



The role of single cell transcriptomics for efficacy and toxicity profiling of chimeric antigen receptor (CAR) T cell therapies

Moritz Thomas^{a,b,1} , Ruben Brabenec^{a,c,1} , Lisa Gregor^c , David Andreu-Sanz^c , Emanuele Carlini^c , Philipp Jie Müller^c , Adrian Gottschlich^{c,d} , Donjete Simnica^c , Sebastian Kobold^{c,e,f,2} , Carsten Marr^{a,*,2} 

^a Institute of AI for Health, Computational Health Center, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

^b School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

^c Division of Clinical Pharmacology, Department of Medicine IV, University Hospital, Ludwig-Maximilians-Universität München, Munich, Germany

^d Department of Medicine III, University Hospital, LMU Munich, Munich, Germany

^e German Cancer Consortium (DKTK), Partner Site Munich, a Partnership between the DKFZ Heidelberg and the University Hospital of the LMU, Germany

^f Einheit für Klinische Pharmakologie (EKLiP), Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

ARTICLE INFO

Keywords:

Single-cell sequencing
Chimeric antigen receptor
Adoptive T cell therapy
CAR T cell therapy
Target antigen
Off-tumor toxicity

ABSTRACT

CAR T cells are genetically modified T cells that target specific epitopes. CAR T cell therapy has proven effective in difficult-to-treat B cell cancers and is now expanding into hematology and solid tumors. To date, approved CAR therapies target only two specific epitopes on cancer cells. Identifying more suitable targets is challenged by the lack of truly cancer-specific structures and the potential for on-target off-tumor toxicity.

We analyzed gene expression of potential targets in single-cell data from cancer and healthy tissues. Because safety and efficacy can ultimately only be defined clinically, we selected approved and investigational targets for which clinical trial data are available. We generated atlases using >300,000 cells from 48 patients with follicular lymphoma, multiple myeloma, and B-cell acute lymphoblastic leukemia, and integrated over 3 million cells from 35 healthy tissues, harmonizing datasets from over 300 donors. To contextualize findings, we compared target expression patterns with outcome data from clinical trials, linking target profiles to efficacy and toxicity, and ranked 15 investigational targets based on their similarity to approved ones. Target expression did not significantly correlate with reported clinical toxicities in patients undergoing therapy. This may be attributed to the intricate interplay of patient-specific variables, the limited amount of metadata, and the complexity underlying toxicity. Nevertheless, our study serves as a resource for retrospective and prospective target evaluation to improve the safety and efficacy of CAR therapies.

1. Introduction

Chimeric antigen receptors (CARs) are fully synthetic receptors that endow immune cells, typically T cells, with targeting and lytic abilities towards desired target cells [1]. CAR-T cell therapy has emerged as a highly dynamic and promising approach in oncology. The field has witnessed rapid breakthroughs, including the first successful clinical trials in patients with refractory B-ALL in 2012, the landmark FDA approval of Kymriah in 2017 for pediatric and young adult patients with relapsed or refractory B-ALL, and the subsequent expansion of CAR-T

therapy to diffuse large B-cell lymphoma (DLBCL) and other B-cell malignancies [1–3]. Based on unprecedented response rates and their curative potential, CARs against CD19 and BCMA have been approved in various B cell malignancies and are thus part of the standard of care for these diseases. There have been efforts to move CARs into other hematological malignancies and, especially, solid tumors, with limited success so far [4,5].

However, despite these advancements, CAR-T therapy is also accompanied by significant controversies and challenges, including severe toxicities, limited success in solid tumors, and issues related to

* Corresponding author.

E-mail address: carsten.marr@helmholtz-munich.de (C. Marr).

¹ Contributed equally to this work.

² Shared senior authorship.

target selection and persistence. A major determinant of CAR therapy safety and efficacy is the target choice. The ideal target is ubiquitously expressed on cancer cells and not expressed at all in healthy tissues [6, 7]. In reality, these characteristics rarely apply and target expression is often shared with other cells and tissues. Such "on-target off-tumor" effects [6] can result in a range of sometimes severe and potentially life-threatening consequences, including cytokine release syndrome (CRS), neurotoxicity, hematotoxicity, and serious organ damage [6,8]. Recent studies have advanced our understanding of CAR-associated toxicities, including CRS and immune effector cell-associated neurotoxicity syndrome (ICANS), by providing frameworks for mitigation and predictive strategies [9]. Majzner & Mackall [10] emphasized difficulties in antigen selection and immune escape mechanisms, while Fesnak et al. [11] explored engineering variability, challenges in non-B-cell malignancies and solid tumors and its impact on therapeutic outcomes.

Furthermore, as highlighted by Ghaffari et al. [12], transcriptomic and epigenomic analyses offer crucial insights into prioritizing targets with high clinical potential while minimizing safety risks, particularly in solid tumors. Huang et al. [13] demonstrated the utility of scRNA-Seq in identifying transcriptional profiles associated with CAR T cell exhaustion and variability in the therapeutic response, providing a foundation for improved patient stratification and CAR engineering.

Anti-CD19 CAR T cells, for example, can result in CRS, neurotoxicity, and hematotoxicity [14,15], which has been partially linked to target co-expression and target availability [5,6]. Importantly, these are retrospective findings from the perspective of product development, which were discovered only after inception of said treatments in the clinic and turned out to be clinically manageable to a large extent [6, 16]. In other cases, however, CAR-induced toxicity might be less manageable or even fatal. Examples include CAIX-targeting CAR T cells used to treat renal cell carcinoma that caused hepatobiliary toxicity, most likely due to shared CAIX expression on the biliary epithelium [17, 18]. Fatal cardiorespiratory failure in a colorectal cancer patient upon treatment with CAR T cells targeting HER2 was potentially caused due to reactivity against pulmonary and cardiac tissue [4,19]. This underscores the need for a strategic evaluation of investigational targets, prioritizing those with clear clinical potential. Thus, a better a priori understanding of CAR target expression could help mitigate risk and prioritize approaches with higher potential for efficacy with controllable safety.

The widespread adoption of single-cell transcriptomics (scRNA-Seq) has revolutionized our understanding of cellular biology and provides unprecedented insights into gene expression [20–22]. Global endeavors such as the Human Cell Atlas (HCA) [23] and the Human BioMolecular Atlas Program (HuBMAP) [24] create comprehensive large-scale single-cell atlases that provide a unique window into tissue heterogeneity. Demonstrating the power of scRNA-Seq data to assess safety properties of CAR targets, Parker et al. [25] have linked neurotoxicity of CD19-targeting CAR T cells with CD19 expression in a small population of brain mural cells that maintain blood-brain-barrier permeability. Similarly, neurotoxicity in a patient suffering from multiple myeloma treated with BCMA-directed CAR T cells was associated with BCMA expression in the caudate nucleus of healthy human brain tissue [26].

These previous studies have predominantly focused on exploring a specific subset of cells expressing the target of interest [25,26]. However, with the rapid increase in clinical trials and the consequent expansion of CAR target sets, a comprehensive analysis of the global expression profiles of CAR targets in both healthy and malignant tissues becomes possible [27,28]. Zhang et al. [28] have provided insights into target antigen expression across tissues to elucidate toxicity mechanisms. Similarly, Jing et al. [29] examined normal tissue expression patterns to predict off-tumor toxicity risks. Building upon these efforts, our study integrates transcriptomic data from both healthy and malignant tissues with clinical trial outcomes, offering a broader framework for understanding CAR T cell efficacy and toxicity correlations.

To that end, we compile targets that are currently tested in clinical trials in (hemato-) oncology with FDA-approval for CAR T cell therapy. We then examine global gene expression profiles of CAR targets across respective tumor and healthy tissues, quantify their similarity to approved targets, and link them with clinical outcome data from patients who underwent CAR T cell therapy.

Our work aims to provide a comprehensive atlas as a resource for both clinical and translational research. While our study cannot establish direct mechanistic insights into CAR therapy limitations, it consolidates key data on target expression and toxicity profiles, enabling future work to refine target selection and mitigate risks. This underscores the essential role of single-cell transcriptomic analyses in driving future strategies in CAR T cell therapy.

2. Results

2.1. Compilation of approved and investigational CAR targets

To investigate whether the observed efficacy and toxicity of CARs could be explained by global target expression profiles, we first compiled a list of approved and investigational CAR targets across clinical trials (Fig. 1a). As of February 2023, there are a total of six FDA approved CAR T cell therapies, all utilizing CD19 or BCMA as targets to treat patients suffering from follicular lymphoma, multiple myeloma, B-cell acute lymphoblastic leukemia, large B cell lymphoma, diffuse large B cell lymphoma, mantle cell lymphoma, primary mediastinal B cell lymphoma, and high-grade B cell lymphoma (Table S1). By screening current clinical trials (www.clinicaltrials.gov) and literature, we accumulated a list of in total 15 investigational targets from clinical trials for the same malignancies (Fig. 1a–Table S2).

2.2. Screening of clinical trials for patient outcome data after CAR T cell therapy

With the accumulation of clinical trials, a body of CAR T therapy outcome data has begun to emerge. These data have the potential to provide a better understanding of the intricate relationship between molecular characteristics, treatment response and toxicity, and potentially identify key factors that influence therapeutic success. We therefore conducted a comprehensive screening of clinical trials involving CAR T cell therapy that reported patient outcome data (Table S3). For each study, we extracted the number of patients with evidence of progressive disease, stable disease, partial response, and complete response, which was used to calculate the overall response rate by dividing the number of patients with a partial or complete response by the total number of patients in the study. The data obtained presents a highly heterogeneous profile with respect to response rates and observed adverse effects across different targets and malignancies (Fig. 1b–e). CAR T cell therapies targeting CS1 resulted in lower overall response rates compared to other targets. Patients who received CD19-directed CAR T cells had relatively low levels of hematotoxicity (Fig. 1c), while showing the highest incidence of neurotoxicity (Fig. 1d). Hematotoxicity was most commonly observed in patients treated with CARs targeting CD22, CS1, and GPRC5D (Fig. 1c). Cytokine release syndrome (CRS) rates (grade 1/2) and severe CRS rates (grade 3/4) showed similar patterns across targets, with no discernible trend towards a given target (Fig. 1e).

2.3. Tumor expression profiles of CAR targets for the treatment of follicular lymphoma, multiple myeloma and B-cell acute lymphoblastic leukemia

The success of CAR T cell therapy is highly dependent on the precise recognition of tumor-specific targets, enabling the subsequent elimination of cancer cells while minimizing damage to healthy tissue. However, if the target is expressed at only very low levels, CAR T cells may

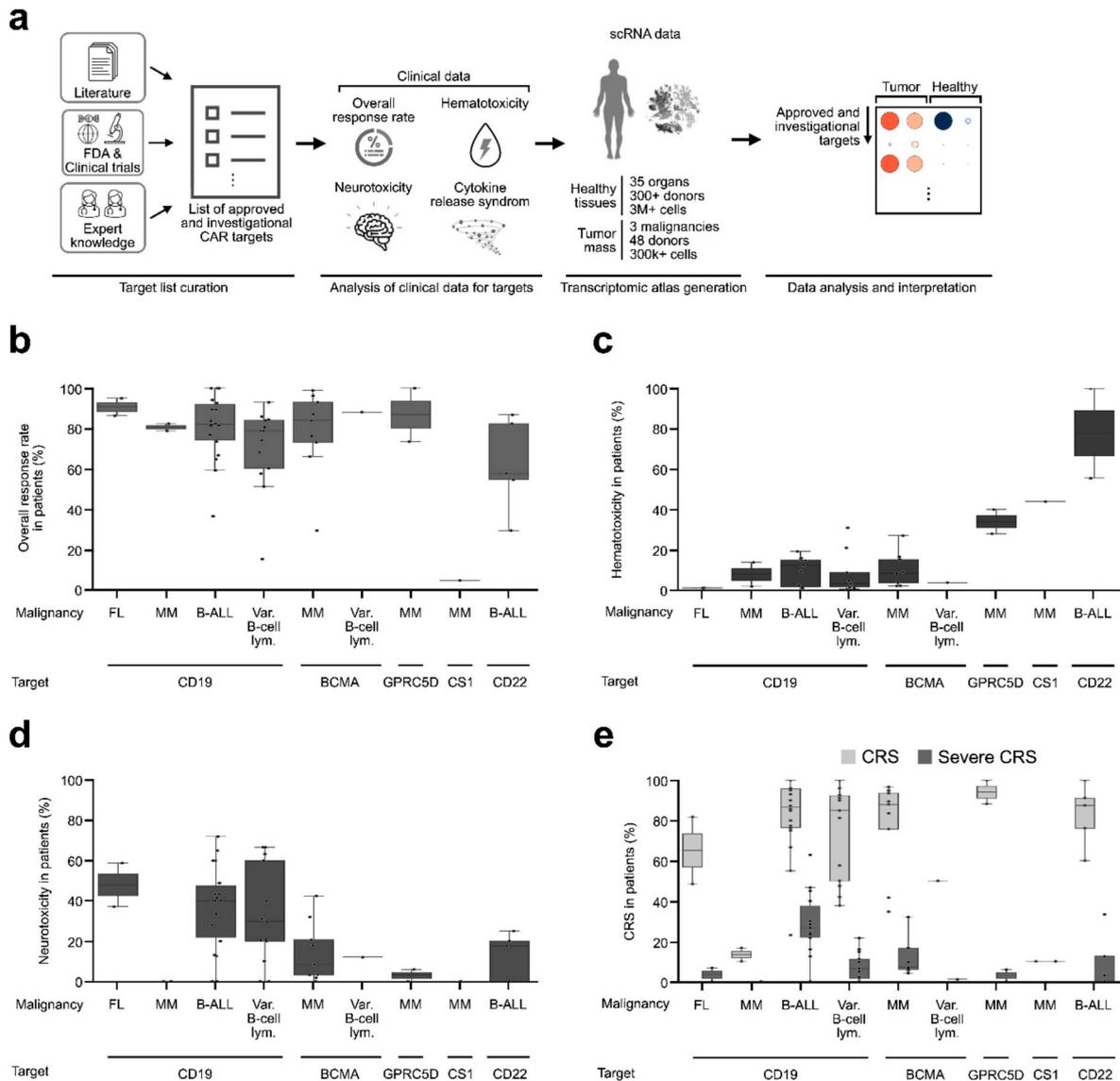


Fig. 1. Patient outcome data following CAR T cell therapies show high response rates, but also high toxicity rates.

a) A list of approved and investigational CAR targets was compiled from literature, FDA data and clinical trials, and expert knowledge. Clinical data provided insights into the efficacy (overall response rate) and safety (neurotoxicity, cytokine release syndrome (CRS), hematotoxicity) of CAR T cell therapy. Gene expression of CAR targets from clinical trials was evaluated computationally on atlases with over 3 million single cells from malignant and healthy tissues. **b**) High overall response rate (77 % on average), calculated by dividing the number of patients exhibiting partial or complete response by the total number of patients involved in the study following CAR T cell therapy, for five targets (approved targets CD19 and BCMA, investigational targets GPRC5D, CS1, and CD22) and nine treated malignancies. **c-e**) A considerable percentage of patients show signs of hematotoxicity (15 % on average) (**c**), neurotoxicity (26 %) (**d**), CRS (75 % defined as stages 1 and 2) and severe CRS (14 % defined as stages 3 and 4) (**e**). Each dot represents one clinical study, bars represent 25/75 quantile, whiskers extend to the furthest data point except for outliers.

fail to recognize and eliminate malignant cells. Achieving target selectivity is therefore critical to reducing side effects and improving the overall safety of the therapy. Therefore, it is essential to thoroughly evaluate expression levels and target density (i.e., the number of cells expressing a target above a threshold level) on malignant cells prior to administering CAR T cell therapy. To that end, we generated a large-scale single cell transcriptional atlas using three single-cell gene expression (scRNA-Seq) datasets with 303,190 cells from follicular lymphoma (FL), multiple myeloma (MM) and B-cell acute lymphoblastic leukemia (B-ALL) tumors [30–32] and analyzed expression levels of approved and investigational targets for the treatment of these malignancies (Fig. 2, Fig. S1, Table S4). For the other five B cell lymphomas mentioned above, we screened target expression across healthy cell populations, as there is currently no scRNA-Seq data available.

For follicular lymphoma, one approved product targets CD19

(Table S1), and two alternative targets are currently investigated in clinical trials: MS4A1 (from here on referred to as CD20) and CD22 (Table S2). Expression of these targets in 63,136 single cells [32] was confined to malignant B cell lymphoma cells, and non-malignant B cells and plasma cells, with differences in target density and mean expression (Fig. 2a, Fig. S1a–b). The two investigational targets were ordered according to their smallest Euclidean distance in gene expression space to the approved target CD19. The global expression pattern of CD22 was closest to CD19, as CD20 was also expressed on erythroid cell types (Fig. 2a, Fig. S1a–b).

For multiple myeloma, BCMA is currently the only target for two FDA approved CAR T cell products (Table S1), but we identified eight additional investigational targets in clinical trials (Table S2). CAR targets for the treatment of multiple myeloma were overall highly expressed on tumor cells and much less on healthy cells of the 212,400

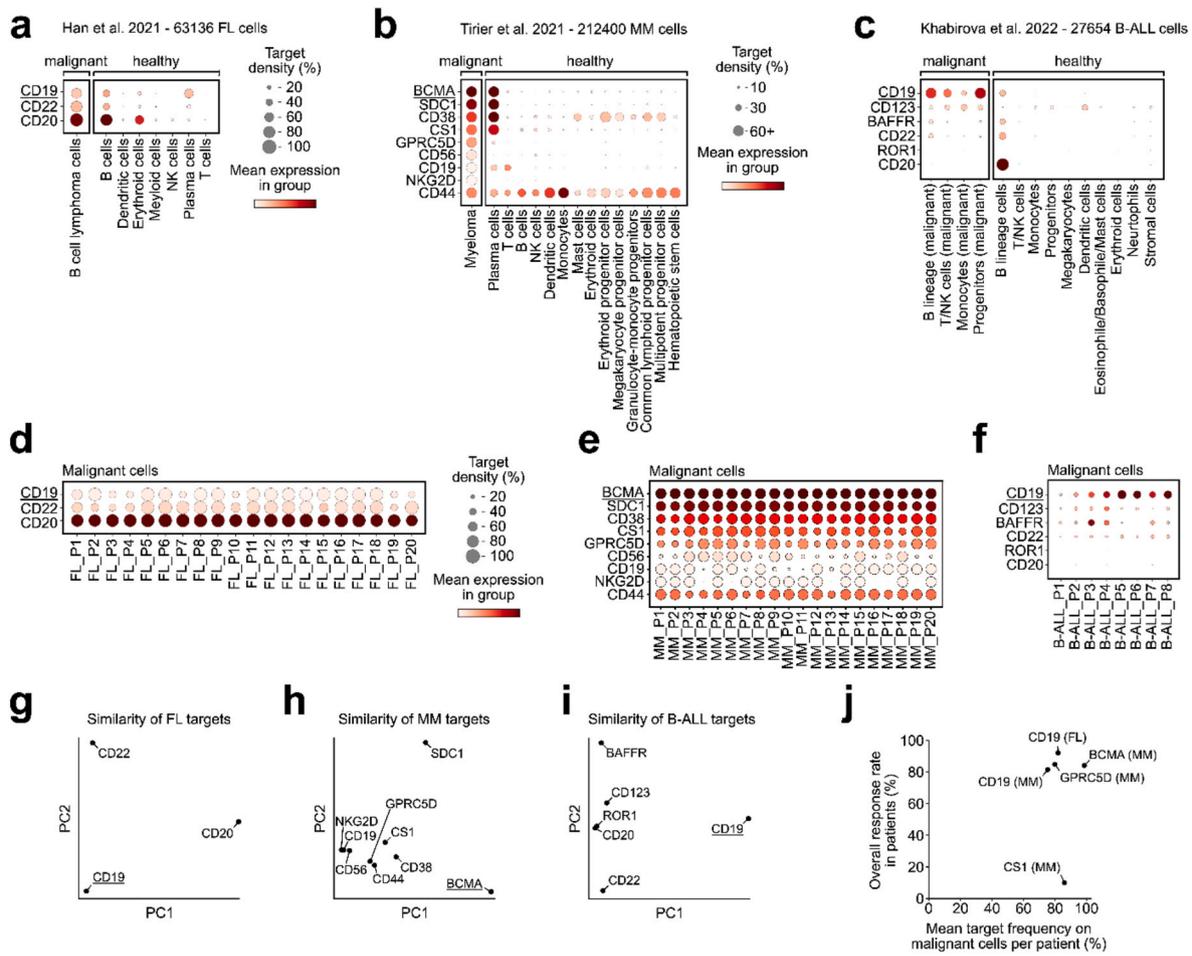


Fig. 2. CAR target expression on single malignant and healthy cells of patients suffering from follicular lymphoma (FL), multiple myeloma (MM) and acute lymphoblastic leukemia (B-ALL).

a,b,c) Variable expression of CAR targets on tumor and healthy cells for FL (**a**), MM (**b**) and B-ALL (**c**). Dot size indicates the fraction of cells per cell type expressing a target, color intensity shows mean normalized gene expression per cell type. FDA-approved targets (CD19 and BCMA) are underlined. Targets were ordered according to their similarity to approved targets by calculating the Euclidean distance in normalized gene expression space. **d,e,f)** Patient wise expression of CAR targets on healthy and malignant cells for FL (**d**), MM (**e**), and B-ALL (**f**). Due to poor sequencing depth, target densities in B-ALL data were generally low. **g,h,i)** Gene expression similarity on malignant cells for FL (**g**), MM (**h**) and B-ALL (**i**) targets in the first two principal components PC1 and PC2. **j)** Overall response rates in patients (calculated by dividing the number of patients exhibiting partial or complete response by the total number of patients involved in the study) seems uncorrelated to mean CAR target frequency of malignant cells per patient. Since target densities in B-ALL data were generally low, B-ALL targets were excluded.

single cell dataset of Tirier et al. [30] (Fig. 2b, Fig. S1c–d). BCMA, SDC1, and SLAMF7 (from here on referred to as CS1) were observed to be co-expressed on plasma or B cells, yet were also found to be highly expressed on tumor cells (Fig. 2b, Fig. S1c–d). KLRK1 (from here on referred to as NKG2D) showed very little overall expression across tumor and healthy cell populations. Investigational targets were ordered according to their smallest Euclidean distance in gene expression space to the approved target BCMA. In terms of global expression pattern, SDC1 was most similar to BCMA.

In the case of B-cell acute lymphoblastic leukemia, while CD19 remains the sole FDA-approved CAR T cell therapy target, we here explored five additional investigational targets to expand the therapeutic spectrum. (Table S1, Table S2). Investigational targets demonstrated only moderate expression levels within malignant cells (Fig. 2c, Fig. S1e–f). CD123 was closest to BCMA in terms of its global expression pattern (Fig. 2c, Fig. S1e–f), but exhibits some on-target off-tumor expression. CD19, IL3RA (from here on referred to as CD123), CD22, and BAFFR were expressed on both malignant and healthy cells belonging to the B cell lineage. Conversely, CD20 exhibited elevated expression levels within healthy cells of the B cell lineage but comparatively lower expression levels within malignant cells of the same

lineage. Investigational targets were ordered according to their smallest Euclidean distance in gene expression space to the approved target CD19. In terms of global expression pattern, CD123 was most similar to CD19.

Patient eligibility for CAR T cell therapy is often determined based on the expression of the target. We therefore evaluated target expression in individual patients suffering from follicular lymphoma, multiple myeloma and B-cell acute lymphoblastic leukemia (Fig. 2d–f). All targets employed in follicular lymphoma and multiple myeloma treatment consistently exhibited high expression levels in malignant cells across 20 follicular lymphoma patients (Fig. 2d) and 20 multiple myeloma patients (Fig. 2e). Notably in multiple myeloma, CD19, NKG2D, and CD56 lacked a consistently high target density across all patients (Fig. 2e). For targets related to B-cell acute lymphoblastic leukemia treatment, target densities were generally low (Fig. 2f). While CD19, CD123, BAFFR, and CD22 were expressed across all or most patients at least slightly, ROR1 and CD20 were hardly expressed at all (Fig. 2f).

Based on the expression of targets across malignant cells in follicular lymphoma, multiple myeloma, and B-cell acute lymphoblastic leukemia samples, we visualized distance and similarity between approved and investigational targets in PCA space. Target distances in PCA space were

comparable to target similarities according to Euclidean distance. After dimension reduction, the targets for follicular lymphoma (CD19, CD22, and CD22) appear similarly distant to each other (Fig. 2g). For multiple myeloma, seven investigational targets cluster, albeit without strong proximity to the approved BCMA (Fig. 2h). For B-ALL, CD19 appears distant from targets in trials (Fig. 2i). Upon investigating the interplay between CAR target expression profiles and patient outcome, we noted a uniform distribution of points across both axes with no clear correlation between the mean frequency of target gene expression on malignant cells per patient and the observed overall response rate (Fig. 2j).

2.4. CAR target expression across healthy cell populations

CAR T cells are designed to recognize and attack cells that express the target antigen upon infusion into patients. However, there is a risk of unintended damage if the target is also expressed on healthy cells. This may result in unwanted toxicity, especially if the target is located in vital organs or tissues [6]. To investigate the potential association between expression of CAR targets in healthy cell populations and observed adverse on-target off-tumor effects, we screened 32 scRNA-Seq datasets spanning a total of 3,629,817 single cells and 35 healthy tissues across the human body [33–63] (Fig. 3, Fig. S2a, Table S4). Targets were grouped according to availability of clinical data and ordered according to their smallest average Euclidean distance to approved targets in gene expression space across all cell populations and tissues (see Methods for details).

Expression of the approved targets CD19 and BCMA was high in B cells or cells of the B cell lineage, such as plasma cells (Fig. 3). For three more targets, clinical data is available: The investigational target GPRC5D exhibited the closest global expression pattern to CD19 and BCMA, although with lower overall expression levels (Fig. 3). In particular, GPRC5D showed overall low expression in brain cell populations, including astrocytes, microglial cells, and neurons (Fig. 3b). CD22 expression was mostly confined to B cells, but was also present in brain oligodendrocytes (Fig. 3b) and myeloid cells, such as basophils or mast cells (Fig. 3c and d). CS1 demonstrated a broader expression pattern across immune cells in multiple tissues, including T cells, NK cells, and myeloid cells in lung, lymph nodes, and spleen (Fig. 3c–e).

For targets without patient outcome data from clinical trials, CD20 was most similar to CD19 and BCMA and was expressed in cells of the B cell lineage, such as B cells and plasma cells, and in a narrow range of immune cells, particularly monocytes (Fig. 3). SDC1 expression was predominantly restricted to plasma cells, but could also be detected in specific cell populations of the liver (Fig. 3g) and kidney (Fig. 3h). CD56 was expressed in endothelial cells (Fig. 3a) and showed high expression in various brain cell types, including astrocytes, cerebellar epithelial cells, microglial cells, neurons, and oligodendrocytes (Fig. 3b). CD38 was observed across a broad spectrum of immune cell types, including lymphatic endothelial cells, innate lymphoid cells, T cells, and NK cells, but also in astrocytes, neurons, and pericytes (Fig. 3b–f). CD123 was predominantly expressed in monocytes, macrophages, dendritic cells, and mast cells but also exhibited expression in neurons (Fig. 3b). NKG2D expression was limited to a narrow range of immune cells, including T cells and NK cells (Fig. 3a). CD44 displayed a broad expression pattern across screened tissues (Fig. 3). ROR1 exhibited a narrow expression pattern across immune cells but was highly expressed in vital cell types of the brain, including astrocytes, neurons, oligodendrocytes, microglia, and was also expressed in cardiomyocytes of the heart (Fig. 3b–f). BAFFR expression was mainly confined to B cells and plasma cells (Fig. 3). CD137 showed low expression levels across immune cell types from multiple tissues, primarily lymphocytes and lymphatic endothelial cells (Fig. 3a–h). CD30 expression was confined to a narrow range of immune cells, but was also expressed in various brain cell types, such as astrocytes, microglial cells, and neurons (Fig. 3b). Based on Euclidean distance to approved targets in gene expression space across all screened cell populations and tissues (see Fig. S2a), CD79A showed the least

similar expression pattern to approved targets and was predominantly restricted to B cells and plasma cells (Fig. 3a–h).

We visualized distance and similarity between approved and investigational targets additionally in a two-dimensional PCA space (Fig. S2b). A substantial proportion of the targets exhibited notable proximity, indicating their similarity in expression patterns. The genes CD56, ROR1, CD38, CD44, and CD79A stood out by exhibiting considerable separation from the rest of the target set, signifying pronounced dissimilar expression patterns and a higher potential for undesired effects (Fig. S2b).

2.5. The impact of CAR target expression for efficacy and toxicity estimation

Naively, target gene expression in healthy tissue is expected to impact on-target off-tumor effects and therapy toxicity [25,26]. While some association between target expression on HSPCs and hematotoxicity (Fig. S2d) as well as target expression on immune cells in the lymph nodes and CRS (Fig. S2h, rightmost plot) seems plausible, our analysis did not reveal any significant correlation between global CAR target expression profiles and reported clinical toxicity effects (all Pearson correlations had Bonferroni-adjusted p-values above 0.05). When assessing hematotoxicity (Fig. S2c–f), neurotoxicity (Fig. S2g), and CRS (Fig. S2h), we observed an overall heterogeneous pattern between scRNA-Seq profiles and the occurrence of toxicity.

3. Discussion

CAR T cell therapy is a powerful approach for the treatment of relapsed or refractory B cell malignancies [64–66]. Despite clinical success, a wider application is hindered by considerable and often life-threatening adverse effects, such as on-target, off-tumor toxicities [8], which have been shown for a number of CARs [25,67–71], posing significant challenges to broad clinical application. Therefore, careful target selection and analysis of safety and potential risks is essential before initiating clinical testing.

We provide a comprehensive review of current targets from reported clinical trials in malignancies with FDA approval for CAR T cell therapy. Additionally, we screened clinical trials to generate patient outcome data following administered CAR T cell therapies. We interpreted observed toxicity effects in conjunction with global target expression profiles by generating a transcriptional landscape of more than 300,000 cells across follicular lymphoma, multiple myeloma, and B-cell acute lymphoblastic leukemia, as well over 3 million cells across 35 healthy tissues and more than 300 donors.

While our findings provide valuable insights into CAR T cell target expression, several effects warrant further investigation. Differences in response rates may stem from antigen heterogeneity, microenvironmental factors, and antigen density variability, while the limited correlation between target expression and toxicity likely reflects the influence of systemic immune activation and cytokine cascades. Additionally, high incidence of cytokine release syndrome (CRS) highlights the predominant role of global immune dynamics over target-specific expression.

Findings in tumor immunology underscores the importance of immune cell exhaustion, regulatory T cells (Tregs), and tumor microenvironment dynamics in shaping immune responses and toxicity outcomes, with recent studies suggesting that immune suppression driven by Tregs and tumor-associated macrophages (TAMs) can modulate CAR T cell efficacy and toxicity, potentially influencing the relationship between target gene expression and clinical outcomes [72,73]. Consistent with these observations, longitudinal data from CAR T cell therapies indicate that CD19-targeted therapies induce durable remissions with minimal long-term toxicity in certain subsets of patients, whereas BCMA-targeted CAR T cells show shorter-lived remissions. These insights highlight the need for refined target selection and therapy optimization by integrating

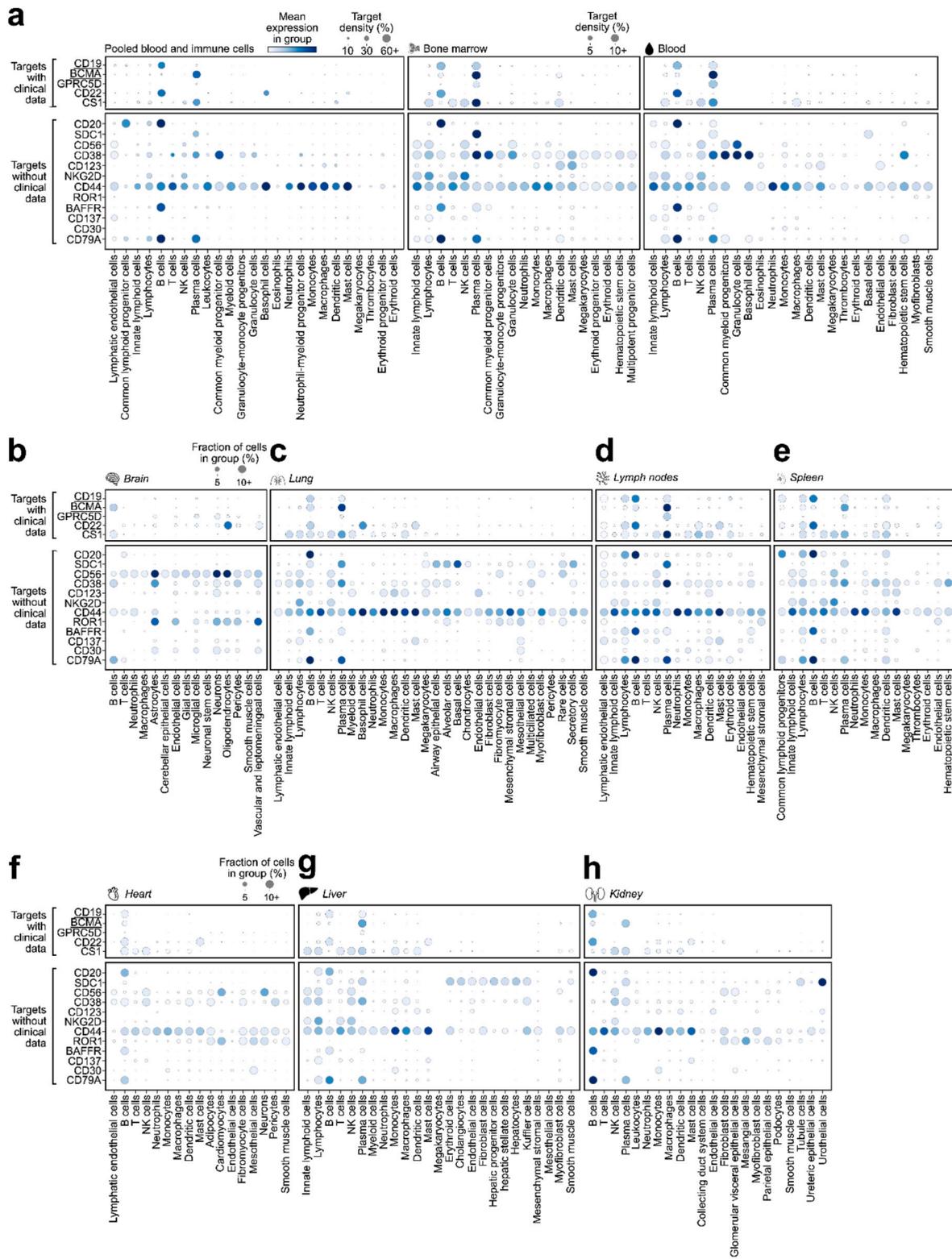


Fig. 3. Gene expression screening of approved and investigational CAR targets on a healthy transcriptional atlas reveals widespread on-target, off-tumor expression.

a Expression of CAR targets with (top) and without (bottom) clinical data on pooled immune cells from all healthy tissues (left), bone marrow (middle) and blood (right). Approved targets CD19 and BCMA are underlined. **b** Expression of CAR targets across brain (b), lung (c), lymph nodes (d), spleen (e), heart (f), liver (g) and kidney (h) tissues. Dot size indicates the fraction of cells per cell type expressing a target, color intensity shows mean normalized gene expression per cell type. Targets with and without clinical data were respectively ordered according to their smallest Euclidean distance in gene expression space to the approved targets.

single-cell transcriptomic data with detailed clinical metadata to better link profiles with outcomes.

We are not the first to utilize scRNA-Seq for assessing the potential of on-target off-tumor toxicities in CAR T cell therapy. However, previous approaches either did not take target expression levels into account [27], lacked gene expression analysis of tumor data [25–28], were limited to a few tissues [28], or focused entirely on a specific subset of cells expressing the target of interest [25,26]. Our study thus provides the first comprehensive global analysis of CAR target expression in tumor and healthy tissues and compares these results with clinical patient outcome data.

Unfortunately, scRNA-Seq data is not available for all malignancies with FDA approval for CAR T cell therapy, most likely due to distortions during the single cell dissociation step in the sampling process [74,75].

Our findings, including those reported in previous work [76], suggest that target expression patterns offer initial insights into patient selection and therapeutic potential. However, significant heterogeneity in publicly available sequencing datasets limits the ability to translate these findings into actionable recommendations for clinicians. Larger datasets with more diverse patient representation and integrated clinical metadata will be essential for deriving metrics that reliably inform patient stratification and therapy outcomes. As of now, the most value of our study comes from the ability to spot potential off-tumor areas of toxicity.

Our analysis might also be limited by the sequencing-induced 3' bias of the chosen public data that may lead to an incomplete characterization of isoforms [77,78]. For example, we were only able to detect the gene CD44, which shows a high expression pattern across a multitude of cell types, rather than specific CD44 isoforms, such as CD44v6, which is specifically implicated in tumor progression and metastasis [79–81]. To address this, complementary approaches, including full-length transcript sequencing (e.g., Smart-Seq3) and long-read sequencing technologies, are needed to capture isoform-specific expression patterns. Furthermore, proteomic validation of isoform expression on the protein level could provide additional translational insights. Also, while scRNA-Seq undoubtedly has great potential to benefit CAR T research [76], it is possible that gene expression results may not always correlate with protein expression. To overcome this limitation, large-scale protein expression screening data may be necessary, which we expect to become accessible in the coming years.

Notably, the full extent of potential risks associated with off-tumor activity of CAR T cells is still not well understood [82]. Therefore, our study does not aim to evaluate the relative merits or shortcomings of any given CAR target, it merely provides a way to identify potential off-tumor activity and guide the selection of CAR targets before translation into clinics.

The lack of a significant correlation between global scRNA-Seq profiles of targets and clinically reported toxicity is not entirely unexpected, considering the complex nature of toxicity effects and our limited understanding of their emergence. Hematotoxicity, neurotoxicity, and cytokine release syndrome associated with CAR T cell therapy are multifactorial in nature, influenced by various factors such as immune system activation, cytokine dysregulation, and interactions with the tumor microenvironment [6,16]. Our findings highlight that the presence of target mRNA expression alone is not a definitive indicator of toxicity, as evident from CSF1R, which was strongly expressed in microglia cells but did not result in observed toxicity upon CAR development [76]. This highlights the complexity of correlating gene expression with clinical outcomes and the necessity for multi-faceted analytical approaches. In a recent study [83], infections (non-target toxicity) were identified as a significant cause of non-relapse mortality post-CAR-T therapy, constituting 47.6 % of such deaths. The observed toxicities in CAR-T therapy therefore might not always be linked to the global expression of target genes in healthy cells, suggesting a more complex interplay. Further investigation into downstream immune responses is essential to fully understand these dynamics.

Although direct correlations between scRNA-Seq target expression patterns and clinical toxicity are complex, our comprehensive analysis provides a foundational understanding for further clinical research. While not providing a direct link between CAR target expression in healthy tissues and risk for clinical toxicity, it will serve as a valuable resource to assess potential off-tumor activity and guide the selection of CAR targets before translation into clinics.

Future clinical trials must focus on integrating single-cell transcriptomic profiling at key stages, e.g. during selection, therapy, and in follow-up, while ensuring the collection of comprehensive and standardized clinical metadata that is critical to assess the predictive value of transcriptomic data for therapy outcomes and toxicity. With the growing complexity of multidimensional datasets, systematic evaluation of integration methods is becoming increasingly important and highly warranted before clinical decision-making. In this study, we selected the scVI algorithm based on its benchmarking performance [84].

Looking ahead, continued accumulation of scRNA-Seq data from patients, along with increasing availability of clinical toxicity data, will play a crucial role in enhancing our understanding of toxicities associated with CAR-T cell therapy.

Early identification of potential safety concerns associated with CAR T cell therapy will facilitate a safer and more efficient treatment development. The careful assessment and monitoring of specific cell populations during therapy will help to ultimately gain a better understanding of potential risks and benefits of novel CAR T targets.

4. Methods

4.1. Collection of CAR target antigens

CAR target antigens with FDA approval were obtained from the website www.fda.gov. For antigens in clinical trials (“Early Phase 1, Phase 1, Phase 2, Phase 3, Phase 4, Not Applicable”) as of February 2023, we screened www.clinicalTrials.gov, applying the search criteria of the intervention/treatment: “CAR T Cell Therapy, CAR T, CAR” to various malignancies. See [Table S1](#) and [Table S2](#) for further details and an overview of these target antigens.

4.2. Screening clinical trials for patient outcome data

We obtained a list of ongoing clinical trials utilizing CAR T cell therapy as of March 2023 by thoroughly examining the websites <https://clinicaltrials.gov/> (search criteria: “car-t” or “chimeric antigen receptor”, filtering by completed and terminated trials) and <https://pubmed.ncbi.nlm.nih.gov/> (search criteria: “car-t” or “chimeric antigen receptor”, filtering by clinical trials). Papers of respective clinical studies were obtained from PubMed and were manually screened to extract the clinical trial identifier, which was then used to eliminate duplicates. As the focus of our study was to evaluate the efficacy of different CAR T cell products, we selected trials evaluating CAR T products and excluded any differing studies. All trials were screened by two independent individuals to avoid human error. Subsequently, these trials were carefully assessed to acquire detailed information of the CAR target and reported efficacy (number of patients exhibiting partial or complete response and patient death) and toxicity (number of patients displaying signs of hematotoxicity, neurotoxicity, or CRS and patient death) data of patients. To define hematotoxicity, we extended our analysis beyond the conventional Immune Effector Cell-Associated Hematotoxicity (ICAH) grading by incorporating additional hematological parameters such as anemia, lymphopenia, thrombocytopenia, infections, and leukopenia, which may provide a more nuanced understanding of the overall toxicity profile. The overall response rate was determined by dividing the number of patients exhibiting partial or complete response by the total number of patients involved in the study. For CRS, mild stages 1 and 2 were defined as “CRS,” while severe stages 3 and 4 were classified as “severe CRS”. The list of clinical trials with reported patient outcomes

is shown in [Table S3](#).

4.3. Single-cell transcriptome analysis

All preprocessing and analysis steps of scRNA-Seq data were run in python 3 using Scanpy [85] v.1.4.6 to 1.6.1 and anndata [86] v.0.7.1 to 0.7.5 except stated otherwise. All scRNA-Seq figures were plotted using matplotlib and seaborn.

4.3.1. Preprocessing publicly available scRNA-Seq data of healthy and malignant cells

We obtained raw, annotated count data for cells of healthy and malignant tissues using the python-based data repository sfaira [87] and cellxgene (<https://cellxgene.cziscience.com>). To ensure consistency and mitigate sequencing biases, 28 out of 35 data collections generated using 3' sequencing platforms (e.g., 10X Chromium) were included, thereby avoiding variability introduced by mixing full-length and 3' sequencing methods. An overview of used scRNA-Seq datasets can be found in [Table S4](#). Count data was converted to the anndata format, if necessary. To quantitatively analyze the expression of CAR target antigens across tissues, comparable preprocessing steps were carried out for each dataset separately. Barcodes were filtered for each sample to retain high-quality cells, defined based on the distribution of UMI counts, gene expression, and mitochondrial content, with specific thresholds determined after manual quality control (QC) and visual inspection.

Respective filtering thresholds were defined after visual inspection of each sample (for exact threshold values, see code provided on GitHub). Cells with more than 20 % of mitochondrial-encoded genes, indicating dying or stressed cells, were excluded after threshold validation through manual QC.

Genes that were detected in less than 20 cells per datasets were excluded for further analysis. UMI counts of each cell were normalized using the SCRAN algorithm as implemented in the R-based package [88, 89]. Briefly, size factors that correlate with the amount of counts of captured cells were estimated by preliminarily clustering the data using the Louvain algorithm implemented in Scanpy with a resolution of 0.5 before running ComputeSumFactors (min.mean = 0.1). The estimated size factors were then used for cell normalization. Finally, the data was log-transformed ($\log(\text{count}+1)$).

4.3.2. Feature selection and target antigen expression analysis

The top 4000 variable genes were identified based on normalized dispersion as described previously [77] using Scanpy's `pp.highly_variable_genes` with `flavor = cell_ranger`. Briefly, genes were ordered along their mean expression in several bins. For each bin, genes with the highest variance-to-mean ratio were selected as highly variable. To efficiently capture the underlying data structure in two dimensions, principal component analysis (PCA) dimension reduction was carried out by computing 15 principal components on highly variable genes using Scanpy's `pp.pca`. Next, a neighborhood graph was computed on the first 50 principal components using Scanpy's `pp.neighbors` with 15 neighbors. For 2D visualization, embedding the neighborhood graph via UMAP [90] was done by running Scanpy's `tl.umap` with an effective minimum distance between embedded points of 0.5. Cell annotation labels were provided by the authors of the respective study and were carefully inspected and relabeled if necessary to facilitate global comparisons of cell populations (see [Table S4](#) for details). To account for technical batch effects, such as sequencing depth or library preparation between datasets, we used scVI (Single-cell Variational Inference) [91] to integrate datasets of respective healthy tissues in an organ-wise fashion. Integration was not performed across distinct malignancies to preserve the integrity of disease-specific biological features. This approach ensures that biological variation specific to a condition remains intact, while minimizing potential artifacts that could arise from aligning datasets with inherently different biological contexts. The scVI method was chosen due to its demonstrated efficacy in benchmarking

studies, where it excelled in addressing complex integration challenges in single-cell datasets [84].

4.3.3. Reference mapping and label transfer of B-cell acute lymphoblastic leukemia data

To obtain reliable cell annotations, we mapped raw count data of unlabeled B-cell acute lymphoblastic leukemia cells [31] to an annotated reference of fetal bone marrow cells [92] using the semi-supervised variational auto-encoder model scANVI [93]. Briefly, the SCVI Model was trained on the raw, annotated reference data with 2 hidden layers and a dropout rate of 0.2 (for the exact parameters, see code provided on GitHub). Next, we initialized the scanVI model from the pretrained scVI model, before training the scANVI model for 20 epochs with 100 samples per label. Afterwards, we created a new query model instance before training the query data with a `weight_decay` of 0 for 100 epochs. The latent representation and the label predictions were obtained using `get_latent_representation()` and `predict()`, respectively. Finally, we computed a neighborhood graph with 15 neighbors using the scANVI representation, before embedding the graph using UMAP as mentioned before.

4.3.4. Target distance evaluation

For FL, MM, and B-ALL samples, Euclidean distances between targets were calculated by comparing expression values of each investigational target to expression values of the approved target across cells. For healthy tissues, Euclidean distances between targets were calculated by comparing expression values of each investigational target to the average expression values of both approved targets across cells. Euclidean distances were calculated using Scipy's `spatial.distance.cdist`. Resulting similarity values were min-max normalized.

Target similarities in PCA space for FL, MM, and B-ALL samples were calculated by first subsetting the data to malignant cells and investigational and approved CAR targets, before transposing the count matrix and calculating a PCA embedding using Scanpy's `sc.tl.pca` with default parameters. For target similarities in PCA space across healthy tissues, the count matrix was subsetted to include only investigational and approved CAR targets, before transposing the count matrix and calculating a PCA embedding using Scanpy's `sc.tl.pca` with default parameters.

CRediT authorship contribution statement

Moritz Thomas: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ruben Brabenec:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lisa Gregor:** Methodology, Data curation. **David Andreu-Sanz:** Formal analysis, Data curation. **Emanuele Carlini:** Formal analysis, Data curation. **Philipp Jie Müller:** Formal analysis, Data curation. **Adrian Gottschlich:** Formal analysis, Data curation. **Donjete Simnica:** Formal analysis, Data curation. **Sebastian Kobold:** Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Carsten Marr:** Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Data and code availability

As this study did not generate new data, publicly available scRNA-Seq can be found in the accession numbers in the respective studies. Python scripts for replicating the figures from the scRNA-Seq analysis are available as jupyter notebooks in a GitHub repository (https://github.com/marrlab/CAR_T_TargetExpression).

Ethics statement

This study involved data analysis and did not involve any experiments with human or animal subjects. All data used in this analysis were obtained from publicly available datasets, and no ethical approval was required for this research. The privacy rights of individuals are respected, and any data included in this study have been anonymized and de-identified in accordance with relevant privacy and data protection laws.

Declaration of competing interest

Parts of this work have been performed for the doctoral thesis of M.T. and R.B. at Technische Universität München and Ludwig-Maximilians-Universität München. S.K. has received honoraria from TCR2 Inc, Novartis, BMS, Miltenyi and GSK. S.K. is inventor of several patents in the field of immuno-oncology. S.K. received license fees from TCR2 Inc and Carina Biotech. A.G. received research support from Tabby Therapeutics for work unrelated to the manuscript. S.K. received research support from TCR2 Inc., Plectonic GmbH, Catalym GmbH and Arcus Bioscience for work unrelated to the manuscript. The remaining authors declare no competing interests.

Acknowledgements

M.T. is funded by the Volkswagen Foundation (project OntoTime). C. M. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 866411) and acknowledges support from the Hightech Agenda Bayern. This study was supported by the Förderprogramm für Forschung und Lehre der Medizinischen Fakultät der LMU (grant number 1138 to A.G.), the Bavarian Cancer Research Center (BZKF) (A.G. and TANGO to S.K.), the Deutsche Forschungsgemeinschaft (DFG, grant number GO 3823/1-1 to A.G.; grant number: KO5055-2-1 and KO5055/3-1 to S.K.), the international doctoral program 'i-Target: immunotargeting of cancer' (funded by the Elite Network of Bavaria; to S. K.), the Melanoma Research Alliance (grant number 409510 to S.K.), Marie Skłodowska-Curie Training Network for Optimizing Adoptive T Cell Therapy of Cancer (funded by the Horizon 2020 programme of the European Union; grant 955575 to S. K.), Else Kröner-Fresenius-Stiftung (to A.G. and IOLIN to S.K.), German Cancer Aid (AvantCAR.de to S.K.), the Wilhelm-Sander-Stiftung (to S. K.), Ernst Jung Stiftung (to S.K.), Institutional Strategy LMUexcellent of LMU Munich (within the framework of the German Excellence Initiative; to S.E. and S.K.), the BMBF Go-Bio-Initiative (to S.K.; grant number 16LW0600 to C.M.), the m4-Award of the Bavarian Ministry for Economical Affairs (to S.E. and S.K.), Bundesministerium für Bildung und Forschung (to S.E. and S.K.), European Research Council (Starting Grant 756017, CoG 101124203 and PoC Grant 101100460 to S.K.), Deutsche Forschungsgemeinschaft (DFG; KO5055-2-1 and 510821390 to S.K.), by the SFB-TRR 338/1 2021–452881907 (to S.K.), Fritz-Bender Foundation (to S.K.), Deutsche José Carreras Leukämie Stiftung (to S. K.), Hector Foundation (to S.K.), Bavarian Research Foundation (BAY-CELLATOR to S.K.), the Bruno and Helene Jöster Foundation (360° CAR to S.K. and C.M.), the Else-Kröner-Fresenius-Foundation (IOLIN to AG and SK).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compbio.2025.110332>.

References

- [1] C.H. June, M. Sadelain, Chimeric antigen receptor therapy, *N. Engl. J. Med.* 379 (2018) 64–73.
- [2] S.L. Maude, et al., Chimeric antigen receptor T cells for sustained remissions in leukemia, *N. Engl. J. Med.* 371 (2014) 1507–1517.
- [3] S.A. Grupp, et al., Chimeric antigen receptor-modified T cells for acute lymphoid leukemia, *N. Engl. J. Med.* 368 (2013) 1509–1518.
- [4] U. Patel, et al., CAR T cell therapy in solid tumors: a review of current clinical trials, *EJHaem* 3 (2022) 24–31.
- [5] C.L. Flugel, et al., Overcoming on-target, off-tumour toxicity of CAR T cell therapy for solid tumours, *Nat. Rev. Clin. Oncol.* 20 (2023) 49–62.
- [6] C.L. Bonifant, H.J. Jackson, R.J. Brentjens, K.J. Curran, Toxicity and management in CAR T-cell therapy, *Mol. Ther. Oncolytics* 3 (2016) 16011.
- [7] F. Perna, et al., Integrating proteomics and transcriptomics for systematic combinatorial chimeric antigen receptor therapy of AML, *Cancer Cell* 32 (2017) 506–519.e5.
- [8] R.C. Sterner, R.M. Sterner, CAR-T cell therapy: current limitations and potential strategies, *Blood Cancer J.* 11 (2021) 69.
- [9] A.W. Rankin, B.B. Duncan, C. Allen, S.K. Silbert, N.N. Shah, Evolving strategies for addressing CAR T-cell toxicities, *Cancer Metastasis Rev.* 44 (2024) 17.
- [10] R.G. Majzner, C.L. Mackall, Tumor antigen escape from CAR T-cell therapy, *Cancer Discov.* 8 (2018) 1219–1226.
- [11] A.D. Fesnak, C.H. June, B.L. Levine, Engineered T cells: the promise and challenges of cancer immunotherapy, *Nat. Rev. Cancer* 16 (2016) 566–581.
- [12] S. Ghaffari, M. Saleh, B. Akbari, F. Ramezani, H.R. Mirzaei, Applications of single-cell omics for chimeric antigen receptor T cell therapy, *Immunology* 171 (2024) 339–364.
- [13] S. Huang, et al., Deciphering and advancing CAR T-cell therapy with single-cell sequencing technologies, *Mol. Cancer* 22 (2023) 80.
- [14] A.V. Hirayama, C.J. Turtle, Toxicities of CD19 CAR-T cell immunotherapy, *Am. J. Hematol.* 94 (2019) S42–S49.
- [15] S. Fried, et al., Early and late hematologic toxicity following CD19 CAR-T cells, *Bone Marrow Transplant.* 54 (2019) 1643–1650.
- [16] J.N. Brudno, J.N. Kochenderfer, Recent advances in CAR T-cell toxicity: mechanisms, manifestations and management, *Blood Rev.* 34 (2019) 45–55.
- [17] C.H. Lamers, et al., Treatment of metastatic renal cell carcinoma with CAIX CAR-engineered T cells: clinical evaluation and management of on-target toxicity, *Mol. Ther.* 21 (2013) 904–912.
- [18] C.H.J. Lamers, Y. Klaver, J.W. Gratama, S. Sleijfer, R. Debets, Treatment of metastatic renal cell carcinoma (mRCC) with CAIX CAR-engineered T-cells—a completed study overview, *Biochem. Soc. Trans.* 44 (2016) 951–959.
- [19] M. Castellarin, et al., A rational mouse model to detect on-target, off-tumor CAR T cell toxicity, *JCI Insight* 5 (2020).
- [20] H. Chen, F. Ye, G. Guo, Revolutionizing immunology with single-cell RNA sequencing, *Cell. Mol. Immunol.* 16 (2019) 242–249.
- [21] A. Wagner, A. Regev, N. Yosef, Revealing the vectors of cellular identity with single-cell genomics, *Nat. Biotechnol.* 34 (2016) 1145–1160.
- [22] P. Guruprasad, Y.G. Lee, K.H. Kim, M. Ruella, The current landscape of single-cell transcriptomics for cancer immunotherapy, *J. Exp. Med.* 218 (2021).
- [23] J.E. Rood, A. Maartens, A. Hupalowska, S.A. Teichmann, A. Regev, Impact of the human cell atlas on medicine, *Nat. Med.* 28 (2022) 2486–2496.
- [24] HuBMAP Consortium, The human body at cellular resolution: the NIH Human Biomolecular Atlas Program, *Nature* 574 (2019) 187–192.
- [25] K.R. Parker, et al., Single-cell analyses identify brain mural cells expressing CD19 as potential off-tumor targets for CAR-T immunotherapies, *Cell* 183 (2020) 126–142.e17.
- [26] O. Van Oekelen, et al., Neurocognitive and hypokinetic movement disorder with features of parkinsonism after BCMA-targeting CAR-T cell therapy, *Nat. Med.* 27 (2021) 2099–2103.
- [27] Y. Jing, et al., Expression of chimeric antigen receptor therapy targets detected by single-cell sequencing of normal cells may contribute to off-tumor toxicity, *Cancer Cell* (2021), <https://doi.org/10.1016/j.ccell.2021.09.016>.
- [28] Y. Zhang, et al., Single-cell analysis of target antigens of CAR-T reveals a potential landscape of 'on-target, off-tumor toxicity', *Front. Immunol.* 12 (2021) 799206.
- [29] P. Mondello, S.M. Ansell, PHOENIX rises: genomic-based therapies for diffuse large B cell lymphoma, *Cancer Cell* 39 (2021) 1570–1572.
- [30] S.M. Tirier, et al., Subclone-specific microenvironmental impact and drug response in refractory multiple myeloma revealed by single-cell transcriptomics, *Nat. Commun.* 12 (2021) 6960.
- [31] E. Khabirova, et al., Single-cell transcriptomics reveals a distinct developmental state of KMT2A-rearranged infant B-cell acute lymphoblastic leukemia, *Nat. Med.* 28 (2022) 743–751.
- [32] G. Han, et al., Follicular lymphoma microenvironment characteristics associated with tumor cell mutations and MHC class II expression, *Blood Cancer Discov.* 3 (2022) 428–443.
- [33] X. Han, et al., Construction of a human cell landscape at single-cell level, *Nature* 581 (2020) 303–309.
- [34] et-al, The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans, *Science* 376 (2022) eabl4896.
- [35] C. Domínguez Conde, et al., Cross-tissue immune cell analysis reveals tissue-specific features in humans, *Science* 376 (2022) eabl5197.
- [36] P.A. Szabo, et al., Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease, *Nat. Commun.* 10 (2019) 4706.
- [37] N. Habib, et al., Massively parallel single-nucleus RNA-seq with DroNc-seq, *Nat. Methods* 14 (2017) 955–958.
- [38] J. Berg, et al., Human cortical expansion involves diversification and specialization of supragranular intratelencephalic-projecting neurons, *bioRxiv* (2020) 018820, <https://doi.org/10.1101/2020.03.31.018820>, 2020.03.31.
- [39] T.E. Bakken, et al., Comparative cellular analysis of motor cortex in human, marmoset and mouse, *Nature* 598 (2021) 111–119.

- [40] S.W. Lukowski, et al., A single-cell transcriptome atlas of the adult human retina, *EMBO J.* 38 (2019) e100811.
- [41] B.J. Stewart, et al., Spatiotemporal immune zonation of the human kidney, *Science* 365 (2019) 1461–1466.
- [42] Y. Muto, et al., Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney, *Nat. Commun.* 12 (2021) 2190.
- [43] J.-E. Park, et al., A cell atlas of human thymic development defines T cell repertoire formation, *Science* 367 (2020).
- [44] M. Xiang, et al., A single-cell transcriptional roadmap of the mouse and human lymph node lymphatic vasculature, *Front. Cardiovasc. Med.* 7 (2020) 52.
- [45] N. Kim, et al., Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma, *Nat. Commun.* 11 (2020) 2285.
- [46] P.A. Reyfman, et al., Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis, *Am. J. Respir. Crit. Care Med.* 199 (2019) 1517–1536.
- [47] K.J. Travaglini, et al., A molecular cell atlas of the human lung from single-cell RNA sequencing, *Nature* 587 (2020) 619–625.
- [48] E. Madisson, et al., scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation, *Genome Biol.* 21 (2019) 1.
- [49] K.R. James, et al., Distinct microbial and immune niches of the human colon, *Nat. Immunol.* 21 (2020) 343–353.
- [50] D. Fawcner-Corbett, et al., Spatiotemporal analysis of human intestinal development at single-cell resolution, *Cell* 184 (2021) 810–826, e23.
- [51] Y. Wang, et al., Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine, *J. Exp. Med.* 217 (2020).
- [52] S.A. MacParland, et al., Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations, *Nat. Commun.* 9 (2018) 4383.
- [53] T.S. Andrews, et al., Single-cell, single-nucleus, and spatial RNA sequencing of the human liver identifies cholangiocyte and mesenchymal heterogeneity, *Hepatology*. 6 (2022) 821–840.
- [54] P. Ramachandran, et al., Resolving the fibrotic niche of human liver cirrhosis at single-cell level, *Nature* 575 (2019) 512–518.
- [55] M. Baron, et al., A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure, *Cell Syst.* 3 (2016) 346–360.e4.
- [56] J. Peng, et al., Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma, *Cell Res.* 29 (2019) 725–738.
- [57] M. Enge, et al., Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns, *Cell* 171 (2017) 321–330.e14.
- [58] M.J. Muraro, et al., A single-cell transcriptome atlas of the human pancreas, *Cell Syst.* 3 (2016) 385–394.e3.
- [59] M. Litviňuková, et al., Cells of the adult human heart, *Nature* 588 (2020) 466–472.
- [60] P. van Galen, et al., Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity, *Cell* 176 (2019) 1265–1281.e24.
- [61] J.B. Cheng, et al., Transcriptional programming of normal and inflamed human epidermis at single-cell resolution, *Cell Rep.* 25 (2018) 871–883.
- [62] G.H. Henry, et al., A cellular anatomy of the normal adult human prostate and prostatic Urethra, *Cell Rep.* 25 (2018) 3530–3542.e5.
- [63] P. Bhat-Nakshatri, et al., A single-cell atlas of the healthy breast tissues reveals clinically relevant clusters of breast epithelial cells, *Cell Rep. Med.* 2 (2021) 100219.
- [64] S.L. Maude, et al., Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia, *N. Engl. J. Med.* 378 (2018) 439–448.
- [65] N. Raje, et al., Anti-BCMA CAR T-cell therapy bb2121 in relapsed or refractory multiple myeloma, *N. Engl. J. Med.* 380 (2019) 1726–1737.
- [66] S.J. Schuster, et al., Tisagenlecleucel in adult relapsed or refractory diffuse large B-cell lymphoma, *N. Engl. J. Med.* 380 (2019) 45–56.
- [67] C.H.J. Lamers, et al., Treatment of metastatic renal cell carcinoma with autologous T-lymphocytes genetically retargeted against carbonic anhydrase IX: first clinical experience, *J. Clin. Oncol.* 24 (2006) e20–e22.
- [68] E. Drent, et al., A rational Strategy for reducing on-target off-tumor effects of CD38-chimeric antigen receptors by affinity optimization, *Mol. Ther.* 25 (2017) 1946–1958.
- [69] K. Feng, et al., Phase I study of chimeric antigen receptor modified T cells in treating HER2-positive advanced biliary tract cancers and pancreatic cancers, *Protein Cell* 9 (2018) 838–847.
- [70] S. Gill, et al., Preclinical targeting of human acute myeloid leukemia and myeloablation using chimeric antigen receptor-modified T cells, *Blood* 123 (2014) 2343–2354.
- [71] A. Mardiros, et al., T cells expressing CD123-specific chimeric antigen receptors exhibit specific cytolytic effector functions and antitumor effects against human acute myeloid leukemia, *Blood* 122 (2013) 3138–3148.
- [72] S.F. Ngiow, A. Young, Re-education of the tumor microenvironment with targeted therapies and immunotherapies, *Front. Immunol.* 11 (2020) 1633.
- [73] K.M. Cappell, J.N. Kochenderfer, Long-term outcomes following CAR T cell therapy: what we know so far, *Nat. Rev. Clin. Oncol.* 1–13 (2023).
- [74] C.B. Steen, et al., The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma, *Cancer Cell* 39 (2021) 1422–1437, e10.
- [75] S.C. van den Brink, et al., Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations, *Nat. Methods* 14 (2017) 935–936.
- [76] A. Gottschlich, et al., Single-cell transcriptomic atlas-guided development of CAR-T cells for the treatment of acute myeloid leukemia, *Nat. Biotechnol.* (2023), <https://doi.org/10.1038/s41587-023-01684-0>.
- [77] G.X.Y. Zheng, et al., Massively parallel digital transcriptional profiling of single cells, *Nat. Commun.* 8 (2017) 14049.
- [78] L. Pan, H.Q. Dinh, Y. Pawitan, T.N. Vu, Isoform-level quantification for single-cell RNA sequencing, *Bioinformatics* 38 (2022) 1287–1294.
- [79] L.T. Senbanjo, M.A. Chelliah, CD44: a multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells, *Front. Cell Dev. Biol.* 5 (2017) 18.
- [80] I. Lodewijk, M. Dueñas, J.M. Paramio, C. Rubio, CD44v6, STn & O-GD2: promising tumor associated antigens paving the way for new targeted cancer therapies, *Front. Immunol.* 14 (2023) 1272681.
- [81] K.-H. Heider, H. Kuthan, G. Stehle, G. Muzert, CD44v6: a target for antibody-based cancer therapy, *Cancer Immunol. Immunother.* 53 (2004) 567–579.
- [82] R.G. Majzner, C.L. Mackall, Clinical lessons learned from the first leg of the CAR T cell journey, *Nat. Med.* 25 (2019) 1341–1355.
- [83] D.M. Cordas dos Santos, T. Tix, K. Rejeski, Infections drive non-relapse mortality following CAR-T therapy across disease entities and CAR products - a meta-analysis of clinical trials and real-world studies, *Blood* 142 (2023) 1064, 1064.
- [84] M.D. Luecken, et al., Benchmarking atlas-level data integration in single-cell genomics, *Nat. Methods* 19 (2022) 41–50.
- [85] F.A. Wolf, P. Angerer, F.J. Theis, SCANPY: large-scale single-cell gene expression data analysis, *Genome Biol.* 19 (2018) 15.
- [86] I. Virshup, S. Rybakov, F.J. Theis, P. Angerer, Alexander Wolf, F. anndata, Annotated data, *bioRxiv* (2021) 473007, <https://doi.org/10.1101/2021.12.16.473007>, 2021.12.16.
- [87] D.S. Fischer, et al., Sfaira accelerates data and model reuse in single cell genomics, *Genome Biol.* 22 (2021) 248.
- [88] A.T.L. Lun, K. Bach, J.C. Marioni, Pooling across cells to normalize single-cell RNA sequencing data with many zero counts, *Genome Biol.* 17 (2016) 75.
- [89] A.T.L. Lun, D.J. McCarthy, J.C. Marioni, A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor, *F1000Res* 5 (2016) 2122.
- [90] L. McInnes, J. Healy, J. Melville, UMAP: uniform manifold approximation and projection for dimension reduction, *arXiv [stat.ML]* (2018).
- [91] R. Lopez, J. Regier, M.B. Cole, M.I. Jordan, N. Yosef, Deep generative modeling for single-cell transcriptomics, *Nat. Methods* 15 (2018) 1053–1058.
- [92] L. Jardine, et al., Blood and immune development in human fetal bone marrow and Down syndrome, *Nature* 598 (2021) 327–331.
- [93] C. Xu, et al., Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models, *Mol. Syst. Biol.* 17 (2021) e9620.