Designed improvement to T-cell immunotherapy by multidimensional single cell profiling

Irfan N Bandey,1 Jay R T Adolacion,1 Gabrielle Romain,1 Melissa Martinez Paniagua,1 Xingyue An,1 Arash Saeedi,1 Ivan Liadi,1 Zheng You,1 Rasindu B Rajanayake,2 Patrick Hwu,3 Harjeet Singh,4 Laurence JN Cooper,4,5 Navin Varadarajan1

ABSTRACT

Background Adoptive cell therapy based on the infusion of chimeric antigen receptor (CAR) T cells has shown remarkable efficacy for the treatment of hematologic malignancies. The primary mechanism of action of these infused T cells is the direct killing of tumor cells expressing the cognate antigen. However, understanding why only some T cells are capable of killing, and identifying mechanisms that can improve killing has remained elusive.

Methods To identify molecular and cellular mechanisms that can improve T-cell killing, we utilized integrated high-throughput single-cell functional profiling by microscopy, followed by robotic retrieval and transcriptional profiling.

Results With the aid of mathematical modeling we demonstrate that non-killer CAR T cells comprise a heterogeneous population that arise from failure in each of the discrete steps leading to the killing. Differential transcriptional single-cell profiling of killers and non-killers identified CD137 as an inducible costimulator molecule upregulated on killer T cells. Our single-cell profiling results directly demonstrate that inducible CD137 is feature of killer (and serial killer) T cells and this marks a different subset compared with the CD107apos (degranulating) subset of CAR T cells. Ligation of the induced CD137 with CD137 ligand (CD137L) leads to younger CAR T cells with sustained killing and lower exhaustion. We genetically modified CAR T cells to co-express CD137L, in trans, and this lead to a profound impact in anti-tumor efficacy in leukemia and refractory ovarian cancer models in mice.

Conclusions Broadly, our results illustrate that while non-killer T cells are reflective of population heterogeneity, integrated single-cell profiling can enable identification of mechanisms that can enhance the function/proliferation of killer T cells leading to direct anti-tumor benefit.

INTRODUCTION

Remarkable progress in the last three decades investigating the use of T cells as living drugs has culminated in the approval of two types of chimeric antigen receptor (CAR) T cells for the treatment of leukemias and lymphomas.1 This clinical success targeting CD19 for hematologic malignancies has launched the field of CAR T cell (and NK (natural killer) cells) for targeting other antigens in both liquid and solid tumors, and there are >200 clinical trials investigating CAR T cell-based therapies.2 This has opened up several challenges associated with the manufacturing of the newer generation of CAR T cells, including the discovery of tumor-associated antigens; optimizing CAR structure, activation, and affinity; and controlling the expression of the CAR transgene.3,4 CAR T cell-based therapy is at a critical juncture, and important advances in CAR T-cell biology will have a significant impact on determining whether this mode of treatment will be restricted to specific cohorts of patients with B cell malignancies or whether it has a broader potential in other cancers.5

It is essential to derive an in-depth and comprehensive understanding of CAR T-cell biology to maximize the efficacy of CAR T cells. Quantitative and systematic profiling of the mechanisms that potentiate the anti-tumor activity of CAR T cells at the cellular and molecular level can identify pathways to the generation of more potent infusion products.6 Not surprisingly, the comprehensive evaluation of the populations of CAR T cell infusion products used for the treatment of chronic lymphocytic leukemia has identified subpopulations that might be associated with clinical benefit.7 From the perspective of anti-tumor efficacy, it is now clear that T cells at least have to: (1) proliferate and persist in vivo,8-10 and (2) overcome and function under immunosuppressive conditions.11 Translating these findings has been accomplished through several novel approaches, including the identification of phenotypic markers of less differentiated cells (central memory (TCM) or stem memory (TSCM) T cells), genetically modifying the T cells to
co-express cytokines and the expression of dominant-negative receptors.\textsuperscript{12–16} These approaches have helped advance CAR T-cell therapy, but the number of different combinations that could be empirically tested is large, and thus not practically feasible. Even if restricted to only the studies of activating/inhibitory receptors, there are >15 individual receptors that are being targeted through either genetic modification or companion antibodies. Data-driven approaches can help engineer the next-generation of cellular immunotherapies with enhanced function under challenging tumor microenvironments.

At the single-cell level, although direct killing and serial killing of tumor cells mediated by CAR T cells is a crucial component of the anti-tumor function of CAR T cells both in vitro and in vivo, systematic studies of the molecules/pathways that can be altered for increasing killing potency have not been investigated.\textsuperscript{17–19} This is primarily due to the difficulty in tracking cytotoxicity mediated by individual CAR T cells while still being able to profile the same T cells at the molecular-level at single-cell resolution. Differential profiling of killers and non-killer T cells, following interaction with tumor cells, can not only further the understanding of CAR T-cell biology but can provide avenues to enhanced anti-tumor efficacy. Here, we directly integrated single-cell profiling of T cell function with transcriptional profiling to rapidly engineer modified CAR T cells with profound improvement in anti-tumor efficacy in liquid and solid tumor models in mice. Our results uncover an underappreciated facet of CD137 expression (cytotoxicity); \textit{CD2}, \textit{HAVCR2}, \textit{CD27}, \textit{TNFRSF9} (activation markers and costimulatory/inhibitory proteins); and \textit{TGFB1} (cytokine) were upregulated in killer T cells in comparison to non-killer T cells (figure 1 and online supplemental figure 2B). By contrast, \textit{GZMA} (cytotoxicity); \textit{LEF1}, \textit{RORA}, and \textit{TCF7} (transcription factors) and \textit{CD69} (activation marker) were upregulated in non-killer T cells in comparison to the other two populations (figure 1). The expression of transcripts corresponding to the CAR, \textit{GZMB}, and \textit{IFNG} were not different between killer and non-killer T cells (online supplemental figure 2). Principal component analysis (PCA) using the differentially expressed genes (DEGs) confirmed that the killers and non-killer T cells segregated into separate clusters, but this was donor-dependent (online supplemental figure 3). The set of DEGs in the killer and non-killer T cell comparisons that also showed at least 1.5-fold difference in each of the donor-derived populations studied included: \textit{FASLG}, \textit{HAVCR2} and \textit{TNFRSF9} with higher expression in killers, and \textit{RORA} and \textit{GZMA} associated with higher expression in non-killers (figure 1E). Collectively, these results demonstrated that a core set of genes can identify killer T cells, but there are donor-specific variations in the expression of these transcripts.

**RESULTS**

**Molecular profiling of individual serial-killer, mono-killer, and non-killer CAR T cells**

CD19-specific CAR (designated CD19R.28z) human T cells comprised predominantly of CD8\textsuperscript{+} T cells with a memory/naïve phenotype were generated (online supplemental figure 1A).\textsuperscript{20} CD19R.28z T cells as effectors and EL4-hCD19 cells as targets were sequentially loaded onto a nanowell grid array, and the kinetics of killing was monitored using timelapse imaging microscopy in nanowell grids (TIMING) (figure 1A). We considered all nanowell containing one CAR T cell and one to three tumor cells and identified serial killers (killed at least two tumor cells, online supplemental movie 1), mono-killers (killed exactly one tumor cell), and non-killers (T cells that conjugate to tumor cells but failed to kill, online supplemental movie 2). We retrieved a total of 256 single T cells (killers and non-killers), expressing two different CD19 CAR designs (modified IgG4 hinge or CD8a hinge, (online supplemental figure 1B), from three separate donor-derived CD19.28z T cell populations and performed direct quantitative PCR (qPCR)-based amplification since it has higher sensitivity than single-cell RNA-sequencing (RNA-seq).\textsuperscript{21} Comparative analyses of serial killers, mono-killers, and non-killer T cells showed no significant differences in the number of transcripts corresponding to housekeeping and T cell genes, including \textit{ALDOA}, \textit{GAPDH}, and \textit{CD3D} (online supplemental figure 2A).

Next, we performed differential testing between killers and non-killers across all donors by calculating the reproducibility optimized test statistic (ROTS) from the normalized transcript values for each gene.\textsuperscript{22} We identified that the transcripts corresponding to the \textit{FASLG} (cytotoxicity); \textit{CD2}, \textit{HAVCR2}, \textit{CD27}, \textit{TNFRSF9} (activation markers and costimulatory/inhibitory proteins); and \textit{TGFB1} (cytokine) were upregulated in killer T cells in comparison to non-killer T cells (figure 1 and online supplemental figure 2B). By contrast, \textit{GZMA} (cytotoxicity); \textit{LEF1}, \textit{RORA}, and \textit{TCF7} (transcription factors) and \textit{CD69} (activation marker) were upregulated in non-killer T cells in comparison to the other two populations (figure 1). The expression of transcripts corresponding to the CAR, \textit{GZMB}, and \textit{IFNG} were not different between killer and non-killer T cells (online supplemental figure 2). Principal component analysis (PCA) using the differentially expressed genes (DEGs) confirmed that the killers and non-killer T cells segregated into separate clusters, but this was donor-dependent (online supplemental figure 3). The set of DEGs in the killer and non-killer T cell comparisons that also showed at least 1.5-fold difference in each of the donor-derived populations studied included: \textit{FASLG}, \textit{HAVCR2} and \textit{TNFRSF9} with higher expression in killers, and \textit{RORA} and \textit{GZMA} associated with higher expression in non-killers (figure 1E). Collectively, these results demonstrated that a core set of genes can identify killer T cells, but there are donor-specific variations in the expression of these transcripts.

**Classification of killers and non-killers based on dynamic interaction data**

We next analyzed the time-dependent features available from TIMING that describes the interaction of the individual CD19R.28z T cells with the target cells (figure 2A, online supplemental table 1). Eleven functional parameters were used to cluster the cell populations using PCA, and they segregated into four separate clusters, three killer rich clusters (1–3) and one non-killer cluster (4) (figure 2B). The classification accuracy for clusters 1 (94% killers) and 4 (95% non-killers), without the aid of annexin V staining, confirmed the validity of using dynamic parameters to stratify T-cell killing propensity.

We utilized ROTS to identify four features that were consistently different across all the donors tested: T-cell polarization with and without conjugation to the tumor cell, and tumor cell polarization with and without conjugation to the T cell (figure 2A,C). Tracking of single-cells within each of these clusters confirmed that the simultaneous polarization of T cell (low aspect ratio), and predominantly circular morphology of tumor cells (high aspect ratio) during conjugation could be used to discriminate the killer and non-killer clusters (figure 2).
To identify the origin of four clusters within the TIMING data, when functional classifications only supported two kinds of T cells (killers and non-killers), we analyzed the killing behaviors of T cells with the aid of annexin V. Interestingly, these data supported a progressive increase in the efficiency of T-cell mediated killing from cluster
Figure 2  Classification of chimeric antigen receptor (CAR) T cells based on the dynamic functional data. (A) Schematic describing the dynamic interaction parameters used to quantify the interaction between individual CAR T cells and tumor cells. The red bar denotes the period of conjugation, and green denotes the induction of apoptosis. The time taken by the T cell to initiate conjugation ($t_{\text{Seek}}$), the duration of conjugation ($t_{\text{Contact}}$) and the time to induce apoptosis ($t_{\text{Death}}$) are computed for the T cell interacting with each tumor cell. The motility of the cells and their polarization (aspect ratio of minor/major axes) were also computed. (B) Unsupervised clustering of T cells based on their dynamic features measured by TIMING leads to the identification of killers and non-killers without the aid of the death marker. (C) The dynamic features that contribute the most variation in discriminating T-cell killing efficacy. (D) Single-cell tracks of the quantitative dynamic profiles comparing killer and non-killer T cells. (E) Qualitative and quantitative differences in the efficiency of the killer T cells identified by the clustering of dynamic features. Cluster 1 comprised the most efficient killers, whereas cluster 4 is dominated by non-killers (no $t_{\text{Death}}$). P value was determined using analysis of variance. (F) Transcripts that are differentially expressed between clusters 1 (efficient killers) and 4 (non-killers). The size of the circle reflects the median normalized gene expression at the single-cell level. Transcripts with boxes around them are the transcripts that are consistently identified as differentially expressed in all three donor-derived CAR T cell populations. All data is derived from the functional/transcriptional profiling of three donor-derived CAR T cell populations. Differentially expressed genes are identified using the ROTS.
4 to cluster 1. This was reflected both qualitatively in the time to tumor cell apoptosis and quantitatively in the number of tumor cells that were killed (figure 2E).

Collectively, these data support that as opposed to simple binary classifications, T cells with varying efficiencies of killing could be identified within heterogeneous populations by dynamic profiling.

Since the TIMING data and molecular profiles were mapped at single-cell resolution, we next sought to identify the comparative gene expression profiles of cells in cluster 1 (most efficient killer T cells) with those of cells in cluster 4 (non-killer T cells). Once again, the same set of housekeeping genes, *ALDOLA*, *GAPDH*, and *CD3D* were not different between these clusters (online supplemental figure 4A). The core set of DEGs conserved across all donor-derived cells included *TGFBI*, *CAR*, *CD58*, *HAVCR2*, and *TNFRSF9* enriched in cluster 1; and *IL2RA*, *IL2RB*, *LAG3*, and *GZMA* enriched in cluster 4 (figure 2F). Although the transcripts for *GZMB* were increased in cluster 1, this difference was not significant (online supplemental figure 4). Three genes were consistently different across all three donor-derived T cells, *HAVCR2* and *TNFRSF9* were upregulated in killer T cells; and *GZMA* upregulated in non-killer T cells (figure 2F). Unlike *GZMA*, since *HAVCR2* and *TNFRSF9* are surface receptors, we prioritized them for further investigation.

We investigated if tumor-cell intrinsic properties might be responsible for killing failures. Non-killer T cells could arise from defects in T cell cytotoxicity or could arise from the interaction of killer T cells with tumor cells that are resistant to killing. The relatively high frequency of failed killing events (42%) is inconsistent with tumor cell resistance since population-level assays routinely lead to complete elimination of all tumor cells even at E:T of 1:1 (data not shown). Furthermore, all non-killer T cells showed evidence of conjugation to tumor cells (figures 1A and 3) and we have previously shown that >99% of tumor cells express CD19 (antigen). To directly address if non-killer T cells can switch to killer T cells when encountering susceptible targets, we set up a TIMING assay with NALM-6 tumor cells and CD19R.28z T cells, followed by the microscopy of killer and non-killer T cells over 6 hours. Immediately after the assay, we added a new aliquot of NALM-6 tumor cells to investigate if the non-killers were capable of killing the added tumor cells (online supplemental figure 5A). Not surprisingly, non-killer T cells rarely transitioned to killer T cells suggesting that non-killer T cells are deficient in cytotoxicity, and this is independent of putative tumor cell resistance mechanisms (online supplemental figure 5).

**The inability to sustain granule polarization underlies CAR T cell failure to kill tumor cells**

Since we established that non-killer T cells are defective in cytotoxicity, we utilized kinetic modeling to understand the origin of the failed killing events. The interaction between T cells and the target cells can be conceptualized as a series of sequential intermediates: receptor ligation and activation, polarization of the granules towards the microtubule organizing center, degranulation, and target cell apoptosis (figure 3A). We hypothesized that the nature in which effector-target associations transition from conjugation to detachment occurs in a gamma-distributed fashion (online supplemental figure 6). This implies that effector-target conjugation times are a multi-step reaction consisting of $N_{\text{steps}}$ intermediate steps with identical rate constants $k_{\text{step}}$. Effector-target conjugation behaviors that possess the same kinetic profile are hypothesized to have a single defined rate-limiting step whereas inconsistent failure/stalling at any intermediate step will lead to an incongruent solution for the model.

For kill events (E:T 1:1 to 1:3), we find estimated values of $N_{\text{steps}} \geq 1$, suggesting that the kinetics for kill events is dominated by a rate-determining step (figure 3B). By contrast, no-kill events have $N_{\text{steps}} < 1$, which is physically impossible for the kinetic system described by the gamma distribution. Consistently across all of the E:T ratios, the non-killers are dominated by T cells with persistent contact with tumor cells indicative of stalling at a kinetic step before detachment. In aggregate, killing events appear to be kinetically homogeneous, whereas no-kill events seem to stratify into kinetically distinct subpopulations indicative of failure at one of the many sequential steps associated with killing.

To validate the implications of the model experimentally, and to understand the mechanistic basis of the failure, we performed TIMING assays for tracking lysosomes in live-cell microscopy. As expected, 100% of killer CD19R.28z T cells polarized the lytic granules to the immunological synapse and showed sustained polarization (figure 3C, online supplemental movie 3). This result is consistent with our previous data that demonstrated that granule exocytosis is essential for the killing mediated by CD19R.28z T cells.

By contrast, non-killer CD19R.28z T cells behavior could be classified into two categories: (1) 80% of the T cells showed only transient but not sustained polarization of the lysosomes towards the immunological synapse (figure 3D, online supplemental movie 4), and (2) 20% of the T cells showed sustained polarization similar to the killer T cells yet failed to kill (figure 3E, online supplemental movie 5). Collectively, the combined results from both modeling and experiments highlight that the failure of CD19R.28z T cells to kill can arise from failure at multiple steps in lysosome polarization and degranulation before killing (figure 3A).

**Inducible expression of TIM3 and CD137 proteins on killer T cells**

Both TIM3 (*HAVCR2*) and CD137 (*TNFRSF9*) are inhibitory/costimulatory receptors that are induced on T-cell activation. We incubated CD19R.28z T cells with NALM-6 tumor cells and evaluated the expression of the receptors TIM3, CD137, PD1 and CTLA4 at the single-cell level using flow cytometry. CD19R.28z T cells demonstrated increased expression of both TIM3 and CD137 but only
minimal increase in PD1 expression on incubation with the NALM-6 cells (online supplemental figure 7A,B).

To link the induction of expression of these proteins to underlying functional capacity, we incubated CD19R.28z T cells with NALM-6 cells and used the surrogate marker for degranulation, CD107a, to indirectly identify killer (CD69+CD107a+) and non-killer cells (CD69+CD107a-). The frequencies of T cells expressing both CD137 and TIM3 was higher within the degranulating subset in comparison to the CD107a- population. Non-killers T cells comprise a heterogeneous population that displays failure in distinct steps before killing. (A)

Figure 3  Non-killers T cells comprise a heterogeneous population that displays failure in distinct steps before killing. (A) Schematic depicting the set of major sequential steps leading to T-cell mediated killing of tumor cells. (B) These sequential steps are modeled as a gamma-distribution based on the contact lifetimes between chimeric antigen receptor (CAR) T cells and tumor cells at different E:T ratios. Non-killers have an apparent number of intermediate steps \( N_{\text{app}} < 1 \) suggesting the existence of kinetically distinct subpopulations indicative of failure at one of the many sequential steps associated with killing. (C–E) Representative examples of lysosome tracking experiments in CAR T cells in their interaction with tumor cells. Killer CAR T cells showed sustained lysosomal polarization (E) leading to killing, whereas the majority of non-killer CAR T cells (80%) failed to display sustained polarization (D) or failed to kill (20%) despite evidence of sustained polarization (E). A minimum of 30 events was analyzed for both killers and non-killer CAR T cells pooled from three different donor-derived populations. The cell boundaries are delineated with dotted lines. Time is denoted as hh:mm:ss. Scale bar is 25 µm.
Figure 4  Chimeric antigen receptor (CAR) T cells with inducible CD137 expression show transcriptional signatures of proliferative cells. (A) Representative flow-cytometry data demonstrating the induction of CD107a, on incubation with NALM-6 tumor cells. (B–C) Both TIM3 (HAVCR2), and CD137 (TNFRSF9) showed increased expression with the degranulating (CD107a, surrogate for cytotoxicity) T cells. (D) Differential induction of CD137 between the CD107apos and CD107aneg CAR T cells. RNA-sequencing data, illustrated as GSEA (gene set enrichment analyses)-derived pathways plotted in Cytoscape comparing CD137apos CAR T cells to CD137aneg CAR T cells after activation by NALM-6 tumor cells. Nodes (red and blue circles) represent pathways, and the edges (green lines) represent overlapping genes among pathways. The size of nodes represents the number of genes enriched within the pathway, and the thickness of edges represents the number of overlapping genes. The color of nodes was adjusted to an false discovery rate (FDR) q-value ranging from 0 to 0.01. Clusters of pathways are labeled as groups with a similar theme. (E) Enrichment of C2-curated pathways related to metabolism and proliferation comparing CD137apos CAR T cells to CD137aneg CAR T cells after activation by NALM-6 tumor cells. (F) Comparative assessments of the enrichment of mitochondrial transcripts (MitoCarta) within the CD137 and CD107a populations of CAR T cells. (G) Heatmap illustrating the pathways related to cellular metabolism in comparing the CD137 and CD107a populations of CAR T cells. The pathways that are significant (FDR q-value <0.05) are marked with an asterisk. For all the panels, a minimum of three donor-derived CAR T-cell populations was used.

T-cell subpopulation (figure 4A-B). Similarly, the degranulating T cells also showed a higher frequency of PD1+ but not CTLA4+ T cells (online supplemental figure 8A,B). Comparisons of the median fluorescence intensity showed that although TIM3 expression was induced on incubation with tumor cells, only small (but significant) differences in expression at the protein-level in comparing the killer and non-killer T cells were observed (online supplemental figure 8). By contrast, however, CD137 showed large differences in comparing the killer and non-killer T cell populations (figure 4B). In aggregate, these flow-cytometric data were used to validate the induction of CD137 and TIM3 expression within killer T cells.

**CD137pos populations demonstrate signatures of proliferation and are more glycolytic**

To facilitate a comprehensive and unbiased understanding of the differences in the functional capacities of cellular subpopulations, we took advantage of our results that demonstrated that CD137 is inducibly upregulated in killer T cells. Accordingly, we flow-sorted T cells after incubation with NALM-6 tumor cells using the standard marker for identifying killers (CD107a) and compared these to killer T cells that were defined by the inducible expression of CD137. Four separate populations of tumor-cell activated CD19R.28z T cells: CD137pos and CD137neg, and CD107apos and CD107aneg (online supplemental figure 9) were isolated and subjected to RNA-seq.

A total of 249 DEGs were identified (online supplemental table 2) in comparing the CD137pos and CD137neg populations, and pathway analyses performed using gene set enrichment analyses (GSEA) in comparing these two populations revealed the presence of four major clusters: an upregulation of pathways associated with cell cycle, DNA replication and carbohydrate/amino acid metabolism; and downregulation of pathways associated with oxidative phosphorylation and housekeeping genes in the CD137pos populations in comparison to the CD137neg populations (figure 4C-D). Similar comparisons of the CD107apos and CD107aneg populations yielded only 100 DEGs, and consistent with our single-cell studies, the expression of GZMA was increased within non-killers but not a statistically significant level (online supplemental table 3).

Since the data indicated enrichment of metabolic-related pathways within the CD137pos populations, we next wanted to systematically examine the mitochondrial transcriptome by utilizing the human MitoCarta database. Comparison of the donor-specific normalized gene expression counts within CD107apos and CD107aneg populations revealed a significant enrichment of mitochondrial transcripts within the CD107apos subset (figure 4E). Similarly, mitochondrial transcripts were enriched in the CD137pos populations in comparison to the CD107apos population (figure 4E).

GSEA analysis of metabolic pathways revealed that in addition to glycolysis and one-carbon metabolism, a number of the amino acid metabolism pathways were differentially induced in the CD137pos subset in comparison to CD137neg T cells, whereas oxidative phosphorylation was suppressed (online supplemental figure 10). This metabolic profile of the CD137pos population was different from the CD107apos population, suggesting that these markers identify cells with non-overlapping metabolic programs (figure 4F). Consistent with the upregulation of glycolysis, we confirmed that when CD19R.28z T cells were stimulated with NALM-6 cells, the CD137pos population showed a small increase in the uptake of the fluorescent glucose analog 2-NBDG in comparison to the CD137neg populations, and this small difference is not surprising since both populations have been stimulated with antigen (online supplemental figure 10). Collectively, our analysis of the mitochondrial transcriptome of CD137pos cells demonstrated that co-incubation with tumor cells leads to increased expression of the pathways of glycolysis and proliferation (figure 4E-F).

**Genetically engineered CD137L co-stimulation in CD19R.28z T cells leads to the expansion of younger cells and enables better control of leukemia in vivo**

Since the combined functional, transcriptional and phenotypic data suggested that CD137 is specifically upregulated within the killer-cell subpopulation, we next investigated if triggering the induced CD137 using its cognate ligand would facilitate improved T-cell function. Accordingly, we incubated CD19R.28z T cells with NALM-6 tumor cells either in the presence or absence of human CD137L trimers (hCD137L). Functional activation in the presence of hCD137L led to a decrease in expression of PD1 and TIM3, whereas CTLA4 was unchanged (online supplemental figure 11). As illustrated with mass cytometry, the decrease in CD69 expression on incubation with hCD137L might be indicative of increased cell proliferation. When we compared the killing capacity of the T cells in either population-level or single-cell killing assays, the presence of hCD137L only induced small changes in killing capacity in short-term assays (4-6 hours) (online supplemental figure 12 A-C). Since TNFRSF9 was specifically induced in the most efficient killers (figure 2F), we investigated if hCD137L stimulation can lead to sustained killing over the long term (36–48 hours). At high target ratios (E:T 1:10 and 1:20), the addition of CD137L enabled the sustained killing of CD19R.28z T cells over a 48 hours time period (online supplemental figure 12). Collectively, these results indicated that activation of the induced CD137 can support the sustained killing of functional CD19R.28z T cells.

Based on these results, we designed genetic constructs to facilitate the expression of CD137L in CD19R.28z T cells (figure 5A). Sleeping Beauty plasmids encoding for CD19 CAR (CD19R.28z) and CD137L were co-electroporated into peripheral blood mononuclear cells. CD19R.28z-CD137L T cells demonstrated a modest subpopulation of T cells co-expressing CD137L; at day 21, 7.6% of CAR T cells also expressed CD137L (figure 5B).
Figure 5  Genetically engineered CD137L costimulation improves the efficacy of CD19R.28z chimeric antigen receptor (CAR) T cells in liquid and solid tumors. (A) The constructs used for the transduction of the CAR and CD137L in T cells using electroporation and the SB system. (B) Expression of CD137L within SB modified CD19R.28z and CD19R.28z-CD137L T cells after expansion (21 days) using artificial antigen-presenting cell (aAPC) and interleukin (IL)-21. (C) Time course of the longitudinal measurements of NALM-6 derived photon flux from the three separate cohorts of mice treated with SB modified CAR T cells. The background luminescence was defined based on mice with no tumor. Error bars represent SEM and p values are computed using the Mann-Whitney U test. (D) Schematic depicting the construct used for CD137L expression using lentiviral transduction. (E) Expression of CD137L within lentivirally modified CD19R.28z and CD19R.28z-CD137L T cells after expansion (21 days) using aAPC and IL-21. (F) Time course of the longitudinal measurements of NALM-6 derived photon flux from the three separate cohorts of mice treated with lentivirally modified CAR T cells. The background luminescence was defined based on mice with no tumor. Error bars represent SEM and p values are computed using the Mann-Whitney U test. (G) Survival curves of mice inoculated with NALM-6 cells and treated with CD19R.28z or CD19R.28z-CD137L T cells. P values are computed using a log-rank test. (H) Schematic depicting the construct used for CD137L expression using retroviral transduction. (I) Expression of CD137L within the CD19R.28z and CD19R.28z-CD137L T cells after expansion (14 days) using IL-15 and IL-21. (J) False-colored images illustrating the photon flux from firefly luciferase (ffLuc) expressing SKOV3-CD19 cells. (K) Time course of the longitudinal measurements of SKOV3-CD19 derived photon flux from the three separate cohorts of mice. The background luminescence was defined based on mice with no tumor. Error bars represent SEM and p values are computed using the Mann-Whitney test. (L) Survival curves of mice inoculated with SKOV3-CD19-ffLuc cells and treated with CD19R.28z or CD19R.28z-CD137L T cells. P values are computed using a log-rank test.

Phenotypic analyses of the differentiation status indicated that the CD19R.28z-CD137L T cells had an increased frequency of naïve-like T cells and, in general, also an increased frequency of CD45RA+ T cells, in comparison to CD19R.28z T cells (online supplemental figure 13). NSG mice were injected with CD19+ NALM-6 human

leukemia cells transduced with firefly luciferase (ffLuc), and subsequently (5 days after tumor-cell engraftment) treated with a suboptimal dose of T cells that has very little efficacy in controlling the tumor genetically modified T cells (online supplemental figure 1A). At suboptimal T cell doses, CD19.28z-CD137L T cells showed significantly enhanced anti-tumor activity compared with CD19.28z T cells (figure 5C). These initial results served as proof-of-concept that even <10% CD19R.28z T cells co-expressing CD137L, lead to modest improvement in tumor control compared with CD19R.28z T cells. To improve the frequency of cells co-expressing CD137L and CAR, a lentiviral construct was designed with the CAR and CD137L separated by a T2A ribosomal skipping sequence (figure 5D,E). We next tested the efficacy of the CD19R.28z-CD137L T cells at suboptimal doses and confirmed that these cells showed better control of NALM-6 tumors and enhanced survival of treated mice as compared with the CD19R.28z T cells (figure 5F,G). In aggregate, these results demonstrate that genetically engineered CD19R.28z-CD137L T cells, regardless of the construct or the gene delivery method used, show better control of tumor in leukemia models, compared with CD19R.28z T cells.

**Genetically engineered CD137L costimulation in CD19R.28z T cells leads to the complete rejection of established solid tumor**

As a more challenging model, we evaluated the ability of CAR T cells expressing CD137L to control established solid tumor. We chose the SKOV3-CD19 ovarian cancer model since CD19R.28z T cells only have limited efficacy in this model, and all mice succumbed to the aggressive tumor.29 Retrovirally transduced CAR T cells were expanded for 10 days in the presence of exogenous interleukin (IL)-7 and IL-15 (figure 5H).30 Not surprisingly, CD19R.28z T cells had a small but significant reduction in tumor burden, but all the mice had tumor outgrowth (figure 5I,J). By comparison, CD19R.28z-CD137L T cells completely rejected these aggressive tumors, and all mice were tumor-free from day 21 and remained tumor free for 100 days (figure 5J). To compare the efficacy of this design to a CAR containing the 41BB (CD137) endodomain, we compared the CD19R.41BBz T cells against the CD19R.28z-CD137L using the SKOV3-CD19 tumors in vivo. Although the CD19R.4-1BBz showed better anti-tumor efficacy than CD19R.28z T cells, wherein 40% of the mice were tumor-free, CD19R.28z-CD137L T cells again demonstrated complete control of the tumors (online supplemental figure 15). Collectively, these results establish that CD19R.28z-CD137L cells are potent effectors in treating both CD19-expressing solid and liquid tumors.

**DISCUSSION**

The successful translation of CAR T cells as therapeutics has invigorated a basic question in CAR T cell (and TCR T cell) biology: why are only a subpopulation of T cells capable of killing and serial killing tumor cells? This is clinically relevant since T-cell mediated killing of tumor cells with exquisite specificity is one of the principal mechanisms of action for T-cell based drugs. Within the context of conventional CD8 T cells (TCR CD8 T cells) and CTL clones, it has been illustrated that individual T cells are highly heterogeneous in the ability to participate in killing and serial killing.31 Two-photon imaging has suggested that in vivo killing by CTLs is limited by MHC-I downregulation and might require cooperation between multiple CTLs.32 Directly in the context of CAR T cells targeting B-cell lymphomas, two-photon imaging illustrated that consistent with our previous and current in vitro single-cell results, individual CAR T cells can participate in serial and simultaneous killing eliminating multiple tumor cells by inducing direct apoptosis.17 18 33 The killing capacity of CAR T cells, however, is heterogeneous, and only ~20% of CAR T cells killed target cells on conjugation.34 35 This validation provided by the two-photon studies set the stage to understand the mechanisms of T cell failures by tracking individual cells in vivo.

Our analysis of CAR T cell killing using single-cell imaging, linked molecular profiling, and mathematical modeling revealed that the failure of T-cell killing is due to failures at multiple steps involved in killing cascade rather than differences in the abundance of the cytotoxic proteins. From a molecular standpoint, defects in the killing cascade can arise from: (1) heterogeneity in the expression of the Rab family of GTPases that are responsible for granule trafficking, and (2) expression of the master transcription factor NFATc1.36 37–39 Regardless of the molecular mechanism, as with any multistep reaction cascade, this, in turn, implies that killer T cells will only comprise a subpopulation of all CAR T cells. Improvements to T-cell therapies should thus focus on strategies to improve the longevity of the killer T cells.

Comprehensive quantification of the molecular differences between killer and non-killer T cells identified transcripts that have been previously documented to modulate T-cell and CAR T cell function. CD58 was identified as one of the essential genes for the effector T cell function in genome-wide CRISPR mutagenesis analysis.39 Interestingly, knockdown of CD58 in patient-derived HLA-A*02+renal cell carcinoma (2245R) cells, transduced with NY-ESO-1, showed at least two single-guide RNAs exhibited >50% resistance to NY-ESO T cell mediated lysis. Similarly, the role of FasL in promoting T cell expansion in vivo and in initiating extrinsic death receptor signaling within tumor cells is also known. Adoptively transferred FasL+CD8+ T cells demonstrated a twofold advantage in Ag-driven expansion over their FasL−counterparts.40 Studies with CD30.CAR T cells showed that while targeting CD30+EC tumor cells through the CAR (antigen-dependent targeting), also eliminated surrounding CD30–EC cells in an antigen-independent manner, via Fas/FasL interaction.41 Similarly, using a genome-wide loss of function screening, impairment of
death receptor signaling in acute lymphocytic leukemia resulted in rapid disease progression despite CAR19 treatment. The authors suggested that inherent dysregulation of death receptor signaling in ALL directly contributed to failure of CAR T cells due to T cell cytotoxicity impairment there by promoting progressive CAR T cell dysfunction.

We utilized the differential gene expression profiling between killer and non-killer T cells to identify pathways that can enhance the survival of killer T cells directly leading to sustained and enhanced anti-tumor efficacy. Our results illustrate that CD137 (41BB) is inducibly expressed on killer (and serial killer) CAR T cells and that ligation of the CD137 with its native ligand, CD137L, can enable the prolonged killing of tumor cells. While the inducible nature of CD137 in T cells and CAR T cells has been well studied and has been shown to promote increased mitochondrial biogenesis and enhanced interferon gamma secretion, our data provides the first direct link between cytotoxicity and CD137 activation at the single-cell level. CAR constructs incorporating a 41BB endodomain have been studied extensively and shown to promote persistence and delay exhaustion, which has also been translated clinically. As opposed to providing CD137 costimulation directly through the CAR (cis), we chose a trans design based on the expression of the CD137L. The significance of the different costimulatory domains containing CD28 and 41BB in cis, and the superiority of delivering 41BBL in trans has been elegantly demonstrated with the NALM6 tumor model. Our data-driven approach also advances these findings by showing prolonged efficacy in refractory ovarian cancer models. Broadly, the multi-dimensional profiling of functional and non-functional CAR T cells at the single-cell level is a powerful approach to improving the efficacy of genetically engineered T cells. Data-driven approaches based on modulating the longevity of killer T cells is an unbiased approach that is also amenable to personalization based on both the T cells and the tumor cells. This enhances the likelihood of success of the adoptively transferred cells as a broadly applicable cancer treatment.

MATERIAL/METHODS

Animal studies
All animal studies were conducted as per approved IACUC protocols at MD Anderson.

Human subjects statement
All work outlined in this report was performed according to protocols approved by the Institutional Review Boards at the University of Houston and MD Anderson.

Sleeping beauty CAR T cells
A second generation CAR containing a CD28 and CD3-ζ endodomains were expressed in healthy donor pan-T cells by electroporation with DNA plasmids from the Sleeping Beauty (SB) transposon/transposase system as described previously. The IgG4 hinge was modified to contain the Ser241Pro mutation to increase stability and decrease Fc receptor binding, as we have described previously. Cells were used between 14 and 28 days after transfection. Where indicated, CAR T cells were co-transduced with an additional transposon vector encoding for CD137L (4-1BBL).

Retroviral transduction of PBMCs
Activated T cells were transduced with retroviral supernatants by centrifugation onto 24-well retronectin-coated plates, in RPMI media supplemented with IL-7 (10 ng/mL) and IL-15 (5 ng/mL). The cells were allowed to expand for 10 days and the percentage of CAR, CD4, and CD8 T cells was verified by flow cytometry.

Lentiviral transduction of PBMCs
For activation, 1×10^6 PBMCs in 1 mL of RPMI media (10% FBS), supplemented with IL-2 (50 U/mL) and IL-21 (30 ng/mL) were seeded in non-treated 24-well plates (previously coated with 1 mg/mL each of anti-CD3 and anti-CD28 antibodies in PBS overnight). After activation for 48 hours, the activated T cells were transduced with lentiviral supernatants by centrifugation onto retronectin-coated plates. To selectively expand CAR T cells, artificial antigen-presenting cells K562 cells were used as described previously every 7 days. Media was subsequently changed and fresh IL-2, and IL-21 every 2 to 3 days. At the end of 3 weeks, the percentage of CAR, CD8, and CD4 T cells was recorded using a flow cytometer.

Transduction of SKOV3-CD19 cells with EGFP-fluc
SKOV3-CD19 cells were transduced with viral particles of HSV-EGFP-Fluc (Imanis Life Sciences) in presence of 8 µg/mL of polybrene (Sigma). The cells sorted as double-positive CD19 and EGFP expression, expanded and confirmed to be CD19 positive (>90% positive for EGFP and CD19).

Timing assays for the multiplexed study of effector cytolytic phenotypes
We previously described a method called TIMING that allows high throughput, timelapse, and single-cell level imaging of thousands of nanowells, each containing 1 to 4 cells. Effector (T cells) and target cells (mouse EL4 cells stably expressing human CD19), were labeled respectively with 1 µM PKH67 (PKH Green) and PKH26 (PKH Red) fluorescent dyes (Sigma-Aldrich) according to the manufacturer’s protocol and loaded on the array. Cell apoptosis was detected by immersing the array in phenol red-free cell-culture media containing a dilution of 1:60 annexin V - Alexa Fluor 647 (AF647) (Life Technologies). Arrays were imaged for 6 hours at an interval of 5 min using an Axio fluorescent microscope (Carl Zeiss) utilizing a 20×0.8NA objective, a scientific CMOS camera (Orca Flash 4.0), a humidity/CO₂ controlled chamber, and the tile function of the Zen software.
Image processing, cell segmentation, and tracking, and data analytics

Image analysis and cell segmentation/tracking were performed as described by us previously. Only the nanowells containing the T cells that were selected for single-cell reverse transcription (RT)-qPCR were processed for image processing and cell segmentation, offline and after the experiment, and spatio-temporal measurement of cell interactions were extracted.

Gene expression profiling using single-cell RT-qPCR

PKH green stained CAR T cells were imaged for 6 hours using TIMING as described above. After carefully washing the cells on the array three times with ice-cold PBS, cells were kept at 4°C until retrieval, which was completed within 2 hours. Time-lapse sequences were manually analyzed to rapidly establish two groups of live T cells according to their function, either cytotoxic or non-cytotoxic. The cells were individually collected using an automated micro-manipulating system (CellCelector, ALS) and deposited in nuclease-free microcentrifuge tubes containing 5 μL of 2×CellsDirect buffer and RNase Inhibitor (Invitrogen). Single-cell RT-qPCR was then performed using the protocol ADP41 developed by Fluidigm. Ninety-two cells spanning both cytotoxic and non-cytotoxic functions were assayed, along with bulk samples of 10 and 100 cells, and with no-cell and no-reverse transcription controls. The panel of 95 genes included genes relevant to T cell activation, signaling and gene regulation, and was designed and manufactured by Fluidigm D3 AssayDesign. For data analysis, we first extracted Log2Ex value by subtracting Ct values from a threshold of 30. We then excluded data from (1) cells that had less than 40% of genes that were amplified and had a mean of Log2Ex out of the range of population mean±3 SD and from (2) genes that were amplified in <10% of cells. The Ct values were quantile-normalized resulting in normalized gene expression values. Post-process analysis was done using Excel (Microsoft), Prism (GraphPad), and GeneMania webtool (http://www.genemania.org/).

Gene expression profiling of CD107a<sup>pos</sup> and CD137<sup>pos</sup> CAR T cells using RNA-seq

CAR T cells from three different healthy donors were co-cultured for 6 hours with PKH Red stained targets in the presence of anti-CD107a antibody (HA43, BD Biosciences). Cells were then fluorescently labeled with Live/Dead Aqua (Life Technologies) and antibodies specific for CD69 (FN50), CD4 (OKT4), and CD8 (RPA-T8) (BioLegend). Live, activated CD69<sup>+</sup> T cells, either with or without CD137 expression and with or without CD107a, were sorted on Aria I sorter (BD Biosciences) into four separate tubes containing 100 μL. RA1 and 2 μL TCEP reagents from the NucleoSpin RNA XS kit (Macherey-Nagel). RNA samples pairs were further extracted using the NucleoSpin RNA XS kit, followed by RNA cleanup XS (Macherey-Nagel). Indexed complementary DNA (cDNA) libraries were prepared using Ultra Low Input RNA kit and cDNA amplification of 16 cycles. cDNA was cleaned with Agencourt AMPure XP PCR Purification Kit (Beckman Coulter). Libraries were pooled and pair-end sequenced using Nextera XT DNA Library Preparation Kit (Illumina). RNA and DNA concentration was measured using Qubit RNA HS and dsDNA HS kits (Life Technologies) and RNA and cDNA quality were assessed using RNA 6000 Nano Kit and High Sensitivity DNA Kits (Agilent). Data quality was checked and pre-processed using Illumina’s BaseSpace FastQC Toolkit, sequences were aligned TopHat, and differentially expressed genes were obtained using and DESeq, and GSEA was applied. After sequencing, the first 15 bases were trimmed from the 3’ and 5’ sites in all reads to remove the adapter sequences and eliminate the non-uniform distribution caused either by the biased selection of reads or contamination of other sequences. The RNA-seq reads were aligned with the HISAT2 (2.0.5) program from Johns Hopkins University and mapped to the human reference genome (GRCf37/hg19). DEGs were identified using the DEseq2 tool that tests for differential gene expression based on a model using the negative binomial distribution. The differentially expressed pathways were then identified by the ‘GAGE’ package using the gene sets downloaded from the ‘Molecular Signature Database’ of the Broad Institute and also by using the GSEA software.

Plate-based T cell functional assays, phenotyping, cytokine secretion, and cytotoxicity assay

Plate based cytotoxicity assays were done by co-culturing unstained T cells and PKH Green stained target cells (NALM-6) at a 1:1 ratio for 6 hours. The experiments were run with CAR T cells from three donors in duplicate wells. Fluorescently labeled anti-CD107a antibodies (clone HA43) at 1.6 μL/well, together with GolgiStop (BD Biosciences) at 0.7 μL/mL was added to the co-culture to stain for degranulating cells. After the assay, cells were washed and stained with Live/Dead Aqua (Life Technologies), and the following antibodies from BD Biosciences: CD2 (RPA-2.10), PD1 (EH12.1), CD58 (1C3), CD244 (2-69), and the following from BioLegend: CTLA4 (L3D10), CD4 (OKT4), CD8 (RPA-T8), CD137 (4B4-1), CD69 (FN50), and TIM3 (344825, RnD Systems), and acquired by flow cytometry, where T cell activation, degranulation and ability to induce target cell death were the readouts. In some experiments, T cells were loaded on human CD137L coated flat-bottom 96-well plates and incubated for 6 hours at 37°C. The 96-well plate was precoated with 10 μg/mL of human recombinant CD137L protein (human Fc conjugated, carrier-free, RnD Systems) in 100 μL PBS per well overnight at 4°C, followed by two times washes with PBS before cell loading.

For phenotyping, CAR T cells were stained for 30 min at 4°C using a panel of human-specific antibodies from BD Biosciences: CD107a (HA43), CD62L (clone DREG-56), CD45RA (HI100), CD45RO (UCHL1), CD95 (DX2), CD3 (SK7), CD27 (L28, MT271), CD28 (L293), CD25 (M-A251), CD127 (HIL-7-R-M21), KLRG1 (2F1/KLRG1)
CD57 (NK-1), CD4 (OKT4), CD8 (RPA-T8), CD69 (FN50) CCR7 (G043H7), were from BioLegend. Cells were co-stained with the viability dye 7-AAD (BD), and with the in-house anti-CD19cFv.\textsuperscript{44} The optimization involves maximizing the reproducibility Z-score of overlapping genes ranked by d across a top list of size k via bootstrapping. All the calculations were done using the R package ROTS (Suomi et al, 2017) with the number of bootstrap and permutation resamplings set to B=50000 at a largest top list size considered of K=30. A false discovery rate of 0.05 was chosen as the cut-off.

Kinetic modeling
TIMING data collected from CD19R.28z CAR T cells were analyzed using R V.3.5.2 (R Core Team 2018). Analysis involved probability distribution fitting by maximum likelihood carried out using fitdistplus 1.0-11 in conjunction with several helper functions.\textsuperscript{35} Only data on the contact times of kill and no-kill events were considered, with ambiguous cases excluded (eg, simultaneous contact of an effector to two or more targets). For an effector with multiple targets, data arising these cases were sorted based on the chronological sequence the effector made contact with a target. Interval and right censoring of the data were accounted for in the fitting, where an uncertainty of 10 min was attributed to interval-censored data and contact time at the maximum experimental time of 350 min were right-censored. An additional 25 min was included for cases where the effector was already in contact with a target at the start of image acquisition, with the condition that the target was still observed to be viable.

In vivo testing of CAR T-cell efficacy
On day 0, 7-week-old NOD.Cg-PrkdcsidIl2rgtm1wjl/SzJ (NSG) mice were injected intravenously via a tail vein with 5×10^4 EGFP + flLuc + NALM-6 cells. Mice in the two treatment cohorts received via tail vein injection (on day 5) of 2×10^6 CAR T cells. One group of mice bearing tumor were not treated with T cells. Anesthetized mice underwent bioluminescent imaging in an anterior-posterior position using a Xenogen IVIS 100 series system (Caliper Life Sciences) 10 min after subcutaneous injection (at neck and shoulder) of 150 µL (200 µg/mouse) freshly thawed aqueous solution of d-Luciferin potassium salt (Caliper Life Sciences) as previously described.\textsuperscript{36} Photons emitted from NALM-6 xenografts were serially quantified using the Living Image 2.50.1 (Caliper Life Sciences) and contact times with multiple targets, data arising these cases were sorted based on the chronological sequence the effector made contact with a target. Interval and right censoring of the data were accounted for in the fitting, where an uncertainty of 10 min was attributed to interval-censored data and contact time at the maximum experimental time of 350 min were right-censored. An additional 25 min was included for cases where the effector was already in contact with a target at the start of image acquisition, with the condition that the target was still observed to be viable.

Statistical analysis
Asterix were denoted according to the standard p values; p<0.05*, p<0.01**, p<0.001*** and p<0.0001****. The statistical tests performed are described in the appropriate figure legends. For the single-cell gene expression analyses, statistical significance was tested by calculating the reproducibility optimized test statistic (ROTS), a modified t-type statistic of the form, d = \(\bar{x}_1 - \bar{x}_2\) / (\(\alpha_1 + \alpha_2\)) where \(\bar{x}_1\) and \(\bar{x}_2\) are group averages, \(\alpha\) is the pooled SE, and \(\alpha_1\) and \(\alpha_2\) are parameters optimized in a data-adaptive manner.\textsuperscript{37}


