	Sequence	Note
A	CGCAATGCTGGCGCCT	Used in Fig 1, 2, S1, S2
B	GAACGCAGTGTTAACG	Used in Fig 1, 2, S1, S2
C	TGGAATACCCGACAGT	Used in Fig 1, 2, S1, S2
D	GACTTAGGTGTAGATC	Used in Fig 1, 2, S1, S2
E	GCTTTTCACGTAGTCT	Used in Fig 1, 2, S1, S2
F	CGTGTACTAAAGGCTC	Used in Fig 1, 2, S1, S2
1	GGCAGGTGATCGCAAT	Used in Fig 3, 4, S3, S4, S5
2	ATTATTGGCGTTGACC	Used in Fig 3, 4, S3, S4, S5
3	CTGGAGAAGAGTTGGC	Used in Fig 3, 4, S3, S4, S5
4	TTCTCCGCACCATTCG	Used in Fig 3, 4, S3, S4, S5
5	TCTGGACACACCGTAA	Used in Fig 3, 4, S3, S4, S5
6	TAGCATTGTCACCTTC	Used in Fig 3, 4, S3, S4, S5

Table S2: List of DNA barcodes used in this study