Research Square

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

A Single Cell Transcriptomic Atlas of CD8+ Cytokine-Induced Killer Cells Generated from Metastatic Colorectal Cancer Patients

Kevin Fenix
kevin.fenix@adelaide.edu.au
Adelaide University https://orcid.org/0000-0003-1619-1406
Elham Shokati
Adelaide University
Celine Man Ying Li
Adelaide University
Runhao Li
Adelaide University
Cathal King
Adelaide University
Paul Drew
Adelaide University
Belinda Evans
Expert by Lived Experience
Christine Schutz
Expert by Lived Experience
Guy Maddern
Adelaide University
Michael Roach
Flinders University
Timothy Price
Adelaide University
Yoko Tomita
Adelaide University
Eric Smith
Adelaide University

Article

Keywords: Cytokine-Induced Killer Cells, Colorectal Cancer, Single-Cell RNA Sequencing, Single-cell TCR sequencing, NK-like T cells

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11560858/v1

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

Additional Declarations: (Not answered)

Abstract

Cytokine-induced killer (CIK) cell therapy is an adoptive cell immunotherapy under investigation for malignancies including metastatic colorectal cancer (mCRC). Clinical activity has been attributed primarily to $CD3^+CD56^+$ NK-like T cells, with $CD8^+$ cells considered a key effector subset. However, the transcriptional and clonal states underlying $CD8^+$ CIK-cell expansion and function remain poorly undefined. We combined single-cell RNA and paired T-cell receptor sequencing to profile 43,047 cells from day-14 CIK cell products generated from five patients with mCRC. Among 14,973 $CD8^+$ CIK cells, ten transcriptional states were identified, comprising six conventional and four NK-like states (C3, C5, C7 and C9). NK-like T-cell frequency increased during CIK expansion in all patients. The ten most expanded $CD8^+$ clonotypes preferentially occupied NK-like states in four patients, while memory-like states predominated in one, revealing inter-patient heterogeneity. Despite this, NK-like states showed significant clonal overlap (permutation test, Benjamini–Hochberg-adjusted $P < 0.05$), consistent with phenotypic plasticity, whereas memory-like states were largely clonally distinct. Trajectory analysis supported a branching model in which intermediate C5 was associated with trajectories toward highly cytotoxic Temra-like C9 or cytotoxic C3, with a further trajectory from C3 toward dysfunctional-like C7. Regulon analysis identified state-specific programs, including KLF2 in C5, GFI1 and IKZF2 in C9, IKZF1 in C7 and FOXM1 in proliferative C8. Together, these findings identify an apparent branch point within NK-like $CD8^+$ CIK cells that may shape cytotoxic and dysfunctional phenotypes. This suggests that targeting pathways governing this transition may improve CIK-cell manufacturing and lead to more effective therapeutic outcomes in mCRC.

INTRODUCTION

Colorectal cancer (CRC) remains a major cause of cancer-related death worldwide, with about 1.9 million new cases diagnosed annually and poor survival outcomes in metastatic disease, underscoring the urgent need for more effective therapies (1). Cytokine-induced killer (CIK) cell therapy has emerged as a promising immunotherapeutic approach with the potential to significantly improve CRC outcomes. Previous clinical trials have demonstrated that adjunct CIK cell therapy can significantly improve short term overall and progression-free survival in CRC patients, but the therapeutic efficacy of CIK cells is not durably sustained (2). A better understanding of the cell subsets and their developmental pathways in the CIK cell product may lead to improvements in the culture techniques with improved short- and long-term therapeutic outcomes.

CIK cells are a heterogeneous population of lymphoid cells expanded *ex vivo* from peripheral blood mononuclear cells (PBMC) (3). Currently they are described as comprising three subpopulations: conventional T cells (CD3 + CD56-), natural killer (NK)-like T cells (CD3 + CD56+, hereafter referred to as NK-like CIK cells), and NK cells (CD3- CD56+). CIK cells exhibit tumour cytotoxicity through both MHC-restricted and MHC-independent mechanisms, mediated by T cell receptors and NK receptors, respectively. MHC-independent tumour recognition is thought to be critical for CIK cell therapeutic efficacy and is primarily be driven by the NK-like CIK cell subpopulation (4, 5), which express NK receptors such as NKG2D and mediates tumour cell killing through interactions with its ligands, MICA/B and ULBPs commonly found on tumour cells(6). In addition, the expression of CD16 in some CIK cells enables antibody-dependent cell-cytotoxicity (ADCC), supporting enhanced tumour clearance when combined with monoclonal antibodies (7). Despite increasing confidence in the clinical application of CIK cells, their transcriptional heterogeneity and functional states remain poorly characterised.

Previous characterisation of CIK cells has relied on flow cytometric immunophenotyping and bulk transcriptomics to study the underlying biology (8, 9). These techniques identified NK-like CIK cells as critical mediators of therapeutic efficacy and subsequently their enrichment during culture is commonly used as a quality metric for clinical use. While these approaches have led to significant advancements in our understanding of CIK cell biology, they lack the resolution required to answer important questions required to advance this field further. These include the unbiased identification of discrete cell states within CIK cell cultures, the transcriptional dynamics that drive CIK cell subset differentiation, the clonal relationships between transcriptional states, and the functional contribution of these states to therapeutic effectiveness. Among CIK subsets, CD8 + T cells are considered the principal mediators of tumour cell killing (10, 11), yet the transcriptional heterogeneity and NK-like differentiation specifically within this compartment remain unknown. Here, we focused on CD8 + CIK cells to resolve their functional and clonal heterogeneity at single-cell resolution, with the aim of defining discrete transcriptional states, differentiation trajectories, and TCR-based clonal features that underpin cytotoxic and NK-like phenotypes.

Recent developments in single-cell RNA sequencing (scRNA-seq), integrated with single-cell TCR V(D)J sequencing (scTCR-seq) and improved computational tools have significantly advanced our understanding of T cell immunobiology by enabling high resolution cell annotation, identification of rare and transitional T cell states, and tracing of differentiation trajectories and clonal expansion We hypothesised that application of scRNA-seq and scTCR-seq on CIK cells would uncover novel and useful insights into CIK cell heterogeneity and their function.

In this study, we generated autologous CIK cells from treatment-naïve mCRC patients for single cell analysis. Day 14 CIK cell products were generated from PBMC isolated from five mCRC patients (12). Paired scRNA- and scTCR-seq on the CIK cells was undertaken using 10X Chromium NextGEM Single Cell 5' Reagent Kits v2 (10X Genomics). Profiling of transcriptional states of CD8 + CIK cells, clonal dynamics, and linking T cell receptor identity to functional phenotype at single-cell resolution were then performed. Together, our findings provide novel insight into the molecular characteristics that underlie the functional identity of therapeutic CIK cells.

RESULTS

scRNA-seq reveals the cellular heterogeneity of CIK cells.

scRNA-seq from pooled mCRC donor PBMCs.

We performed scRNA-seq on Day 14 CIK cell products generated from five treatment- naïve patients with metastatic CRC (mCRC) (Fig. 1a). CIK cell cultures were verified to contain NK-like CIK cell expansion by flow cytometry before sequencing **(Supp. Fig. S1**).Following quality control and filtering, we retained

43,047 cells across our pooled CIK samples for analysis. We identified 17 distinct “Global Cell Clusters” (G) present across all donors by unsupervised Louvain clustering, and then defined the transcriptional identity of each cluster by the top 100 differentially expressed genes (DEGs) (**Supp. Table S1**, **Supp. Fig. S2c)**. These cell clusters were then grouped into seven main cell subtypes based on shared expression patterns before matching to known cell types (Fig. 1b). Analysis of individual donors show similar proportions of these cell populations across all samples (Fig. 1c).

Identification and annotation of 7 major cell populations found in CIK cells

The seven main subtypes were identified as CD4 + T cells, CD8 + T cells, Mixed CD4/CD8 + T cells, NK cells, γ9δ2 T cells, B cells, and Cycling T Cells (Fig. 1b-c). As expected, the majority of our CIK cells were T cells as shown by robust *CD3E* expression in groups labelled CD4 + T cells, CD8 + T cells, Mixed CD4/CD8 + T cells, Cycling T cells, and γ9δ2 T cells (Fig. 1d). Within the T cell subtypes, *CD4* and *CD8A* expression was utilised to identify CD4 + and CD8 + T cells respectively. We were able to identify cell clusters containing only single-positive CD4 and CD8 T cells, however some contained a mixture of single positive CD4 and CD8 T cells and were labelled “Mixed CD4/CD8 T cells”. “Cycling T cells” similarly contained a mixture of CD4 and CD8 T cells and were identified by the robust expression of the proliferation marker *MKI67* (Fig. 1d**)** and an enrichment for a cell cycle gene signature (**Supp. Fig. S3a**) (13). In addition to αβ T cells, we also identified a population of γ9δ2 T cells through *TRDV2* and *TRGV9* expression (Fig. 1d). The γ9δ2 T cell clusters were enriched for cytotoxicity-associated genes (*GNLY, NKG7, FCGR3A)* and NK receptors *(NCAM1, KLRC1, KLRD1, NCR3)* (**Supp. Table S1**). Notably, the γ9δ2 T cells expressed *ZBTB16* which encodes PLZF, a critical transcription factor for effector γδ T cell function (**Supp. Table S1**) (14). Other than T cells, we also identified a small population of NK cells characterised by high expression of *NCAM1*(encoding CD56) and *TYROBP1* (Fig. 1d) and enriched for a gene signature for human NK cells (**Supp. Fig. S3b**) (15). Finally, we also identified some remaining B cells characterised by their expression of *CD19* and *MS4A1*(Fig. 1d). Taken together, our scRNA-seq analyses confirm previous flow cytometric based CIK cell studies, which showed that the cell product is predominantly composed of αβ T cells. As expected, cytotoxic CD8 + T cells are the main cells expanded in our CIK cell culture and due to their highly potent cell killing capacity should be critical for CIK cell mediated tumour clearance (16). To define cytotoxic CD8 + T cells composition in our CIK cell products, we performed sub-analyses of cytotoxic CD8 + T cells (15) .

scRNA-seq profiling of CD8 + T cells in CIK cultures

To assesses CD8 + T cell heterogeneity, we computationally isolated all CD8 single-positive αβ T cells from 43,047 cells of the pooled samples (n = 14,973). Unsupervised clustering of this CD8 + T cell population led to the identification of 10 distinct cell clusters (Fig. 2a, b). Top differentially expressed genes for each “Cytotoxic Cluster” (C) were identified (**Supp. Table S2**), followed by manual annotation to match these with known CD8 + T cell phenotypes (Fig. 2a-b). By initially dividing these cell clusters based on *NCAM1* expression (Fig. 2c**)**, we identified conventional (C0, C1, C2, C4, C6, C8) and NK-like (C3, C5, C7, C9) CD8 + subtypes. The proportional changes of CD8 + T clusters across parents shown in Supp. Fig. S2a.

Transcriptional profiling defines functional states in conventional CD8 + T cells.

C0 and C1 shared a memory-associated transcriptional core, including *TCF7, IL7R, LEF1, CCR7, LTB*, and *SELL*, together with intermediate expression of GZMK and GZMM and lack of terminal cytotoxicity genes such as *GZMB, NKG7*, and *PRF1* (Fig. 2a; **Supp. Fig. S4–S5**). C0 was further distinguished by enrichment of interferon-stimulated genes (ISGs) such as ISG20, indicating an ISG-responsive program within a memory-like transcriptional state (**Supp. Table S3**). Thus, rather than representing an independent lineage, C0 likely reflects an ISG+ state of this memory-like population (17). Pathway-enrichment analyses supported this interpretation: C0 was enriched for interferon-response signatures, whereas C1 showed hypoxia and glycolysis-associated programs, including *HILPDA, BNIP3, PFKFB4*, and *MIR210HG* (**Supp. Fig. S6**). Together, these findings suggest that C0 and C1 represent related but functionally divergent central memory (Tcm)-like states (ISG + Tcm-like and Tcm-like).

Interestingly, while most of Cytotoxic Clusters displayed elevated expression of canonical type 1 (Tc1) signature genes such as *TNF*, *TBX21* and *IFNG* (**Supp. Fig. S4)**, C2 had low Tc1 gene signatures and instead had high expression of *GATA3*, along with elevated expression of *PTGDR2, IL13* and *IL4*, consistent with type 2 cytotoxic T cells (Tc2) signature and was therefore annotated as a “Tc2-like” subtype (17).

Cell clusters C4, C6, and C8 showed enrichment of cycling cells (Fig. 2a). Cell cycle phase distribution showed that C4 was predominantly composed of G2/M phase, whereas C6 was enriched for S phase, and C8 showed a broader distribution across S and G2/M phases **(Supp. Fig. S7a).** To determine whether these clusters exhibited transcriptional identities beyond their cycling state, differential expression analysis was performed on C4, C6, and C8, followed by exclusion of cell-cycle-associated genes (**Supp. Table S4, Supp. Fig. S7b**). C4 was characterised predominantly by *MKI67* expression with no additional transcriptional identity markers and was therefore labelled “Cycling”. In contrast, C6 expressed memory-associated genes (*AQP3, CD27, LEF1, LTB, IL7R*) and was labelled “Cycling, Memory-like”. C8 co-expressed cytotoxic effector genes (*GZMB, PRF1, GNLY*) and exhaustion-associated checkpoint genes (*ENTPD1, LAG3, CTLA4*), consistent with an effector-to-exhaustion phenotype among actively cycling cells and hence was labelled “Cycling, Teff/Tex-like”.

Identification of NK-like CD8 + T cell states.

NK-like CIK cells are reported to be critical for CIK cell therapeutic efficacy (5). NK-like CD8 + T cells, identified by expression of NCAM1 and NK-associated activating and inhibitory receptors, were primarily localised in C3, C5, C7, and C9 (Fig. 2d). NCAM1 expression was highest in C9 (63.6% of cells), with intermediate expression in C3, C5, and C7 (10.6–12.2%), compared with background levels in the remaining clusters (C0-2; 2.2–6.1%) (**Supp. Table S5**). C3, C5, and C9 resembled effector memory (Tem) states, showing high enrichment of effector-related genes (*GZMB, GZMA, PRF1, NKG7, GNLY, CST7, FASLG, IFNG, GZMK, GZMH)* (**Supp. Fig. S8**).

C3 and C5 shared cytotoxicity-associated transcriptional features, with no significant difference in patient-level cytotoxicity scores between the two subtypes (paired Wilcoxon p = 0.281; adjusted p = 0.281; **Supp Table S6**). C3 showed intermediate expression of *CD226, KLRK1, KLRD1*, and *BHLHE40* (Fig. 2d; **Supp.**

Fig. S8). *BHLHE40* sustains CD8 + T cell effector and exhausted phenotypes (18). In support, C3 displayed elevated expression of exhaustion-associated genes (*LAG3, and CTLA4*) with a consistent trend toward greater exhaustion than C5 across all 5 patients, although the difference did not reach statistical significance (paired Wilcoxon p = 0.059; adjusted p = 0.118; **Supp Table S6**). We therefore annotated C3 as a "Teff NK-like" subtype (**Supp. Fig. S8**). In contrast, C5 showed comparable cytotoxicity marker expression but lower expression of exhaustion markers and more pronounced expression of adhesion-associated genes, including *ITGB7, ITGAL, ITGA2*, and *ITGB2* (**Supp. Fig. S8**). Together, these features were consistent with a less exhausted, intermediate Tem phenotype, therefore C5 was annotated as an intermediate Tem NK-like subtype.

C9 had elevated expression of *CX3CR1 and FCGR3A*, (**Supp. Fig S7**), suggesting a terminally differentiated effector memory (Temra)-like state (19). Interestingly, C9 expressed the highest cytotoxic-associated genes, *NK*-activating receptors (*NCR3, KLRK1, KLRC2*), NK-inhibitory receptors (*KLRD1, KLRC3, KIR2DL1, KIR3DL1*, and NCAM1 (Fig. 2d). Thus, C9 was annotated as a "KIR+ Temra NK-like" subtype (Fig. 2a).

C7 showed downregulated expression of effector genes as well as high expression for multiple exhaustion markers, including *HAVCR2, ENTPD1, CTLA4* and *TOX* (Supp. Fig. S7). *HAVCR2, ENTPD1*, and *TOX* are associated with driving cells toward a terminally exhausted (Tex)-like state (20). Therefore, C7 was annotated as Tex NK-like (Fig. 2a).

To further validate these identity assignments, we scored CD8 + cell subtypes using NK-like, Tex, and Temra gene module scores derived from gene sets reported in published single-cell studies. Consistent with our marker-based annotation, C7 showed the highest mean Tex module score among the ten Cytotoxic Clusters, with significant elevation relative to C0, C2, C4, C6 and C8 (Tukey-adjusted $P < 0.05$), although it did not differ significantly from C1, C3, C5 or C9 (**Supp. Fig. S9a**). C9 exhibited a significantly higher Temra module score than all other subtypes (Tukey-adjusted $P < 0.05$), consistent with C9 representing the most Temra-like cytotoxic state (**Supp. Fig. S9b**). Among the NK-like subtypes, C9 also showed the highest NK-like module score, significantly exceeding all other clusters including C3, C5 and C7 (Tukey-adjusted $P < 0.05$), while C3 and C5 scored significantly higher than most but were not significantly different from each other (**Supp. Fig. S9c**). C7 showed an intermediate NK-like score, significantly higher than C2, C4, and C6, but significantly lower than C3, C5, and C9 (**Supp. Fig. S9c**). Thus, our module-score analysis further supported the manual annotations and marker-based assignment of both conventional and NK-like CD8 + T-cell subtypes

Taken together, we provide for the first time a comprehensive annotation of CD8 + T cell subtypes including comprehensive characterisation of previously unknown transcriptional states within the NK-like CD8 + T cell subtypes in our CIK cell products.

scTCR analysis reveals clonal connectivity among CD8 + CIK cell states

We next investigated the clonal relationships between these transcriptional states using TCR sequences as genetic barcodes to track shared differentiation trajectories in our CIK cell cultures. In total, 34,525 TCR clonotypes were successfully mapped across the CIK cell subtypes, including 12,892 clonotypes detected in CD8 + CIK cells **(Supp. Tables S7-8)**. Hyperexpanded clonotypes (> 20 cells) were predominantly observed within the NK-like subtypes C3, C5, C7 and C9 **(**Fig. 3a, b**)**. Differential expression analysis comparing expanded (> 10 cells) and non-expanded (< 9 cells) clonotypes further showed increased expression of cytotoxicity genes, including *GZMB, PRF1, NKG7* and *IFNG*, together with exhaustion genes including *LAG3, TOX, ENTPD1* and *HAVCR2* **(**Fig. 3c**).** These results are consistent with clonal expansion of CD8 + NK-like cytotoxic subtypes upon activation. consistent with enrichment of cytotoxic and exhaustion-associated programs among expanded CD8 + CIK clonotypes (21).

We next quantified pairwise clonal overlap between transcriptional states to determine the extent to which distinct CIK subtypes shared TCR clonotypes. The greatest overlap was observed between C9 and C3 (27.9%), C9 and C5 (25.6%), C9 and C7 (20.9%), respectively **(**Fig. 3d**)**, indicating substantial clonal sharing among NK-like states. Interestingly, C8 also showed substantial overlap with C9 (23.3%), suggesting a clonal relationship between the cycling C8 state and NK-like cells. Additional overlap was observed between C9 and other cycling subtypes, C4 (16.3%) and C6 (16.3%). In contrast, C0–C2 showed limited clonal sharing with the NK-like subtypes, with no detectable overlap between C9 and C0, C1 or C2. This descriptive analysis identified the transcriptional state pairs with the greatest degree of clonal sharing

We next investigated whether these observed pairwise overlaps were greater than expected by chance. Using a patient-stratified permutation test (5,000 permutations), we compared the observed clonal overlap for each transcriptional-state pair with the overlap expected under the null model (see Methods; permutation-based evaluation of clonal relationships as in LRT) (22). Multiple state pairs showed significantly greater clonal sharing than expected by chance (BH-adjusted $P < 0.05$; **Supp. Table S9**), including relationships among the NK-like (C3, C5, C7 and C9) and cycling (C4, C6 and C8) CD8 + subtypes. We further examined whether enriched clonal relationships were consistently observed across individual patients. C3–C5 and C3–C8 showed significant clonal overlap in four of five patients and were also significant after combining patient-level P values using Fisher's method (BH-adjusted $P < 0.05$; **Supp. Table S10**). However, the extent of clonal connectivity varied across patients, with several state-pair associations detected in only one or two samples, indicating inter-patient heterogeneity in clonal relationships.

To further resolve these relationships at the individual-clonotype level, we identified clonotypes spanning multiple transcriptional states. Two unique clonotypes were detected across all five states examined (C3, C5, C7, C8 and C9), comprising 29 and 205 cells, respectively. Additional clonotypes spanned four states, with state combinations involving C3, C5, C7, and C8 or C3, C5, C8 and C9; these clonotypes were detected within individual patients (**Supp. Table S11**). These findings provide clonal-level evidence that individual TCR clonotypes can populate multiple transcriptional states, complementing the pairwise overlap and cross-patient reproducibility analyses. Taken together, this analysis at the individual-patient level revealed marked heterogeneity in clonal fate **(Supp. Fig. S10a)**. In four of five patients (QEH234, QEH235, QEH236 and QEH238), the top 10 expanded clonotypes per-patient were predominantly distributed across NK-like and cycling states and generally showed higher cytotoxicity and exhaustion signature scores **(Supp. Fig. S10b)**. In contrast, expanded clonotypes in QEH237 were dominated by conventional states C0–C2 and showed comparatively lower cytotoxicity as well as higher memory and

Tc2 signature scores **(Supp. Fig. S10b)**. Notably, the NK-like expanded clonotype in QEH237 retained a high cytotoxicity signature, indicating that the comparatively lower cytotoxicity across the top expanded clonotypes reflected their distinct transcriptional composition rather than reduced cytotoxicity of the NK-like state itself.

Pseudotime analysis reveals NK-like CD8 + T cells differentiation

To further investigate potential differentiation relationships among NK-like CD8 + CIK cells during *ex vivo* expansion, we performed pseudotime trajectory analysis (Fig. 4a). We focused on clonally related states C3, C5, C7, C8 and C9, which showed significant clonal sharing. The highly cycling subtypes C4 and C6 were not included because their transcriptional profiles were dominated by cell-cycle-associated programs, which can represent a major source of variation that confounds inference of differentiation trajectories (23). In contrast, C8 was retained despite its cycling phenotype because, in addition to cycling features, it displayed intermediate effector and exhaustion-associated transcriptional programs, placing it within the broader transcriptional continuum of NK-like CIK cell states examined here. Additionally, although significant clonal sharing was detected between C9 and the cycling subtypes C4 and C6 (**Supp Table S9-10**), these associations not accompanied by the multi-state clonotype patterns observed among C8 and the NK-like states (**Supp Table S11**). C8 was therefore selected as the starting state for trajectory inference, resulting in three major branching pseudotime trajectories (Fig. 4a-b). The Tem-int NK-like subtype C5, consistent with our previous characterisation of an intermediate-effector phenotype (Fig. 2a), occupied an intermediate position along all three trajectories, which terminated in C3, C7 or C9. Two terminal states, C3 and C9, were characterised by effector NK-like T-cell programs, with C9 displaying Temra-like features and C3 (Teff NK-like) displaying an early exhausted effector phenotype. The third trajectory terminated in C7 (Tex NK-like), characterised by an exhaustion profile, and progressed through both C5 and C3 as intermediate states (Fig. 4c**)**.

Gene expression patterns along pseudotime further supported these distinct trajectories. Effector-associated genes, including *GZMA, GNLY, NKG7* and *PRF1*, increased along the trajectories toward cytotoxic states, whereas their expression was markedly reduced toward C7 (Fig. 4d; **Supp. Fig. S11a**). Conversely, exhaustion-associated genes, including *HAVCR2* and *TOX*, progressively increased toward C7 and reached their highest levels in this state (Fig. 4d; **Supp. Fig. S11b**). Together, these findings support a model in which clonally related NK-like CD8 + CIK cells occupy divergent transcriptional trajectories, with C5 and C3 representing closely related intermediate/effector states, C9 representing a highly cytotoxic Temra-like state, and C7 representing an exhaustion-associated state with reduced cytotoxic features (24).

To investigate regulatory programs associated with the inferred trajectories, we examined regulon activity across pseudotime (Fig. 4e**)**. Proliferation-associated *FOXM1* activity was enriched toward early pseudotime, consistent with the cycling C8 subtype (25). In contrast, effector-associated regulons, including *TBX21, EOMES*, and *CEBPD*, showed enrichment in cytotoxic subtypes (C9, C5, C3) across intermediate-to-late pseudotime, coinciding with the emergence of cytotoxic NK-like transcriptional subtypes (26). Distinct regulatory programs were also associated with the late pseudotime state. *SMAD4, NFATC2, RUNX3*, and *IKZF1* showed preferential regulon activity in the Tex-like NK-like subtype (C7), relative to the other subtypes examined (C8, C5, C9, and C3) (**Supp. Table S12**). Cluster-level analysis further identified regulons with preferential higher activity in individual NK-like subtypes (**Supp. Table S13**). Notably, preferential activity of *IKZF1* in C7, *GFI1* and *IKZF2* in C9, and *KLF2* in C5 aligned with the distinct transcriptional states of these subtypes. *IKZF1* activity was associated with the Tex-like C7 subtype, GFI1 with the highly cytotoxic terminal effector C9 subtype, and *KLF2* with the intermediate C5 subtype. These associations identified *IKZF1*, *GFI1*, and *KLF2* as candidate regulatory factors associated with distinct effector and dysfunctional states in NK-like CIK cell differentiation.

Together, these dynamic regulon activity patterns identify distinct regulatory programs associated with the transcriptional continuum from a cycling state toward cytotoxic and exhausted states, providing candidate regulatory mechanisms associated with the divergent phenotypes observed among clonally related CD8 + CIK cells.

DISCUSSION

Here, we constructed a single-cell transcriptome atlas of CD8 + CIK cells by integrating scRNA-seq and scTCR-seq. Focusing on NK-like T cells, we resolved NK-like CD8 + CIK cells's transcriptional heterogeneity, clonal relationships, and trace their lineage trajectories to a functional branch point at which an intermediate state diverges into functional cytotoxic or dysfunctional exhausted fates. This analysis provides a detailed characterization of CD8 + NK-like populations within CIK cell products, identifies clonal and transcriptional features underlying this divergence, and offers a rationale for strategies to guide CIK cell manufacturing toward a more effective therapeutic.

CIK cells comprise a heterogeneous population of immune cells (27). We first defined the major cellular populations within bulk day-14 autologous CIK cell products, a commonly used manufacturing end point for CIK cells. We confirmed that our CIK cells were predominantly composed of CD8 + and CD4 + αβ T cells, with smaller populations of NK, γ9δ2 T, and B cells. Furthermore, the ratio of CD4 + to CD8 + T cells was at approximately 50:50, suggesting that CD4 + T cells may have a role in CIK cell therapeutic efficacy (28–30). However, since CD8 + T cells are established key contributors to CIK-mediated cancer control, we have focused in this study on analysing the CD8 + CIK cells, characterising their transcriptional and functional heterogeneity, with a particular emphasis on the NK-like T cell populations (31).

As expected, CD8 + CIK cells could be broadly divided into conventional and NK-like T cells. Conventional T cells were primarily identified as memory-like populations. Surprisingly, we also identified a conventional Tc2-like population characterised by type 2 cytotoxic signatures, including *GATA3*, *GZMB*, IL*4*, and *IL13*(32, 33). As this Tc2-like subtype was consistently found in all donors, these Tc2-like cells may have currently unknown roles in CIK cell efficacy. A recent study demonstrated that type 2 cells within CAR-T infusion products can suppress a dysfunctional subset and contribute to population homeostasis (32), raising the possibility that Tc2-like cells in our CIK cell products could similarly influence the functional state of other populations, including NK-like cells, and so warrant further investigation.

Beyond the identification of conventional CD8 + CIK subtypes, we revealed for the first time that NK-like CD8 + CIK cells comprise 4 functionally distinct subtypes. These subtypes are proposed to arise from proliferative cells with combined effector and exhaustion-associated features (C8), that lead to NK-like subtypes with an effector memory-like intermediate subtype (C5), a highly cytotoxic terminal subtype (C9), an effector subtype with cytotoxic and early exhaustion features (C3), and a dysfunctional subtype expressing highly expressing exhausted features with significantly reduced cytotoxicity (C7) consistent with previous reports describing dysfunctional T-cell phenotypes (34).

In four of the five patients, the most expanded clonotypes were predominantly associated with four NK-like subtypes (C3, C5, C7, and C9), whereas in the other patient the dominant clonotypes were primarily those associated with conventional CD8 + CIK subtypes (C0, C1, and C2). This occurred despite an increase in the overall NK-like population across all patients, suggesting that phenotypic enrichment and clonal expansion may represent partially distinct features of CIK product evolution. Compared with non-expanded clonotypes, expanded clonotypes within these NK-like subtypes exhibited higher expression of cytotoxicity- and exhaustion-associated genes, whereas non-expanded clonotypes showed greater expression of memory-associated genes including *LEF1* and *CCR7*. These findings align with previous observations that expanded T cell clonotypes predominantly exhibit increased expression of cytotoxicity and exhaustion genes(35).

We also observed that conventional CD8 + CIK cell subtypes (C0, C1, and C2) showed limited clonal overlap with NK-like subtypes. In contrast, shared clonotypes were detected across multiple NK-like subtypes (C3, C5, C7, C9) and cycling subtypes (C4, C6, C8). This suggests that clonally related CD8 + CIK cells, can occupy distinct transcriptional states, rather than representing discrete, clonally independent populations. Consistent with this observation, inferred trajectory analysis of these subtypes, using C8 as a starting state, positioned the C5 intermediate state as a branch point associated with 3 NK-like transcriptional states, which include an effector memory (C3), Temra (C9) and exhausted (C7) subtype (Fig. 5).

Regulon activity analysis further identified transcriptional regulatory features associated with distinct NK-like subtypes. While some regulons were shared across multiple subtypes, *KLF2*, *IKZF1* and *GFI1* showed preferential activity for C5, C7 and C9 respectively. The preferential *KLF2* activity observed in C5 is notable given previous evidence that *KLF2* supports functional effector potential while limiting terminal dysfunction (36, 37). Together with the intermediate position of C5 in the inferred trajectory, these findings suggest that *KLF2* activity may be associated with maintenance of an intermediate transcriptional state with potential to adopt distinct effector phenotypes. Preferential *IKZF1* activity in C7 was associated with the Tex-like transcriptional state. Consistent with our findings, previous studies have identified *IKZF1* as a regulator associated with T cell exhaustion (38, 39), raising the possibility that increased *IKZF1* activity contributes to the dysfunctional phenotype of exhausted CIK cells. Finally, *GFI1* showed preferential activity in effector associated Temra subtype (C9), consistent with previous studies indicating that *GFI1* contributes to terminal effector T cell differentiation and antitumor responses (40, 41). These findings highlight distinct regulatory programs associated with the transcriptional states of NK-like CIK cells. We hypothesise that manipulating candidate regulatory programs that distinguish effector from dysfunctional states may provide a strategy to influence the functional composition of CIK products.

This study has several limitations. First, our analysis focused on transcriptional and TCR profiling of CIK cell products at a single timepoint. Although we identified substantial cellular and clonal heterogeneity and findings consistent with relationships among transcriptional states, a single timepoint cannot establish the directionality of differentiation. We could not determine whether a blood-derived progenitor population gives rise to the NK-like CD8 + CIK cell subtypes (42) or whether additional transitional states are present at earlier stages of culture. Longitudinal profiling across multiple timepoints during CIK cell expansion will be important to define the emergence and progression of these NK-like states. Second, although scRNA-seq was performed on bulk CIK cell products, this study did not present analysis on the CD4 + T cell compartment. As CD4 + T cells constitute a substantial proportion of our CIK cell product (28), defining their composition and potential functional contribution will provide further insight into the determinants of CIK cell therapeutic efficacy (*manuscript in preparation*). Third, our dataset comprised five donors with metastatic CRC. Although we observed broadly consistent T cell subset composition across donors, validation in larger cohorts and across different cancer types will be required to determine the generalisability of these findings to CIK cell products generated for different clinical indications. Importantly, our identification of defining protein markers associated with these transcriptional states provides a potential approach for flow cytometric identification and monitoring of these populations. Finally, functional validation of the transcription factors highlighted by regulon activity analysis, including through CRISPR-based perturbation(43), will be important to determine whether these factors contribute causally to CIK cell phenotypic diversification.

Protocols for the manufacture of CIK cells vary in detail, and lack of standardisation continues to be an important challenge in the field. A more detailed understanding of the cellular composition and transcriptional states within CIK cell products is an important prerequisite for developing rational approaches to manufacturing standardisation. Our study defines the heterogeneity of CD8 + CIK cells, identifying both expected cell populations and previously unappreciated transcriptional states. By linking these states to distinct clonal and regulatory features, our findings provide a framework for systematically evaluating how culture conditions and other manufacturing parameters influence CIK cell differentiation. Our findings of distinct regulatory programs associated with the transcriptional continuum from a cycling state toward cytotoxic or exhausted states, suggests that it may be possible to manipulate culture conditions to favour the development of cells with specific phenotypes. Such approaches may lead to more consistent CIK cell products and inform strategies to enrich for cellular states with greater therapeutic potential.

MATERIALS AND METHODS

Subjects

A total of five treatment-naïve patients diagnosed with metastatic colorectal cancer (mCRC) enrolled in the study which was conducted at The Queen Elizabeth Hospital (TQEH, Woodville, South Australia). Clinical details of the patients can be found in Table 1. Written informed consent was obtained from all participants. Blood samples were taken prior to the patients undergoing treatment. The research team assigned patient identifiers (QEH234-238) that cannot

be linked to patient identity outside of the study group. The Human Research Ethics Committee of the Central Adelaide Local Health Network granted ethical approval for this research. (HREC/14/TQEHLMH/164).

Table 1
Patient Characteristics

Patient ID	Sex	Age Range	Cancer Primary Origin
QEH 234	M	70–75	Caecum
QEH 235	M	65–70	Transverse colon
QEH 236	M	55–60	Caecum
QEH 237	M	65–70	Hepatic Flexure
QEH 238	F	45–50	Sigmoid

CIK cell generation

Cytokine induced killer (CIK) cells from blood of mCRC patients were generated using our previously reported CIK cell expansion protocol (12). Briefly, peripheral blood mononuclear (PBMC) cells were isolated from whole blood using Ficoll Paque™ PLUS (GE Healthcare Bio-sciences AB, Sweden,) density gradient centrifugation. Freshly isolated PBMCs were initially seeded at a concentration of 5×10^5 cells/mL in serum free media X-VIVO 15 (LONZA, Switzerland) supplemented with 1000 U/mL of interferon gamma (*IFN-γ*) (Miltenyi Biotec, USA) in 24 well tissue culture grade plates. After one day, 0.05µg/mL of anti-*CD3* (Miltenyi Biotec, USA) and 300 IU of interleukin *(IL)-2* (Miltenyi Biotec, USA) were added to the culture. After three days, an additional 500 µL of media containing 300 IU of *IL-2* was added. Every 3 to 4 days, cells were divided to maintain a density of 1×10^6 cells/mL, while 300 IU of *IL-2* was supplemented during the 14-day culture period.

Flow cytometry

Flow cytometry to detect expansion of CD3 + CD56+ CD8 + T cells (CD8 + NK-like) within CIK cell cultures was performed on day 14 as a quality control prior to scRNA-sequencing (**Supp. Fig. S1**). To do this, freshly isolated PBMCs and day 14 expanded CIK cells were collected and stained with BD Horizon Fixable Viability Stain 780 (FVS780; Biolegend, USA), followed by an incubation with 50 µL Fc block (BD Biosciences, USA) for 10 min in 4° C to prevent non-specific binding. To stain for cell surface markers, the cells were incubated with a panel of fluorophore-conjugated antibodies in FACS buffer (2% FBS, 0.05% sodium azide, 1 mM EDTA in sterile 1× PBS), including anti-human CD3 (clone UCHT1) BV510, CD56 (clone 5.1H11) PECy7, CD4 (clone A161A1) FITC, CD8 (clone SK1) Alexa Fluor 647, CD226 (clone TX25) PE, and CD314 (clone 1D11) BV421 (Biolegend, USA). Next day, the cells were washed in FACS buffer and acquired on a Cytek Aurora flow cytometer (Cytek Biosciences, USA). Data were analysed using FlowJo (v10). Plots depicting the expansion of NK-like CIK cells, as well as the frequencies of CD8 + cells within CD3 + CD56+ populations (NK-like) and within CD3 + CD56- T cells over 14 days of culture are shown in **Supp. Fig. S1c-d**.

Single-cell RNA sequencing

On day 14 of CIK cell culture, cells were collected and resuspended at approximately 1×10^3 cells/ul, targeting the capture of about 1×10^4 cells per library for single-cell 5' gene expression and TCR sequencing using the 10X Chromium platform based on NextGEM Single Cell 5' Reagent Kits v2 (Dual Index) protocol (CG000331 Rev F). Library preparation and sequencing were performed by the South Australian Genomics Centre (SAGC, Adelaide, Australia) using MGI DNBSEQ-G400 sequencer. Two libraries were prepared :a gene expression library that contains sequencing reads and a T cell receptor library containing TCR information as nucleotide sequence (V(D)J library).

Single-cell RNA sequencing (scRNA-seq) analysis

Sequence preprocessing and demultiplexing

The FASTQ files of five samples (QEH234-238) obtained from sequencing were processed using the 10X Genomics *Cell Ranger* (v6.1.2, 10X Genomics) multi pipelines. This pipeline outputs single cell feature counts, TCR sequence assignments, and associated annotations for each library. Alignment of reads was performed with the GRCh38 human reference genome. Filtered annotated contig files, clonotype information, features, barcodes, and matrix generated by the Cell Ranger were used as input files for downstream analysis.

Analysis workflow for single-cell transcriptomic data

All gene expression analyses were performed using R (v4.5.1) and Seurat package (v5.4.0). Other R packages used during this analysis include dplyr (v1.0.8), patchwork (v1.1.1), ggplot2 (v3.3.5), pheatmap (v1.0.13), and RColourBrewer (v1.1.2). To correct for ambient RNA contamination, we applied SoupX (v1.6.2) to each sample independently. Following RNA ambient correction, Seurat objects were created for each sample, and doublets were individually removed from all samples using the scDblFinder package (v1.24.10). The cleaned samples were then merged for downstream analysis. To ensure removing low-quality cells, following quality control parameters were used. Cells exhibiting more than 5% mitochondrial genes, and fewer than 400 or more than 4000 unique feature counts were excluded from the analysis. The quality filtered gene expression matrices were normalized using Seurat's *NormalizeData* function with the LogNormalize method. Highly variable features were identified using the vst method, selecting the top 2,000 variable genes across all cells. The data were then scaled by *ScaleData* function of Seurat library and principal component analysis (PCA) was performed on the top 30 principal components. To account for batch effects across samples, harmony integration was calculated on the PCA embeddings using the harmony correction algorithm (v1.2.4). The

harmony adjusted components were then used to generate a shared nearest-neighbour graph with the *FindNeighbors* function. It was followed by unsupervised clustering using the Louvain algorithm at a resolution of 0.8. To identify optimal clustering resolution, we performed clustering across a range of resolutions (0.1 to 1.4, step = 0.1) and evaluated cluster stability using the clustree package (v 0.5.1) (**Supp. Fig. S2a**). A resolution of 0.8 was selected as it provided the highest resolution while preserving robust cell subtypes (**Supp. Fig. S2a**). The detailed information about identified cell subtypes was discussed in the result section of the main text. Finally, Uniform Manifold Approximation and Projection (UMAP) was performed on the batch corrected data using the top 30 dimensions for visualization of cells. In total, 43,047 high-quality cells were obtained from five samples for downstream analysis.

Identification of subtype-specific marker genes and cell subtypes

We determined cluster-specific marker genes using the *FindAllMarkers* function in Seurat with the parameter test.use set to "wilcox", facilitating differential expression analysis between each cluster and all other cells. Next, we ranked the top 100 differentially expressed genes (DEGs) for each subtype based on their average log2 fold change. Subsequently, the cell subtypes were annotated using the top 100 differentially expressed genes from each cell subtype. A comprehensive list of the top differentially expressed genes for all cell subtypes is available in **Supp.** Table 1, 2. Visualisation of joint density function and expression intensity of specified genes were generated using the Nebulosa package (v 1.20.1)(44). Detailed descriptions of additional analyses steps are provided in the **Supplementary Methods**.

Isolation of CD8 + CIK cells

CD8 + αβ CIK cells (n = 14,973) were extracted from the integrated CIK cells dataset (n = 43,047) using Seurat. The subsetting process was based on LogNormalized expression values. CD8 + αβ T cells were defined as cells with the following threshold: CD8A > 0 | CD8B > 0) & CD4 = = 0 & TRGV9 = = 0 & TRDV2 = = 0). This threshold excluded CD4 + T cells, γ9δ2T cells, and any potential doublets, yielding 14,973 CD8 + αβ T cells (34.8% of the total dataset) for downstream analysis. To validate the specificity of this population, expression of the pan-T-cell marker CD3 was assessed independently of the subsetting criteria: 99.0% of cells classified as CD8 + CIK cells expressed at least one CD3 subunit transcript (CD3D, CD3E, or CD3G), confirming the T-cell identity of the selected population.

TCR data integration and clonotype annotation

We analysed T cell receptor (TCR) sequences and clonal expansion using the Cell Ranger V(D)J pipeline. TCR data from all five CIK samples were merged with the combineTCR function in the scRepertoire R package (v 2.6.2). We cleaned the data by removing barcodes without any TCR information, cells with more than one TCR, and non-productive rearrangements (using removeNA = TRUE, removeMulti = TRUE, and filterNonproductive = TRUE). Therefore, only T cells that contain at least one functional α-chain and one β-chain remained for downstream analysis. Clonotypes were defined by matching the exact V(D)J gene segments and the nucleotide sequence of the CDR3 region. We then combined these clonotype labels to our main Seurat object with *combineExpression* function. Hence we could look at gene expression and clonal behaviour at the same time. Of the 43,047 high-quality cells after preprocessing, 34,525 cells had a detectable clonotype. These formed 27,915 unique clonotypes, of which 2,460 clonotypes were comprised of more than one cell. Clonotype sizes range from 1 cell to 1146 cells. Any T cells that shared the same α or β-chain sequence were considered as belonging to the same clonotype. To measure clonal expansion, we calculated how many cells belonged to each clonotype. We grouped clones by size: singlets (1 cell), small (2–9 cells), expanded (10–20 cells), and hyperexpanded (> 20 cells). Cells where no αβTCR could be detected were labelled "No TCR". We highlighted individual large clones or single cells directly on the UMAP.

Clonal overlap between transcriptional states

To quantify pairwise clonal overlap between transcriptional states, we computed, for each state pair (A, B), the proportion of unique clonotypes in the smaller state that were also present in the other state:

$$\varvec{C}\varvec{l}\varvec{o}\varvec{n}\varvec{a}\varvec{l}\varvec{o}\varvec{v}\varvec{e}\varvec{r}\varvec{l}\varvec{a}\varvec{p}_{(\varvec{A},\varvec{B})} = \v$$

To assess whether the observed clonal overlaps were greater than expected by chance, we performed a patient-stratified permutation test. Cell-state labels were randomly shuffled within each patient while preserving the original clonotype assignments. The overlap between each pair of states was recalculated across 5,000 permutations to generate a null distribution, from which empirical one-sided P values were calculated. P values were adjusted for multiple testing using the Benjamini–Hochberg procedure. This approach is conceptually similar to methods used to assess non-random clonal relationships in integrated scRNA-seq and scTCR-seq analyses (22). To assess whether clonal overlaps were reproducible across patients, we repeated the permutation analysis separately for each patient. For each state pair within each patient, cell-state labels were randomly shuffled 5,000 times to generate a null distribution and an empirical one-sided P value. P values from the patient-level tests were adjusted for multiple testing using the Benjamini–Hochberg procedure. We then combined the patient-level P values for each state pair using Fisher's method and applied a second Benjamini–Hochberg correction across state pairs. State pairs with an adjusted Fisher P < 0.05 were considered reproducibly enriched across patients.

Individual-clonotype state spanning analysis

To examine whether individual TCR clonotypes were detected across multiple transcriptional states, we analysed cells from C3, C5, C7, C8 and C9. Clonotypes were defined using the paired TCR α- and β-chain sequence identifier (CTstrict). For each clonotype, we counted the number of cells assigned to each state. A clonotype was considered present in a state when it was detected in at least two cells. Clonotypes detected in two or more states were classified as multi-state clonotypes. For each multi-state clonotype, we recorded the states occupied, the total number of cells, and the number and identity of patients in which it was detected.

Pseudotime Trajectory Analysis

To reconstruct the transcriptional continuum of clonally related CD8 + CIK cells, Monocle 3 (v1.4.29; https://github.com/cole-trapnell-lab/monocle3) was applied (45). Cells were first subsetted to TCR-sharing subtypes (C3, C5, C7, C8, C9) identified in Seurat, and the default assay was set to raw RNA counts to preserve untransformed expression values for downstream modeling. Following conversion to a Monocle cell_data_set (CDS) object, we retained the Seurat-derived UMAP coordinates. For gene selection, we combined a balanced set of subtype-specific markers (the top 20 positively differentially expressed genes per subtype, identified using *FindAllMarkers* with min.pct = 0.25 and logfc. threshold = 0.5) with a curated panel of biologically relevant genes associated with NK-like cytotoxicity, memory, and exhaustion phenotypes (*GZMB, PRF1, KLR family, TCF7, PDCD1, TOX, LAG3*). Dimensionality reduction was then performed on this curated gene set using preprocess_cds with 50 principal components. Cells were clustered using Louvain community detection (cluster_cells), and a principal graph was learned in the UMAP space using learn_graph. Pseudotime was computed by designating subtype C8 as the root state, with the root principal node defined as the graph vertex closest to the majority of C8 cells using a custom function. Module scores for cytotoxicity (*GNLY, PRF1, GZMB, GZMH, GZMK, IFNG, NKG7*) and exhaustion gene signatures (*CTLA4, TOX, HAVCR2, TIGIT, LAG3, LAYN*) were calculated using *AddModuleScore* in Seurat and visualized as a function of pseudotime to assess their relationship along the trajectory. Density distributions of pseudotime per subtype were visualized using kernel density estimation. Individual genes associated with exhaustion and cytotoxicity were further visualised as normalized expression along pseudotime.

PySCENIC analysis

Regulon activity was evaluated using the Python implementation of the SCENIC algorithm, pySCENIC (v0.12.1). Co-expression modules between transcription factors and potential target genes were inferred from the normalized expression matrix of αβCD8 + CIK cells using a gradient boosting-based approach conceptually based on *GRNBoost2*, applied to the union of the top 5,000 highly variable genes and all detected transcription factors. Modules were filtered to retain direct targets by cis-regulatory motif enrichment (hg38 v10 databases). Regulons define a transcription factor together with its target genes. Regulon activity in single cells was quantified using AUCell. The resulting AUC score matrix was integrated with the Seurat object for downstream analysis and visualization.

Statistical analysis

For the experimental data, statistical analyses were performed using GraphPad Prism (V 10.1.0) using the test described in the figure legend. The number of samples included is specified in the figure legends. The detailed statistical methods are listed in the figure legends. For the scRNA-seq, statistical analysis and figure generation were conducted using R version 4.5.1 and python version 3.13.5.

Declarations

Author Contributions

E.Sh.: conceptualization, methodology, investigation, formal analysis, funding acquisition, writing –original draft, and writing – review and editing. C.L., R.L.: methodology, investigation, writing – review and editing. G.M., P.D., T.P.: Resources, Writing – review and editing. C.K.: formal analysis, writing – review and editing. B.E., C.S.: consumer partners, conceptualisation, writing – review and editing. M.R: formal analysis, supervision, writing – review and editing. E.S., Y.T..: resources, supervision, writing – review and editing. K.F: Conceptualization, investigation, supervision, project administration, funding acquisition, writing – original draft, writing – review and editing.

Acknowledgements

We would like to thank the medical staff from The Queen Elizabeth Hospital for their assistance and to the patients who provided samples for the study. Illustrated figures were generated using BioRender.

Funding

This work was supported by a Cure Cancer Research Grant (K.F.), a Cancer Council SA Research Fellowship (K.F) and an Adelaide University Research Training Program Scholarship (E.S.).

Conflict of Interest

The authors declare no conflict of interest.

Ethics Statement

Ethics approval for the collection and storage of patient samples was granted by Central Adelaide Local Health Network Human Research Ethics Committee (Adelaide, South Australia) (HREC/14/TQEHLMH/164), and all patients had signed written informed consent.

Supplementary Information

All supplementary materials (Data and Figures) are included in this manuscript.

Data Availability

The raw sequencing data generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) repository under accession code [GSE349184].

References

1. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Translational Oncology. 2021;14(10):101174.
2. Li CMY, Tomita Y, Dhakal B, Li R, Li J, Drew P, et al. Use of cytokine-induced killer cell therapy in patients with colorectal cancer: a systematic review and meta-analysis. J Immunother Cancer. 2023;11(4).
3. Li CMY, Tomita Y, Dhakal B, Tin T, Li R, Wright JA, et al. Generation and assessment of cytokine-induced killer cells for the treatment of colorectal cancer liver metastases. Cancer Immunology, Immunotherapy. 2024;73(1):6.
4. Sangiolo D, Martinuzzi E, Todorovic M, Vitaggio K, Vallario A, Jordaney N, et al. Alloreactivity and anti-tumor activity segregate within two distinct subsets of cytokine-induced killer (CIK) cells: implications for their infusion across major HLA barriers. Int Immunol. 2008;20(7):841-8.
5. Kim HM, Kang JS, Lim J, Kim JY, Kim YJ, Lee SJ, et al. Antitumor activity of cytokine-induced killer cells in nude mouse xenograft model. Arch Pharm Res. 2009;32(5):781-7.
6. Tan G, Spillane KM, Maher J. The Role and Regulation of the NKG2D/NKG2D Ligand System in Cancer. Biology (Basel). 2023;12(8).
7. Ying Li CM, Li R, Drew P, Price T, Smith E, Maddern GJ, et al. Clinical application of cytokine-induced killer (CIK) cell therapy in colorectal cancer: Current strategies and future challenges. Cancer Treatment Reviews. 2024;122:102665.
8. Meng M, Li L, Li R, Wang W, Chen Y, Xie Y, et al. A dynamic transcriptomic atlas of cytokine-induced killer cells. J Biol Chem. 2018;293(51):19600-12.
9. Hsiao CH, Chiu YH, Chiu SC, Cho DY, Lee LM, Wen YC, et al. Isolation and Expansion of Cytotoxic Cytokine-induced Killer T Cells for Cancer Treatment. J Vis Exp. 2020(155).
10. Liu J, Wang L, Wang Y, Zhang W, Cao Y. Phenotypic characterization and anticancer capacity of CD8+ cytokine-induced killer cells after antigen-induced expansion. PLoS One. 2017;12(4):e0175704.
11. Ghanbari Sevari F, Mehdizadeh A, Abbasi K, Hejazian SS, Raeisi M. Cytokine-induced killer cells: new insights for therapy of hematologic malignancies. Stem Cell Research & Therapy. 2024;15(1):254.
12. Li CMY, Tomita Y, Dhakal B, Tin T, Li R, Wright JA, et al. Generation and assessment of cytokine-induced killer cells for the treatment of colorectal cancer liver metastases. Cancer Immunol Immunother. 2024;73(1):6.
13. Jiang X, Rudqvist NP, Jiang B, Ye S, He S, Liang Q, et al. Depletion of Effector Regulatory T Cells Associates with Major Response to Induction Dual Immune Checkpoint Blockade. Cancer Discov. 2025;15(8):1569-92.
14. Cheng ZY, He TT, Gao XM, Zhao Y, Wang J. ZBTB Transcription Factors: Key Regulators of the Development, Differentiation and Effector Function of T Cells. Front Immunol. 2021;12:713294.
15. Rebuffet L, Melsen JE, Escalière B, Basurto-Lozada D, Bhandoola A, Björkström NK, et al. High-dimensional single-cell analysis of human natural killer cell heterogeneity. Nature Immunology. 2024;25(8):1474-88.
16. Raskov H, Orhan A, Christensen JP, Gögenur I. Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. British Journal of Cancer. 2021;124(2):359-67.
17. Chen W, Luo J, Ye Y, Hoyle R, Liu W, Borst R, et al. The Roles of Type 2 Cytotoxic T Cells in Inflammation, Tissue Remodeling, and Prostaglandin (PG) D(2) Production Are Attenuated by PGD(2) Receptor 2 Antagonism. J Immunol. 2021;206(11):2714-24.
18. Saha A, Minowa T, Shavkunov AS, Salmon AJ, Keshari S, Jarjour NN, et al. Bhlhe40 Coordinates T Cell Programs with Distinct CD4 and CD8 T Cell Requirements for Anti-PD-1 Versus Anti-CTLA-4. bioRxiv. 2025.
19. Ma J, Wu Y, Chen L, Zheng X. CD8(+) T(EMRA) cells: A double-edged sword in immunity and disease-Mechanisms and therapeutic targets. Int Immunopharmacol. 2026;179:116614.
20. Hoyt-Miggelbrink AM, Waibl Polania J, Wachsmuth L, Lorrey S, Mohan A, Hardigan A, et al. Upregulation of TNFR2 Precedes TOX Expression by Exhausted T cells and Restricts Antitumor and Antiviral Immunity. Clin Cancer Res. 2026;32(4):782-800.
21. Sawabe A, Okazaki S, Nakamura A, Goitsuka R, Kaifu T. The orphan G protein-coupled receptor 141 expressed in myeloid cells functions as an inflammation suppressor. J Leukoc Biol. 2024;115(5):935-45.
22. Xie J, Jeon H, Xin G, Ma Q, Chung D. LRT: Integrative analysis of scRNA-seq and scTCR-seq data to investigate clonal differentiation heterogeneity. PLOS Computational Biology. 2023;19(7):e1011300.
23. Charrout M, Reinders MJT, Mahfouz A. Untangling biological factors influencing trajectory inference from single cell data. NAR Genomics and Bioinformatics. 2020;2(3).
24. Kasmani MY, Zander R, Chung HK, Chen Y, Khatun A, Damo M, et al. Clonal lineage tracing reveals mechanisms skewing CD8+ T cell fate decisions in chronic infection. Journal of Experimental Medicine. 2022;220(1).
25. Raghuwanshi S, Gartel AL. FOXM1 induces therapy resistance and inhibits apoptosis in a variety of human cancers. Cell Death & Disease. 2026;17(1):230.
26. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature. 2018;564(7735):268-72.
27. Lin C, Lee A, Lee K, Chiu H, Huang Y. THE CYTOKINE-INDUCED KILLER CELLS-BASED IMMUNOTHERAPIES FOR ADVANCED SOLID TUMORS: CLINICAL TRANSLATION AND EFFICACY EVALUATION. Cytotherapy. 2024;26(6, Supplement):S180-S1.
28. Liu S, Meng Y, Liu L, Lv Y, Yu W, Liu T, et al. CD4(+) T cells are required to improve the efficacy of CIK therapy in non-small cell lung cancer. Cell Death Dis. 2022;13(5):441.

29. Meng Y, Yu Z, Wu Y, Du T, Chen S, Meng F, et al. Cell-based immunotherapy with cytokine-induced killer (CIK) cells: From preparation and testing to clinical application. Human vaccines & immunotherapeutics. 2017;13(6):1379-87.
30. Li CMY, Li R, Drew P, Price T, Smith E, Maddern GJ, et al. Clinical application of cytokine-induced killer (CIK) cell therapy in colorectal cancer: current strategies and future challenges. Cancer Treatment Reviews. 2024;122:102665.
31. Chen D, Sha H, Hu T, Dong S, Zhang J, Liu S, et al. Cytokine-induced killer cells as a feasible adoptive immunotherapy for the treatment of lung cancer. Cell Death & Disease. 2018;9(3):366.
32. Bai Z, Feng B, McClory SE, de Oliveira BC, Diorio C, Gregoire C, et al. Single-cell CAR T atlas reveals type 2 function in 8-year leukaemia remission. Nature. 2024;634(8034):702-11.
33. Do J-S, Choi Y-H, Shin S-H, Yi HK, Hwang PH, Nam S-Y. Committed memory effector type 2 cytotoxic T (Tc2) cells are ineffective in protective anti-tumor immunity. Immunology Letters. 2004;95(1):77-84.
34. Collier JL, Weiss SA, Pauken KE, Sen DR, Sharpe AH. Not-so-opposite ends of the spectrum: CD8+ T cell dysfunction across chronic infection, cancer and autoimmunity. Nature Immunology. 2021;22(7):809-19.
35. Shorer O, Pinhasi A, Yizhak K. Single-cell meta-analysis of T cells reveals clonal dynamics of response to checkpoint immunotherapy. Cell Genom. 2025;5(5):100842.
36. Fagerberg E, Attanasio J, Dien C, Singh J, Kessler EA, Abdullah L, et al. KLF2 maintains lineage fidelity and suppresses CD8 T cell exhaustion during acute LCMV infection. Science. 2025;387(6735):eadn2337.
37. DePauw TA, Gai K, Shen J, Maurice NJ, Rhee KH, Valente WJ, et al. KLF2 overrides the resident memory CD8 T cell differentiation program, in opposition to KLF3. Proc Natl Acad Sci U S A. 2026;123(13):e2533700123.
38. Tay T, Bommakanti G, Jaensch E, Gorthi A, Karapa Reddy I, Hu Y, et al. Degradation of IKZF1 prevents epigenetic progression of T cell exhaustion in an antigen-specific assay. Cell Rep Med. 2024;5(11):101804.
39. Chiu H, Zhao J, Basavanhally T, Hsu C-C, Amatangelo M, Jain G, et al. Mezigdomide reverses T-cell exhaustion through degradation of IKZF1/IKZF3 and reinvigoration of cytokine production pathways. Blood. 2026;148(9):1115-28.
40. Ojo OA, Shen H, Ingram JT, Bonner JA, Welner RS, Lacaud G, et al. Gfi1 controls the formation of effector-like CD8+ T cells during chronic infection and cancer. Nature Communications. 2025;16(1):4542.
41. Chaudhry MZ, Chen E, Man HO, Jones A, Denman R, Yu H, et al. GFI1-driven transcriptional and epigenetic programs maintain CD8+ T cell stemness and persistence. Nature Immunology. 2025;26(6):894-907.
42. Franceschetti M, Pievani A, Borleri G, Vago L, Fleischhauer K, Golay J, et al. Cytokine-induced killer cells are terminally differentiated activated CD8 cytotoxic T-EMRA lymphocytes. Exp Hematol. 2009;37(5):616-28.e2.
43. Park B-S, Lee M, Kim J, Kim T. Perturbomics: CRISPR–Cas screening-based functional genomics approach for drug target discovery. Experimental & Molecular Medicine. 2025;57(7):1443-54.
44. Alquicira-Hernandez J, Powell JE. Nebulosa recovers single-cell gene expression signals by kernel density estimation. Bioinformatics. 2021;37(16):2485-7.
45. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32(4):381-6.

Figures

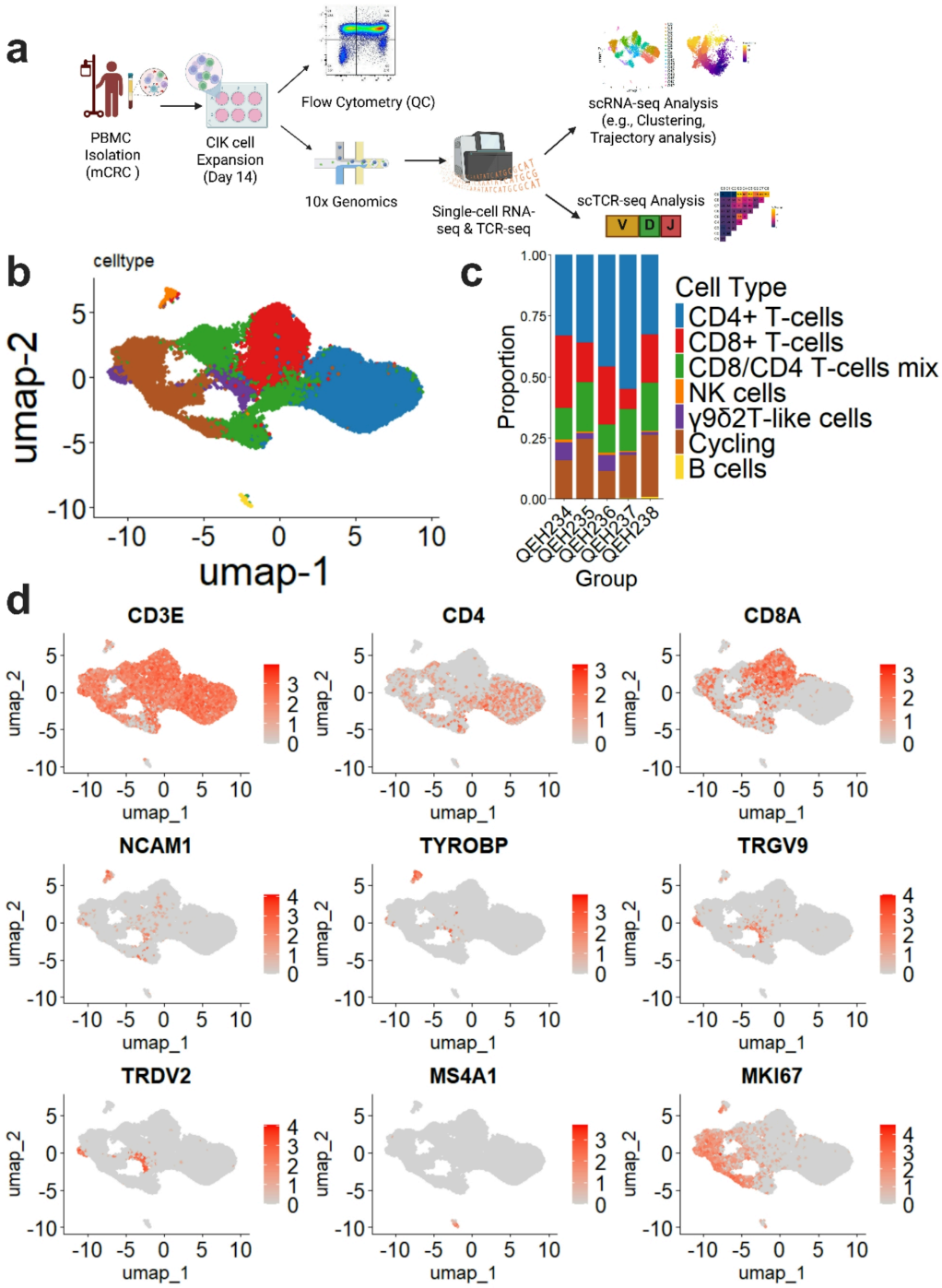


Figure 1

scRNA-seq analysis workflow and integrated clustering of CIK cells.

(**a**) Study design overview for paired scRNA-seq and scTCR-seq of cytokine-induced killer (CIK) cells from metastatic CRC donors (n=5). Illustration prepared using BioRender.com (**b**) UMAP visualization with labelling by cell-type annotations. (**c**) Bar chart depicting proportional cell-type compositions of individual samples. (**e**) Feature plot overlay of marker genes defining the seven main groups, projected onto the integrated UMAP.

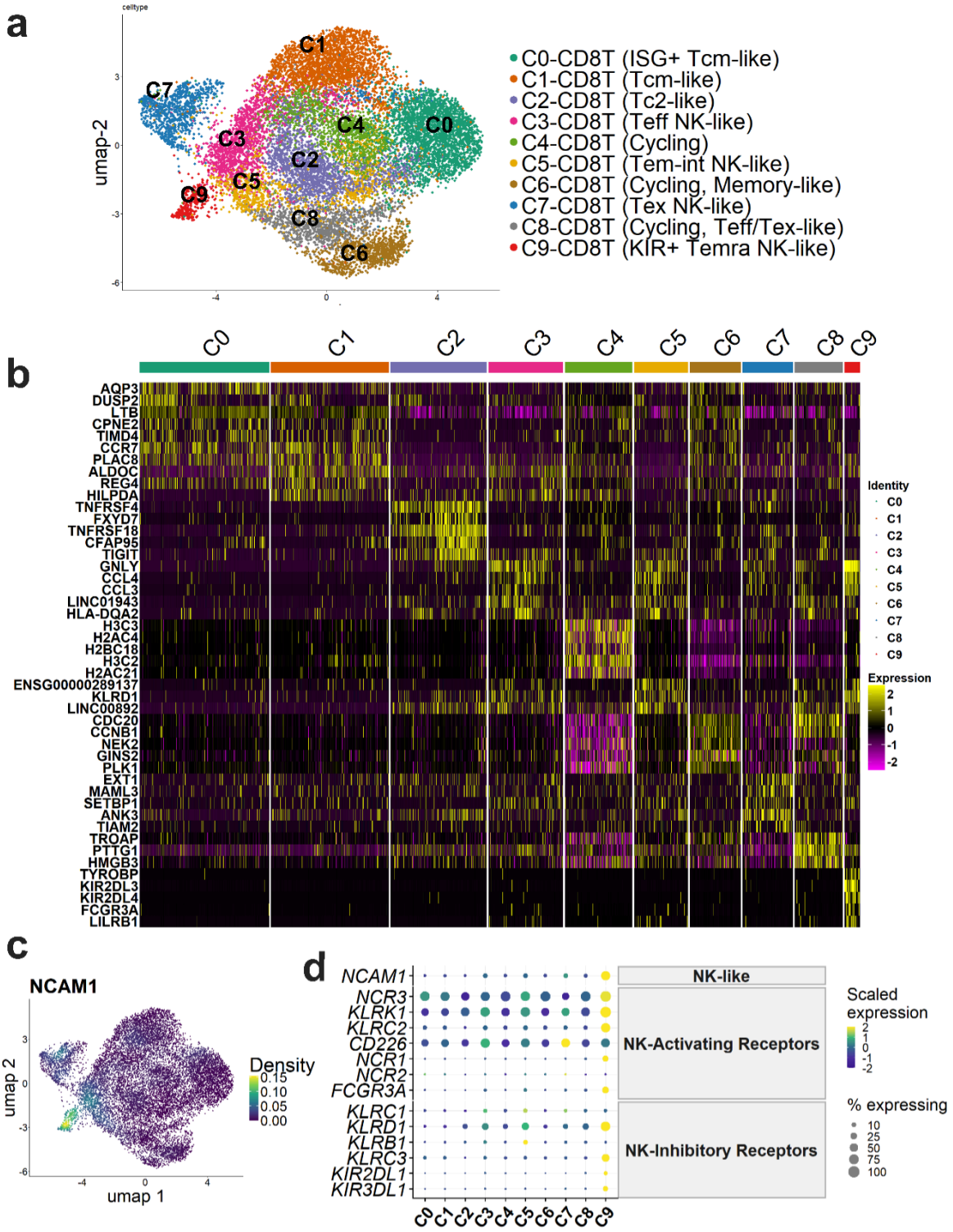


Figure 2

Phenotypic heterogeneity and functional signatures of CD8+ CIK cell clusters.

(**a**) UMAP embedding displaying 10 "Cytotoxic Clusters" (C0–C9) of CD8+ CIK cells, defined by unsupervised clustering. (**b**) Discriminative gene sets across each cluster were visualized as a scaled expression heatmap with the top 5 discriminative genes per cluster, ranked by average log2 fold-change, listed in **Supp. Table S2**. (**c**) Joint density plot of *NCAM1*(CD56) and CD8 expression projected onto the CD8+ CIK UMAP, highlighting cells with an NK-like transcriptional phenotype. (**d**) Dot plot of NK-activating (*NCR3, KLRK1, CD226, NCR2, FCGR3A*) and NK-inhibitory (*KLRC1, KLRB1, KLRD1, KIR2DL1, KIR3DL1*) receptor genes across C0–C9. Dot color indicates scaled average expression, dot size indicates the percentage of cells expressing each gene per cluster.

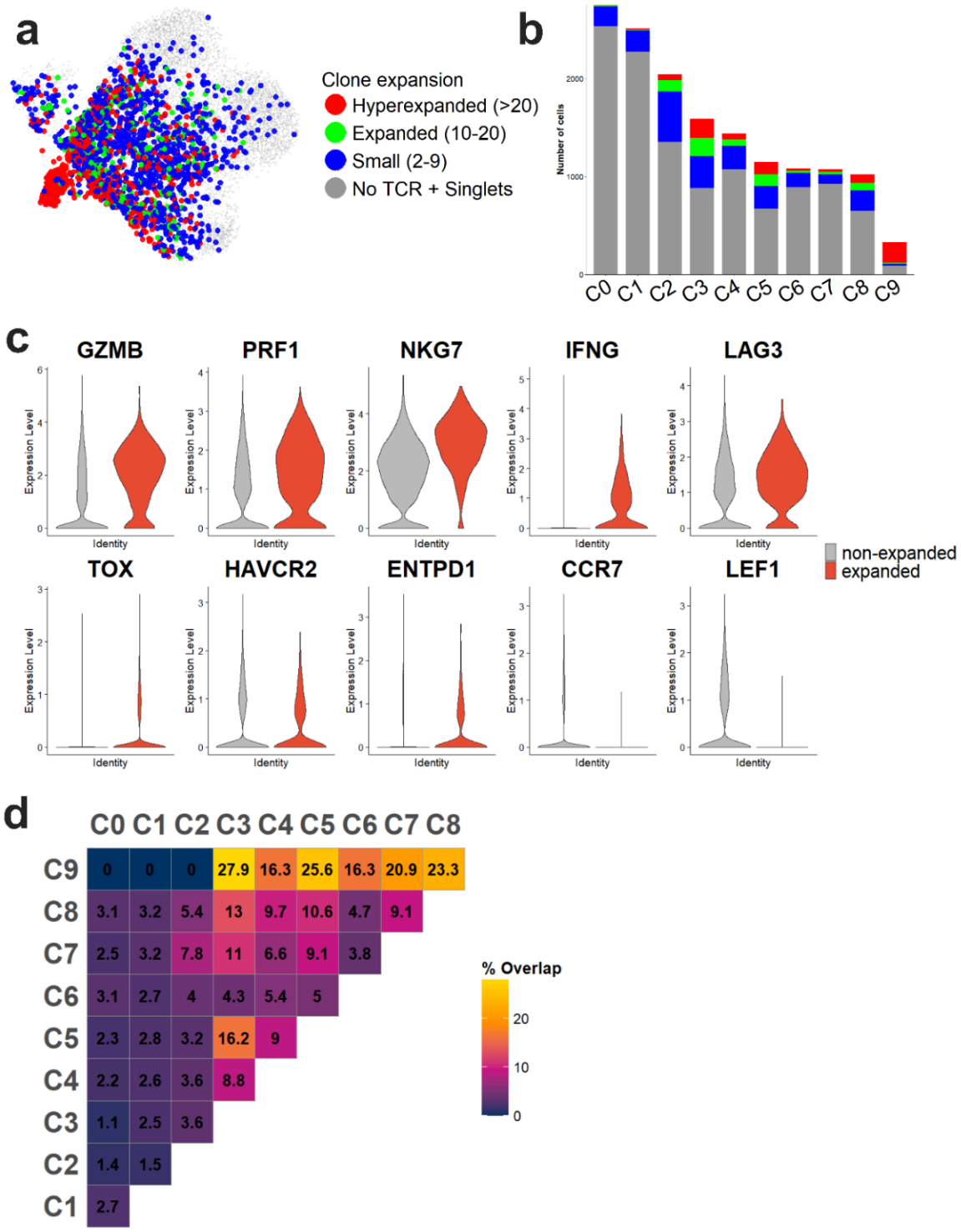


Figure 3

Clonal expansion patterns and transcriptional features in CD8+ CIK cells.

(**a**) UMAP embedding of CD8+ CIK cells by clonal expansion size (hyperexpanded >20 cells, expanded 10–20, small 2–9, singlets and no TCR). (**b**) Bar chart summarising clonal expansion distribution across cell clusters. (**c**) Differential gene expression analysis was performed comparing hyperexpanded and expanded clonotypes (>=10 cells per clonotype) with small and singlet clonotypes (<10 cells per clonotype). Cells without detected TCRs were excluded from this analysis. (**d**) Heatmap depicting clonal overlap matrix between cluster pairs, by proportion of shared clones

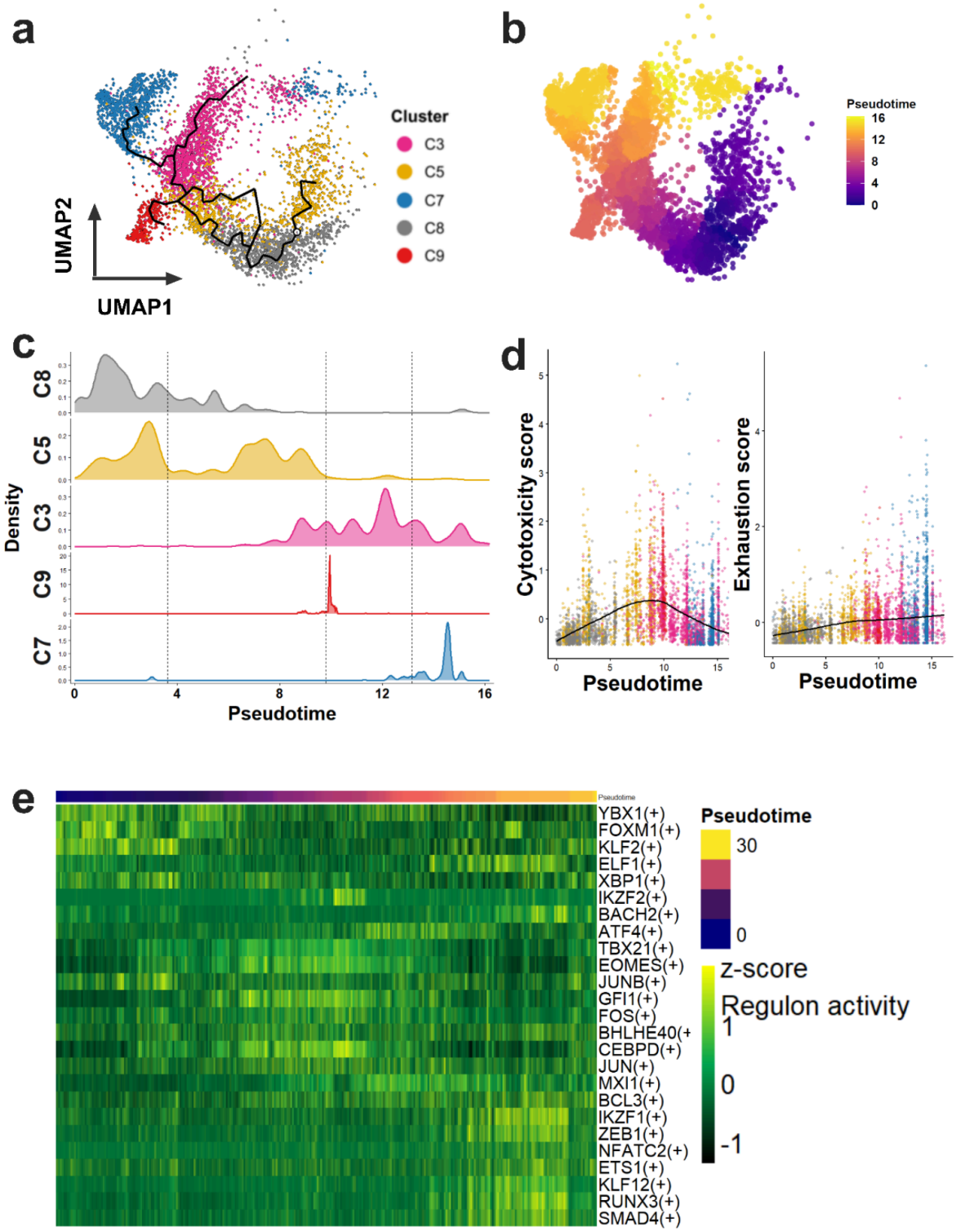


Figure 4

Pseudotime trajectory and regulon dynamics of clonally linked CD8+ CIK cells.

(**a**) UMAP showing the trajectories of Cytotoxic Clusters C3, C5, C7, C8, and C9, selected based on extensive shared TCR clonotype overlap, over pseudotime inferred by Monocle3. Cells are coloured by cell subtype identity, trajectory rooted at C8. (**b**) The same trajectory coloured by pseudotime value, illustrating the ordering of cells from the trajectory origin (dark blue) to terminal branches (light yellow). (**c**) Density distribution of cells along pseudotime, shown separately for each cluster, depicting the relative position of each cluster along the inferred trajectory. (**d**) Exhaustion and cytotoxicity score across pseudotime, coloured by cluster. (**e**) Heatmap showing activating regulon (+) along the pseudotime trajectory. Only a manually curated activating regulon selected for established roles in T cell proliferation, activation, effector differentiation, and exhaustion is displayed.

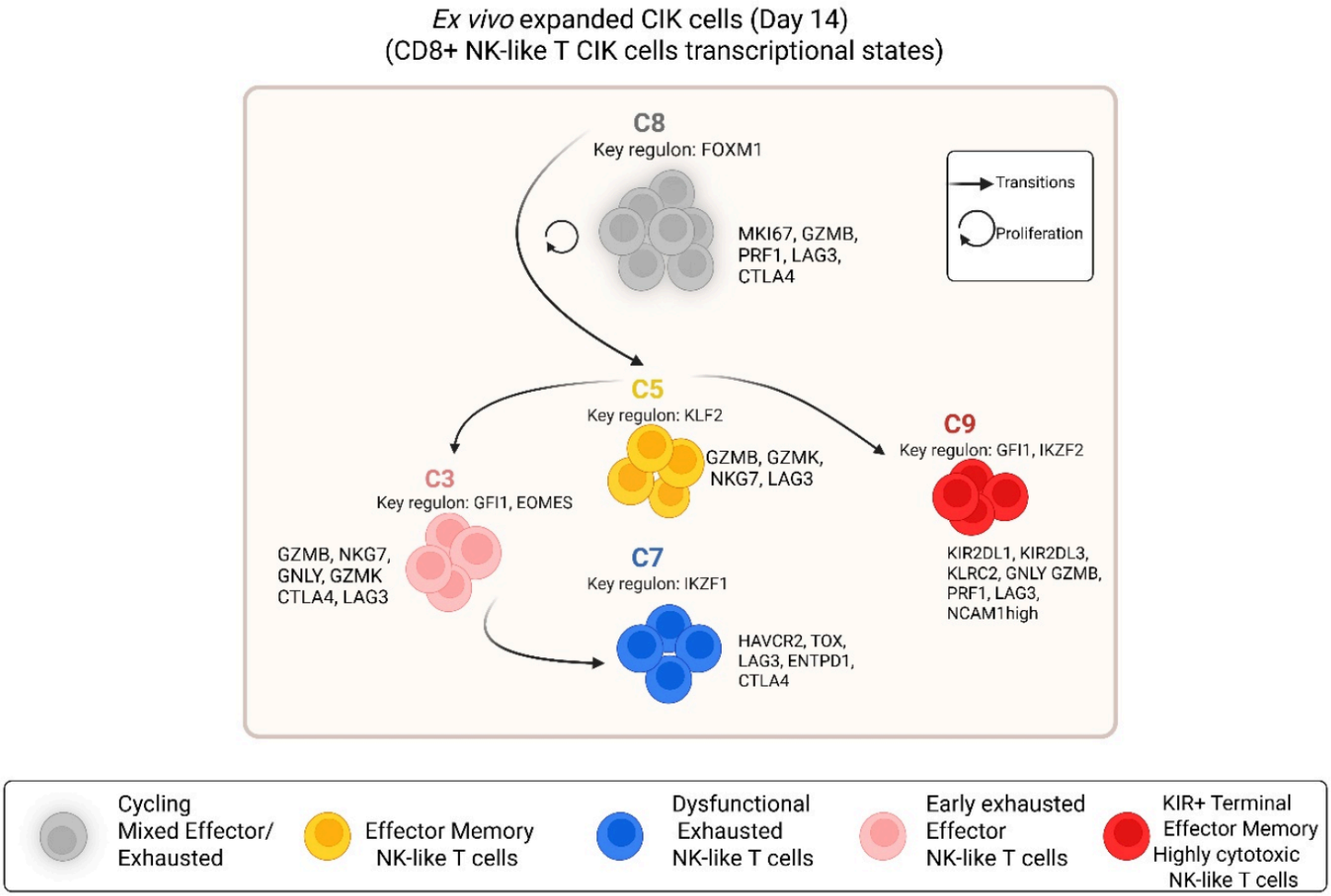


Figure 5

Proposed CD8+ NK-like CIK transcriptional states in CIK cell culture.

The model depicts the 4 NK-like transcriptional states, and a proposed progenitor identified in CIK culture, key genes and regulons of each state, and the proposed transitional relationships between these states.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ShokatiSupplementaryData.xlsx
- ShokatiCIKscRNAseqSupplementaryMaterials.docx