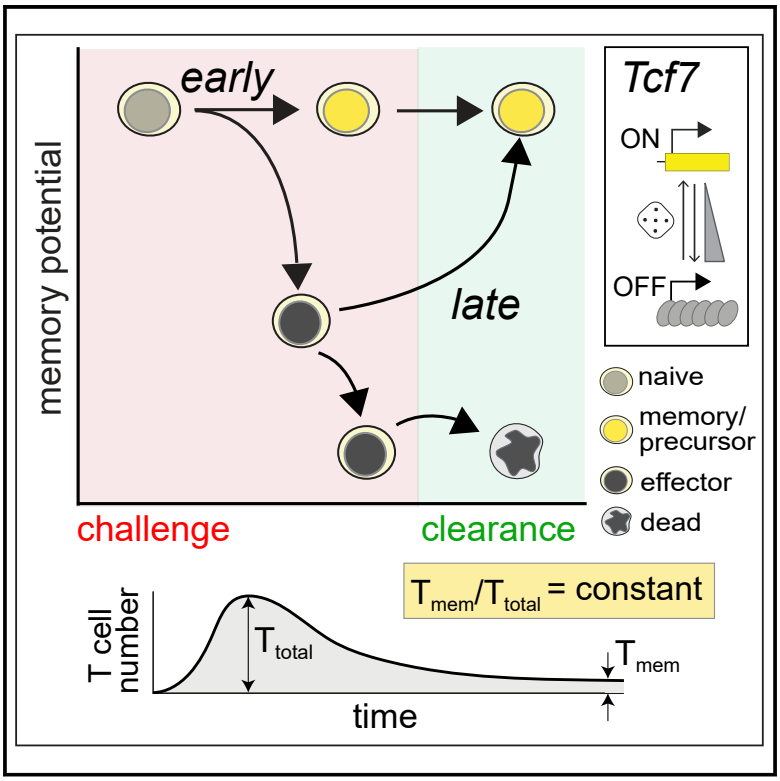
Article

Immunity

Reversible, tunable epigenetic silencing of TCF1 generates flexibility in the T cell memory decision

Graphical abstract



Authors

Kathleen Abadie, Elisa C. Clark, Rajesh M. Valanparambil, ..., Jay Shendure, Junyue Cao, Hao Yuan Kueh

Correspondence

shendure@uw.edu (J.S.), jcao@rockefeller.edu (J.C.), kueh@uw.edu (H.Y.K.)

In brief

When and how T cells make effector or memory decisions is not fully resolved. Abadie, Clark, and Valanparambil et al. uncover a reversible epigenetic switch enabling memory cells to diverge early after activation or dedifferentiate from effector cells later during antigen clearance. Thus, memory populations can scale with infection severity.

Highlights

- CD8$^+$ T cell memory decisions occur early upon antigen encounter and after clearance
- A reversible epigenetic switch controlling TCF1 enables flexible decision making
- Switching is stochastic, generating heterogeneous outcomes within clonal lineages
- Flexible decision making enables robust scaling of memory to infection severity

Abadie et al., 2024, Immunity *57*, 271–286
February 13, 2024 © 2023 Elsevier Inc.
https://doi.org/10.1016/j.immuni.2023.12.006

CellPress

Immunity

CellPress

Article

Reversible, tunable epigenetic silencing of TCF1 generates flexibility in the T cell memory decision

Kathleen Abadie,[1,13] Elisa C. Clark,[1,13] Rajesh M. Valanparambil,[2,13] Obinna Ukogu,[3] Wei Yang,[4] Riza M. Daza,[4] Kenneth K.H. Ng,[1] Jumana Fathima,[1] Allan L. Wang,[1] Judong Lee,[2] Tahseen H. Nasti,[2] Avinash Bhandoola,[5] Armita Nourmohammad,[3,6,7] Rafi Ahmed,[2] Jay Shendure,[4,8,9,10,12,*] Junyue Cao,[4,11,*] and Hao Yuan Kueh[1,12,14,*]

[1]Department of Bioengineering, University of Washington, Seattle, WA 98195, USA
[2]Emory Vaccine Center and Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, GA 30322, USA
[3]Department of Applied Mathematics, University of Washington, Seattle, WA 98105, USA
[4]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
[5]T-Cell Biology and Development Unit, Laboratory of Genome Integrity, Center for Cancer Research, National Cancer Institute, National Institute of Health, Bethesda, MD 20892, USA
[6]Department of Physics, University of Washington, Seattle, WA 98105, USA
[7]Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[8]Brotman Baty Institute for Precision Medicine, Seattle, WA 98195, USA
[9]Allen Discovery Center for Cell Lineage Tracing, Seattle, WA 98109, USA
[10]Howard Hughes Medical Institute, Seattle, WA 98195, USA
[11]Laboratory of Single-Cell Genomics and Population Dynamics, The Rockefeller University, New York, NY 10065, USA
[12]Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA 98109, USA
[13]These authors contributed equally
[14]Lead contact
*Correspondence: shendure@uw.edu (J.S.), jcao@rockefeller.edu (J.C.), kueh@uw.edu (H.Y.K.)
https://doi.org/10.1016/j.immuni.2023.12.006

SUMMARY

The immune system encodes information about the severity of a pathogenic threat in the quantity and type of memory cells it forms. This encoding emerges from lymphocyte decisions to maintain or lose self-renewal and memory potential during a challenge. By tracking CD8$^+$ T cells at the single-cell and clonal lineage level using time-resolved transcriptomics, quantitative live imaging, and an acute infection model, we find that T cells will maintain or lose memory potential early after antigen recognition. However, following pathogen clearance, T cells may regain memory potential if initially lost. Mechanistically, this flexibility is implemented by a stochastic *cis*-epigenetic switch that tunably and reversibly silences the memory regulator, TCF1, in response to stimulation. Mathematical modeling shows how this flexibility allows memory T cell numbers to scale robustly with pathogen virulence and immune response magnitudes. We propose that flexibility and stochasticity in cellular decisions ensure optimal immune responses against diverse threats.

INTRODUCTION

The immune system stores information about the nature and severity of prior infections through the generation of long-lived pathogen-specific memory lymphocytes during an immune response. This memory is encoded by the numbers and types of memory lymphocytes generated upon challenge. The quantity of memory T cells, in particular, scales with the magnitude of a prior infection, such that the memory population is a fixed fraction of the T cell number at the infection peak, across a range of pathogenic challenges.[1–3] This scaling in memory production is robust across T cell clones with different epitope specificities and allows the body to generate memory proportional to the severity of the pathogenic challenge. The regulatory mechanisms that enable this critical feature of adaptive immunity are not well understood.

The size and characteristics of the memory compartment are determined by the lineage decisions of T cells responding to an acute infection.[4] As naive CD8$^+$ T cells respond to antigens, some maintain long-term viability and self-renewal potential and thereby persist to form memory cells as the infection is cleared, while the majority terminally differentiate to form cytotoxic effectors. One class of models posits that memory cells form directly from naive cells without passing through an effector phase, through an early lineage bifurcation that concurrently gives rise to either memory-precursor or terminal effector cells.[5–7] A second class of models posits that this decision to progress toward memory or become short-lived effectors occurs

later, only after cells have undergone effector differentiation and upon resolution of an infection.[8–10] However, in contrast with both models, it is also possible that this process is inherently flexible,[11] such that T cells have multiple opportunities to commit to the memory state. In a number of stem cell systems, cell differentiation decisions are sometimes reversible,[12,13] and this plasticity may enable organisms to robustly maintain cell population sizes amid different settings. During an immune response, flexibility could optimize T cell memory formation for threats whose properties may only manifest as they unfold over time. It is unclear whether there exists such flexibility in T cell memory formation and, if so, what the underlying mechanisms and functional roles are.

T cell effector and memory differentiation is controlled by a circuit of transcription factors (TFs) and chromatin regulators that enable transitions between different states in response to external signals. A central node in this regulatory circuit is T cell factor 1 (TCF1, encoded by *Tcf7*), a TF essential for memory cell generation and self-renewal.[14] *Tcf7* is expressed in naive and memory cells, where it is crucial for maintaining self-renewal, and is silenced during effector differentiation, resulting in loss of memory potential and entry into a short-lived state.[5,15]

To follow *Tcf7* regulation and memory decision making in a controlled environment where cells can be continuously observed and signaling inputs carefully manipulated, we developed an *ex vivo* system to study T cell differentiation in response to stimulatory signals present during an acute infection. Using this system and by complementary testing in an *in vivo* acute infection model, we uncovered a flexible decision-making strategy: T cells can gain or lose memory potential at multiple junctures after antigen encounter and do so in a stochastic and reversible manner. Mathematical modeling revealed that this flexible decision-making strategy allows for the number of memory cells to scale linearly with total numbers of expanded T cells at the peak of infection, thereby encoding information about the severity of the prior threat. These findings unify two major models for memory lineage specification that are often regarded as mutually opposed and provide a quantitative framework for understanding immunological memory encoding.

RESULTS

A minimal *ex vivo* system for effector and memory differentiation of CD8$^+$ T cells

To disentangle candidate models for memory T cell formation (Figure 1A), we developed a minimal system (Figure 1B) in which naive (CD44$^-$CD62L$^+$) CD8$^+$ T cells with a yellow fluorescent protein (YFP) reporter for *Tcf7*[16] were activated with plate-immobilized anti-CD3 and anti-CD28 antibodies and interleukin (IL)-2, together with additional cytokines, present during acute infection (IL-12, IL-7, and IL-15[17–19]). These conditions minimized variability in the exposure of individual cells to stimulatory signals, enabling cell-intrinsic lineage control mechanisms to be studied apart from environmental heterogeneity.

In this system, all cells began dividing rapidly after 24 h and increased expression of the transmembrane glycoprotein CD44, indicating uniform activation (Figure 1C). Activated cells reduced transcription of *Tcf7* and the lymph-node-homing adhesion molecule CD62L, consistent with effector differentiation. The inflammatory cytokines IL-12 and interferon (IFN)-β1 enhanced *Tcf7*-YFP silencing (Figures 1D, S1C, and S1D), consistent with their roles in driving effector differentiation.[20,21] When T cell receptor (TCR) stimulation (anti-CD3/CD28) and inflammation (IL-12) were removed to mimic pathogen clearance, the cells demonstrated a population-level increase in CD62L and *Tcf7*-YFP while continuing to divide, as previously observed.[4] *Tcf7* and CD62L protein levels were heterogeneous, both during stimulation and after removal, suggestive of an early memory and effector differentiation decision. YFP expression closely matched TCF1 protein levels throughout activation, validating use of the reporter in this system (Figures S1A and S1B).

Naive CD8$^+$ T cells bifurcate early into effectors and memory precursors

To determine whether the heterogeneity in *Tcf7* and CD62L regulation reflected early memory and effector programming (Figure 1), we analyzed *ex vivo* activated cells using the temporally resolved single-cell transcriptome sequencing method, *sci-fate*.[22] Here, metabolic labeling of newly synthesized transcripts reveals a cell's current activity state apart from its history[22,23] (Figure 2A). We subjected cells at days 1, 2, and 4 to 4-thiouridine (4sU) pulse-labeling for 2 h, followed by sequencing and analysis as previously described.[22] We obtained old and new transcriptomes for ~17,000 single cells, with a median of 17,574 total and 2,529 new transcripts detected per cell (Figure S2A). To disentangle effector and memory gene programs from other activation-induced programs, we performed an integrative analysis of our temporally resolved transcriptome data and existing TF binding data[24] to identify TF modules, consisting of co-regulated groups of TFs and their cognate target genes (see STAR Methods). This analysis revealed two main TF modules: a cell cycle module and a T cell differentiation module, the latter further separable into submodules that included known regulators of effector and memory differentiation (Figures 2B and S2C).

By visualizing cell states using genes in the T cell differentiation module for uniform manifold approximation and projection (UMAP) dimensionality reduction, we resolved distinct effector and memory states with coherence between time points (Figures 2C, S2B, and S2D). Unsupervised clustering and differential gene expression analysis revealed early and late effector (E1 and E2) and memory precursor (MP) states. E1 and E2 cells exhibited higher expression of the effector-associated genes *Gzmb*, *Ifng*, *Tbx21*, *Zeb2*, and *IL12rb2*, while MP cells had higher expression of the stem- and memory-associated factors *Bach2*, *Lef1*, *Tcf7*, *Sell*, and *Slamf6*, and lower expression of effector-associated genes (Figures 2D, 2E, S2E, and S2F; Table S1).[26] These differential gene expression patterns were present at day 2 and amplified at day 4.

Consistent with an early fate bifurcation, RNA velocity vectors calculated using reads from newly synthesized transcripts originate from the undifferentiated state (U) and flow along separate effector and memory branches[25,27] (Figure 2F). To gain insight into the dynamics of genes differentially regulated between divergent trajectories, we visualized their expression over pseudotime along each trajectory (Figure 2G; Table S1). This analysis, together with RNA velocity and TF activity analysis (Figures 2H, 2I, and S2G), identified effector and memory regulators with greatest differential regulation along their

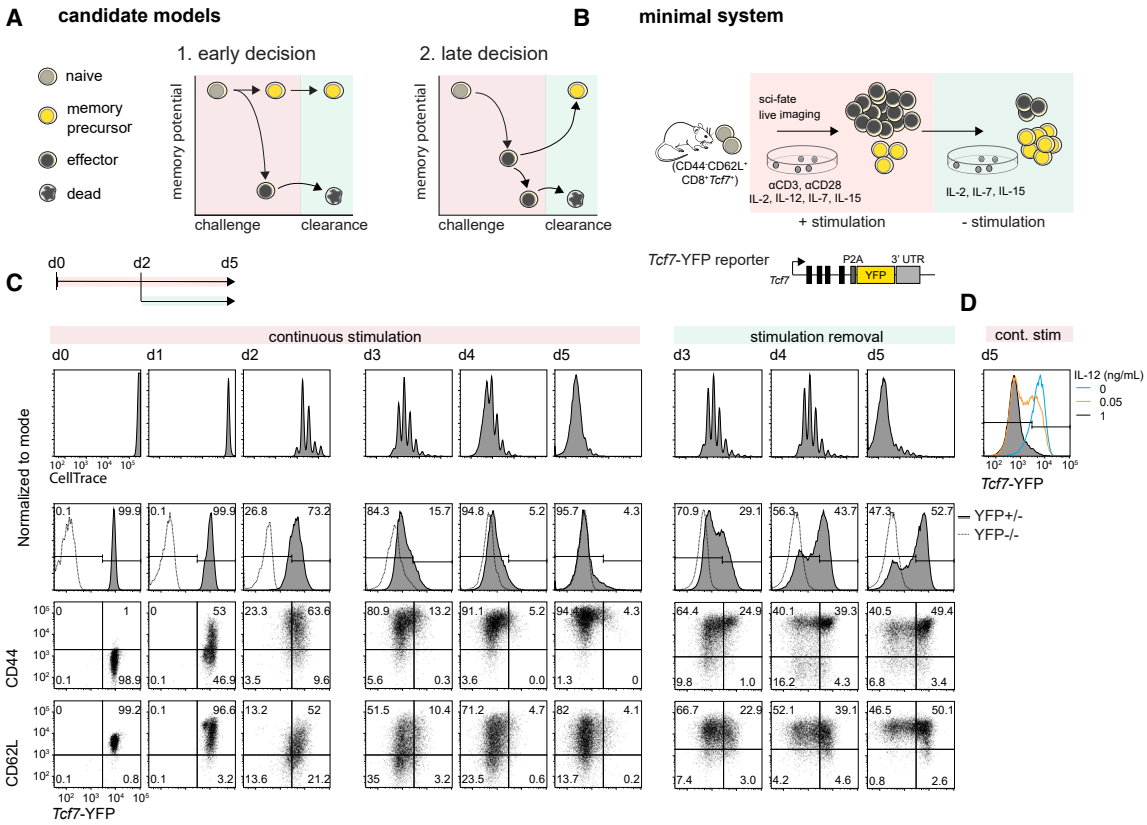


Figure 1. A minimal *ex vivo* system to track CD8$^+$ T cell effector and memory decision-making dynamics

(A) Candidate decision-making strategies for CD8$^+$ T cell memory generation.

(B) A minimal *ex vivo* system for tracking memory decision-making dynamics at the single-cell level.

(C and D) Naive CD8$^+$ T cells were isolated from *Tcf7*-YFP reporter mice, then cultured using this *ex vivo* system. (C) Flow cytometry plots show analysis of cultured cells during initial stimulation for 2 days (left) and continued stimulation to day 5 (middle), or after stimulation withdrawal (removal of αCD3/αCD28 after day 2 and IL-12 after day 3) (right). From top to bottom: CellTrace Violet (CTV) dilution; *Tcf7*-YFP histograms (shaded histograms are from *Tcf7*-YFP$^{+/-}$ mice and open histograms are from wild-type mice); flow plots of CD44 against *Tcf7*-YFP and of CD62L against *Tcf7*-YFP. (D) *Tcf7*-YFP silencing is tunable by IL-12 concentration. (C and D) Data are from a single experiment representative of at least 3 independent experiments.

respective trajectories. *Tbx21*, *Egr1*, and *Irf4*, among other effector regulatory genes, were specifically active along the E1 trajectory, while a distinct set of effector regulators, including *Eomes*, *Bhlhe40*, *Stat5a*, and *Stat3*, characterize the E2 trajectory. This effector heterogeneity and its potential influence on downstream differentiation will be interesting to investigate in future studies but is not further pursued here. Finally, regulators of T cell stemness and survival, including *Tcf7*, *Myb*, *Mxd4*, and *Fli1*, were active in the MP trajectory. *Tcf7* was the most significantly differentially expressed gene between trajectories, increased early along the MP trajectory, and was absent in both E1 and E2 trajectories. Its expression furthermore coincided with that of target genes identified through TF linkage that promote self-renewal, such as *Ikzf2*, *Sesn3*, *Aff3*, and *Pecam1* (CD31). Thus, *Tcf7* is a critical driver of this early divergent memory trajectory in our system.

The early effector and memory decision occurs heterogeneously within CD8$^+$ T cell clones

The divergence of cells into effector and memory lineages, occurring even under the strong, uniform stimulatory conditions of our *ex vivo* system, is suggestive of a cell-intrinsic regulatory mechanism involving *Tcf7* that generates heterogeneity in fate outcomes. To elucidate the degree to which this decision is heterogeneous within cell lineages amid constant environmental signals, we acquired multi-day time-lapse movies of clonal CD8$^+$ T cell lineages during activation, with continuous measurement of *Tcf7*-YFP expression (Figure 3). As T cells are difficult to track with live imaging due to their high mobility, tendency to adhere to one another, and rapid proliferation, we optimized adhesion conditions and computational analyses that allow continuous tracking of a fate-regulating TF across clonal CD8$^+$ T cell lineages (Figures 3 and S3; see STAR Methods).[28] We note that adhering T cells to plate-bound stimulatory molecules is expected to create differences compared with stimulation by antigen-presenting cells. However, we chose this minimal system to ensure that all cells received a uniform signaling environment, to disentangle cell-extrinsic versus intrinsic sources of heterogeneity. Using this method, we tracked 104 lineages over 4 days and an average of 4.4 cell generations, where each lineage is a family of imaged cells derived from a single naive precursor.

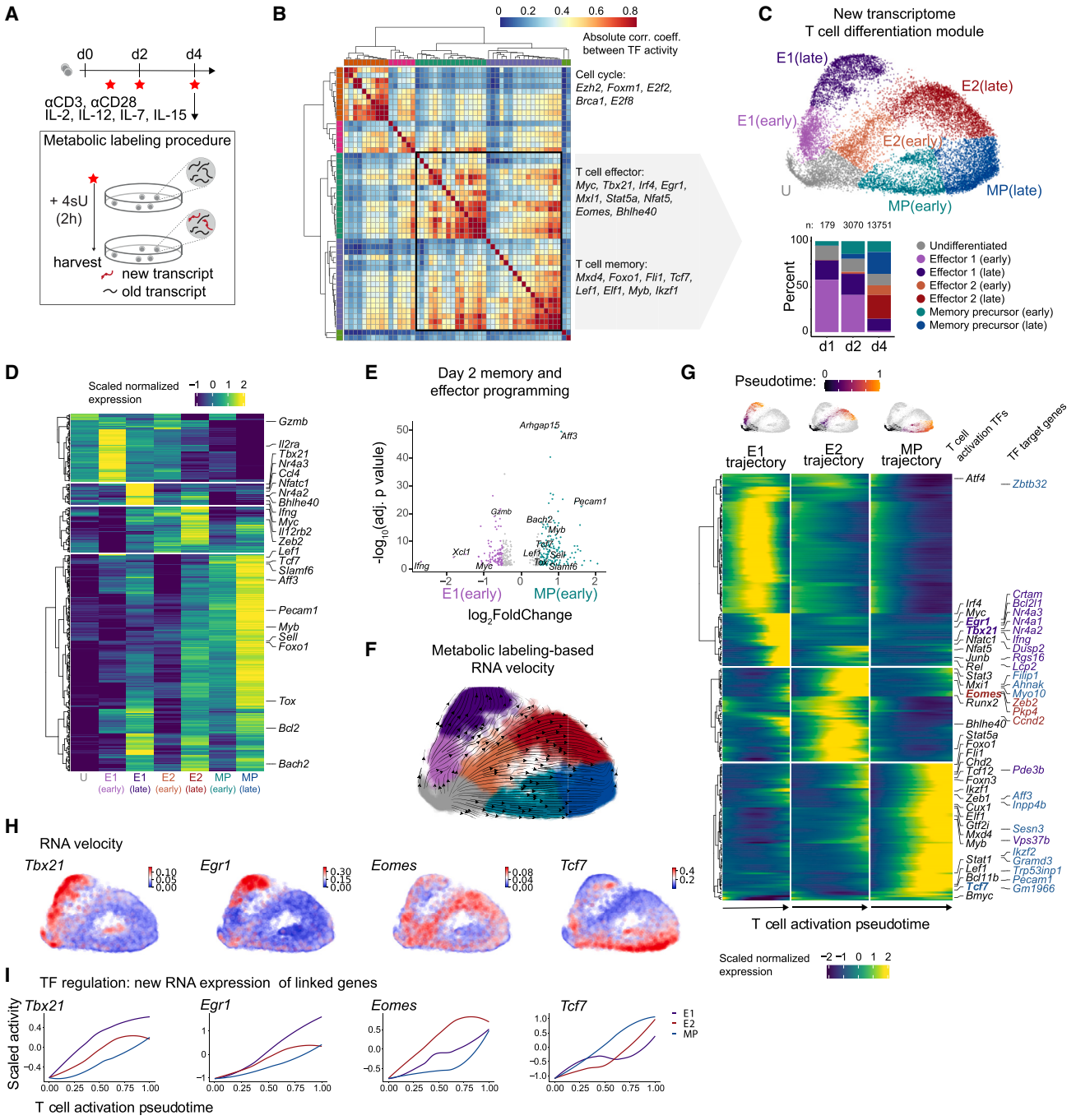


Figure 2. Naive cells diverge into effector and memory states early after activation

(A) Naive CD8$^+$ T cells were activated with αCD3 and αCD28 and 0.05 ng/mL IL-12, as well as IL-2, IL-7, and IL-15. After 1, 2, and 4 days, cells were treated with 4sU for 2 h to label new transcripts, then harvested for time-resolved transcriptomics using *sci-fate*.

(B) Heatmap showing the absolute Pearson's correlation coefficient between the activities of pairs of TFs. Key TFs in each module are labeled at right. T cell differentiation module used for subsequent analysis is boxed.

(C) UMAP visualization of cells based on the activity of T cell differentiation-related TF module, using newly synthesized mRNA, colored by cluster ID (top). Percentage of cells in each T cell differentiation state cluster after indicated days (bottom).

(D) Aggregated expression (scaled, $\log_{10}$ normalized) of top 400 differentially expressed (DE) genes between clusters (adj. $p < 3 \times 10^{-45}$ for all genes except for *Ifng*, adj. $p = 7.3 \times 10^{-29}$).

(E) DE genes between E1 (early) and MP (early) at day 2 only; $\log_2$FC > 0.5 and adj. $p < 0.05$.

(F) UMAP visualization as in (C), characterized by labeling-based RNA velocity analysis. Streamlines indicate the integration paths that connect local projections from the observed state to the extrapolated future state.[25]

(legend continued on next page)

Naive cells in these time-lapse movies started small, adhered to the antibody-bound plate, acquired expression of the activation marker CD69, increased dramatically in size, and divided rapidly after 1–2 days (Figure S3G; Video S1). Individual activating T cell clones often gave rise to *Tcf7*-YFP^{hi} and *Tcf7*-YFP^{lo} subpopulations (Figures 3A and S3J; Video S1), indicating that the effector and memory decision is made heterogeneously within clones. *Tcf7*-YFP^{hi} and *Tcf7*-YFP^{lo} cells showed similar degrees of attachment to the surface, indicating that these intraclonal differences were not due to differences in TCR stimulation but more likely due to cell-intrinsic mechanisms generating heterogeneity in *Tcf7* silencing.

Differences in *Tcf7*-YFP expression after multiple cell divisions likely stemmed from earlier *Tcf7* silencing events propagated through dilution of the stable fluorescent protein by cell division. To pinpoint the timing of early regulatory events that gave rise to these differences in *Tcf7*-YFP expression, we calculated the *Tcf7* promoter activity over time in single cells, defined as the rate at which total *Tcf7*-YFP abundance increased over time, using a hidden Markov model (HMM) to assign *Tcf7* promoter activity states to each cell at each time point and identify switching points between those states (Figures 3A–3C and S3A–S3F; see STAR Methods). This analysis revealed that cells silenced *Tcf7* expression at variable times after the onset of stimulation and could do so as early as the first cell division, as well as at later generations. Cells activated the *Tcf7* promoter prior to the first cell division, reflecting exit from quiescence, and then proceeded to switch the *Tcf7* promoter to a silent state. The timing at which the *Tcf7* promoter transitioned to the silent state varied between cell tracks both within and between cell lineages, consistent with observed heterogeneity in *Tcf7*-YFP expression within clones (Figures 3A–3D). Lineages with variability in *Tcf7* promoter states in their terminal progeny were more prevalent than those having a uniform ON or OFF promoter state (80%, n = 79), indicating that intraclonal heterogeneity in *Tcf7* regulation is the norm (Figures S3J and S3K). Removing IL-12 increased the fraction of cells in an active promoter state (Figures 3C and 3E). Silent *Tcf7* promoter states persisted across multiple cell divisions (Figures 3A, S3I, and S3J) and thus represent heritable regulatory changes as opposed to more transient dynamics such as transcriptional bursting. These results provide evidence that a cell-intrinsic *Tcf7* silencing event, occurring heterogeneously within clones, underlies the early divergence in effector and memory states.

A stochastic epigenetic switch controlling *Tcf7* silencing underlies the early CD8$^+$ T cell effector and memory decision

Heterogeneity in *Tcf7* silencing, as observed in our imaging assay, could derive from asymmetric cell division,[6,29] whereby cell fate determinants partition unequally, giving rise to discordant behavior between two sister cells. Alternatively, this heterogeneity could result from other molecular mechanisms that operate in an inherently stochastic manner[30–33] and would thus cause two sisters to make *Tcf7* silencing decisions independently. Although two sister cells could still silence *Tcf7* discordantly, they would do so no more frequently than expected by chance. To test these predictions, we analyzed the fractions of daughter cell pairs that silenced *Tcf7*, either discordantly (ON/OFF) or concordantly (OFF/OFF), doing so for cell pairs across all cell generations, with or without IL-12 (Figure 3F). By plotting concordant (OFF/OFF) versus discordant (ON/OFF) sister pair fractions, we found that all data points adhered to a theoretical curve representing the expected relationship between sister pair fractions for independent regulation (Figure 3G). Consistently, by statistical analysis using a modified Cohen's kappa coefficient (κ'), we found that daughter cells were no more likely to make discordant decisions than expected by chance (Figure 3H; Table S2). These findings support the view that *Tcf7* silences in a stochastic manner to drive divergent decisions within clones. These results do not rule out asymmetric division as a mechanism for memory and effector heterogeneity, whereby dendritic cell proximity and thus antigen exposure drives divergent fate decisions; however, they argue that differences in external signals are not necessary for the emergence of clonal heterogeneity in lineage decisions.

Epigenetic switching mechanisms, involving changes in chromatin modifications or conformation at gene loci, can introduce stochastic rate-limiting steps to gene activation or silencing.[34–36] Although epigenetic switching mechanisms are initiated by upstream TFs, they occur in an inherently stochastic manner, such that these upstream TFs do not directly modulate the levels of transcription but control the probabilities of all-or-none gene activation or silencing. As *Tcf7* silencing involves repressive DNA or histone methylation,[14,20,37,38] it could be gated by such a mechanism. Epigenetic mechanisms act in *cis* at individual gene loci and therefore would silence each *Tcf7* locus independently. To test for this mechanism, we compared *Tcf7*-YFP silencing kinetics in cells from mice homozygous (*Tcf7*-YFP$^{+/+}$) and heterozygous (*Tcf7*-YFP$^{+/-}$) for the reporter, with the prediction that *Tcf7*-YFP$^{+/+}$ reporter cells would yield a smaller population of *Tcf7*-YFP^{lo} cells, because both loci need to silence for loss of reporter expression (Figures 3I, 3J, and S3L). Indeed, the *Tcf7*-YFP^{lo} population was smaller in *Tcf7*-YFP$^{+/+}$ reporter cells and increased with IL-12, consistent with a *cis*-epigenetic silencing mechanism modulated by inflammation. Together, these results indicate that a stochastic *cis*-epigenetic switch, tunable by external stimuli, enables clonally related cells to make divergent effector and memory decisions, even when they are subject to uniform external signals.

Reversibility of *Tcf7* silencing enables a late CD8$^+$ T cell memory decision

Tcf7 silencing has been proposed to be an irreversible event that marks a "point of no return" for effector differentiation and loss of memory potential.[5,39,40] Conversely, various studies show that

(G) Pseudotemporal ordering of top 200 DE genes and additional genes of interest (adj. $p < 1.4 \times 10^{-17}$) between trajectories. Gene labels correspond to all DE TFs in the T cell differentiation TF module (left text) and DE target genes linked to *Tbx21*, *Egr1*, *Eomes*, and *Tcf7* (right text).
(H and I) (H) RNA velocity magnitude and (I) loess smoothed TF activity over pseudotime for four of the most DE genes between trajectories. TF activity is calculated as the normalized aggregation of newly synthesized mRNA for all TF target genes, scaled across all cells. Cells in the undifferentiated (U) cluster are set to pseudotime = 0 for each trajectory.

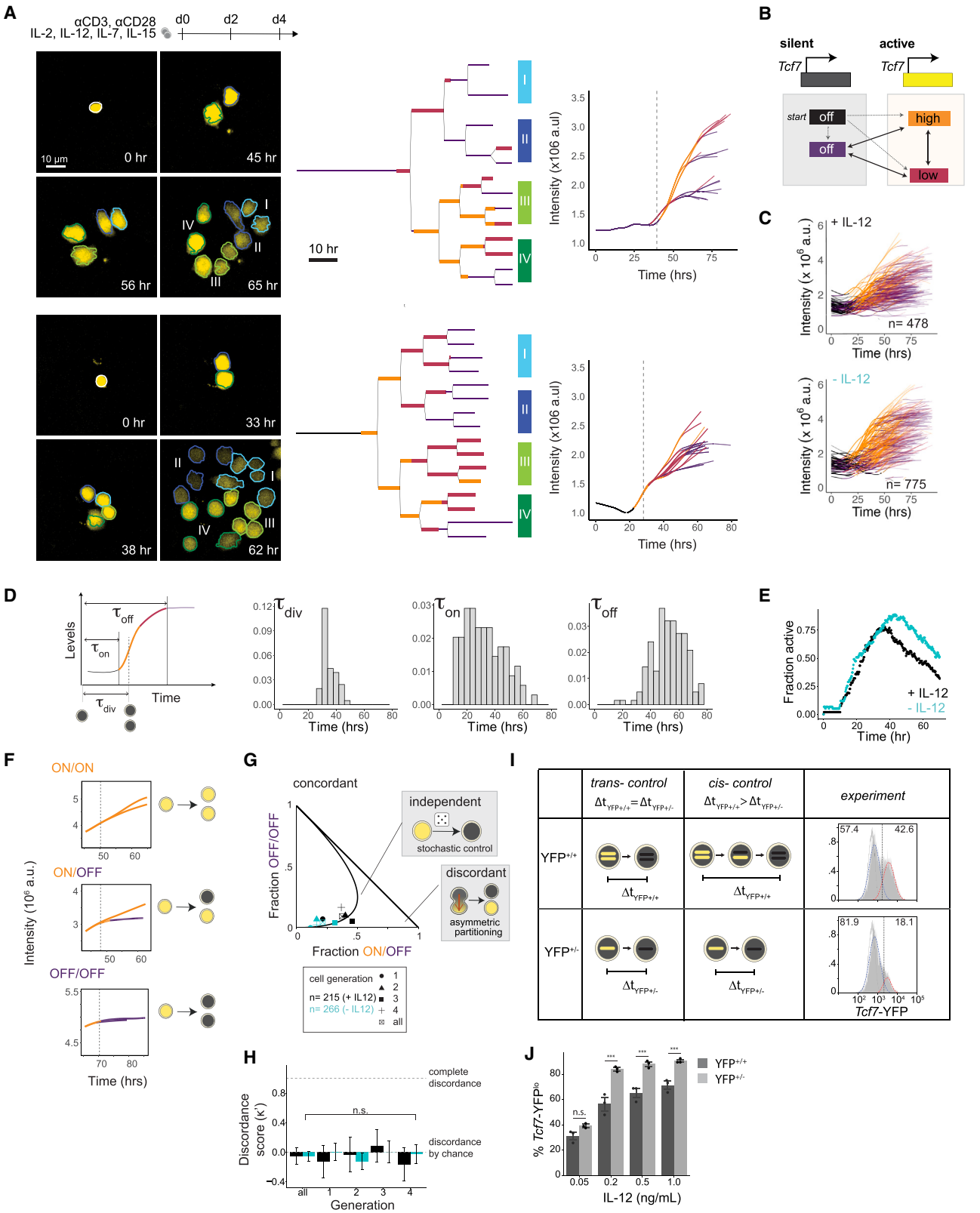


Figure 3. Heterogeneous *Tcf7* silencing within clones is controlled by a stochastic epigenetic switch

(A) Top: naive CD8$^+$ T cells were activated with αCD3, αCD28, and IL-12, as well as IL-2, IL-7, and IL-15. Cells are cultured with 1 ng/mL IL-12 unless otherwise indicated. Bottom: representative lineages demonstrating clonal heterogeneity in *Tcf7*-YFP silencing: image snap shots (left), lineage trees (middle), and reporter

(legend continued on next page)

cells that undergo effector differentiation are able to populate memory compartments after an infection is resolved,[8,9,41] suggesting that *Tcf7*-silenced effectors may still be able to reactivate *Tcf7* and reacquire memory potential. Our data thus far provide evidence for an early T cell decision to lose or maintain memory potential, driven by stochasticity in antigen-driven *Tcf7* silencing, but do not exclude the possibility that effector cells can regain memory potential later after withdrawal of stimulation.

To test this possibility, we sorted *Tcf7*-YFP^{lo} and *Tcf7*-YFP^{hi} cells after initial culture and subjected them to reculture with variable stimulation conditions *ex vivo* (Figure 4A). As expected, sorted *Tcf7*-YFP^{hi} cells maintained *Tcf7*-YFP expression without stimulation but underwent heterogeneous silencing under continuing stimulation (Figures 4B, 4C, and S4A). Furthermore, *Tcf7*-YFP^{lo} cells maintained a silent state upon continued stimulation, as observed. However, upon stimulation withdrawal, *Tcf7*-YFP^{lo} cells reactivated *Tcf7*, with the fraction of *Tcf7*-expressing cells increasing over 6 days. *Tcf7* reactivation upon stimulation withdrawal coincided with reduced CD25 and elevated CD62L expression, suggesting re-entry into a memory state (Figure 4D).

We next used clonal live imaging of sorted *Tcf7*-YFP^{lo} cells confined in microwells to test whether *Tcf7* reactivation was heterogeneous within individual effector clones, as would be expected if reactivation occurs via reversal of stochastic *cis*-epigenetic silencing (Figure 3). Consistent with reactivation observed from bulk starting populations, we found that a subpopulation of starting cells could reactivate *Tcf7* to give rise to *Tcf7*-YFP^{hi} cells (Figures 4E and 4F; Table S3; Videos S2 and S3). Similar to the initial *Tcf7* silencing event, reactivation was heterogeneous within clones. Overall, these results indicate that cells that have silenced *Tcf7* and lost memory potential can reverse this decision later, after resolution of an immune challenge.

A subpopulation of starting cells in these clonal imaging experiments did not reactivate *Tcf7* during imaging observation, suggesting that they may have irreversibly entered a *Tcf7*-silenced state and lost their memory potential (Figures 4E and 4F). T cells differentiate into a terminal effector state upon continued stimulation with inflammation[21]; as such, we asked whether the ability to reactivate *Tcf7* was dependent on the duration of antigen stimulation. We stimulated cells *ex vivo* for 2, 4, and 6 days, respectively, and analyzed *Tcf7* reactivation after 7 days of stimulation withdrawal. We observed that longer stimulation reduced the fraction of *Tcf7* reactivated cells (Figure S4B). We then asked whether this loss in plasticity was associated with inaccessibility at the *Tcf7* locus due to polycomb repressive complex 2 (PRC2)-mediated *cis*-epigenetic changes, as previously described.[37] We observed increased *Tcf7* reactivation in cells recultured in the presence of Tazemetostat or GSK126, EZH2/PRC2 inhibitors, suggesting that polycomb-mediated repression indeed plays a role in this loss of reactivation potential (Figure S4C). Taken together, these observations suggest that *Tcf7* repression is initially reversible, enabling effector cells to regain memory potential, but that this flexibility diminishes with continued stimulation in a polycomb-dependent manner.

Tcf7-YFP^{hi} CD8^{+} T cells formed through early and late decisions acquire a common memory program

From the above experiments, we find that naive cells in our *ex vivo* system can form memory through two pathways: a "naive to memory" (NM) pathway, whereby some cells maintain *Tcf7* expression during initial antigen stimulation, and a "naive to effector to memory" (NEM) pathway, by which cells that have silenced *Tcf7* and entered an effector state can turn expression back on after stimulation removal. To determine whether *Tcf7*-YFP^{hi} cells emerging through these two pathways both have genomic and functional memory programs, we subjected them to transcriptomic, epigenomic, and cytokine secretion analysis, alongside control *in vivo* naive (CD44^{-}CD62L^{+}), memory (Mem, CD44^{+}CD62L^{+}), and *ex vivo*-generated effector (Eff) cells (Figures 4G–4M).

Despite their different *Tcf7* regulatory history, NM and NEM cells showed similar memory characteristics. They were both more similar to naive and memory *in vivo* controls compared with *ex vivo*-generated effector cells in their shared expression of memory-defining genes, though they also maintained some effector characteristics, in line with their recent stimulation (Figures 4G–4I and S4D; Table S4). Similar to memory controls, NM and NEM cells demonstrated greater tumor necrosis factor alpha (TNF-α) and IFN-γ secretion upon restimulation compared with naive cells (Figures 4J and 4K). NM and NEM cells were most similar in global chromatin accessibility to memory controls (Figures 4L, S4E, and S4F). NEM cells recovered similar *Tcf7* accessibility to NM cells (Figure 4M). At the *Ifng* locus, intermediate accessibility of NM and NEM cells between naive and effector controls suggests that both were poised for rapid recall

intensity (area × median YFP fluorescence) over time for each track (right), with the first cell division marked by a vertical dashed line. Cell borders in snapshots are colored and labeled to match their corresponding leaves in the lineage trees. Scale bar, 10 μm. Lineage trees and tracks are colored by HMM-derived promoter state, outlined in (B).

(C) Reporter intensity for all overlaid cell tracks, colored by promoter state, where n is equivalent to the number of progeny across all lineages in the dataset at the end of observation.

(D) For each track, from left to right: time of first division, time of first transition to a stable active state, time of first transition to a stable silent state (stable state ≥10 h).

(E) For all lineages combined, fraction of cells in an active promoter state over time.

(F–H) Each division of a parent cell with the *Tcf7* promoter ON was categorized as giving rise to zero, one, or two daughters that transition to an OFF state. (F) Examples of each division category. (G) The OFF/OFF fraction by ON/OFF fraction is plotted separately for each generation to distinguish concordant, independent, and asymmetric silencing mechanisms. (H) Modified Cohen's kappa test for division events in (G).

(I) Comparison of YFP$^{+/+}$ and YFP$^{+/-}$ reporters to distinguish *cis* and *trans* regulation of *Tcf7* silencing (left). YFP distributions for YFP$^{+/+}$and YFP$^{+/-}$ reporters cultured for 5 days with 0.2 ng/mL IL-12 (right). *Tcf7*-YFP^{lo} fractions are calculated from Gaussian fits to distributions.

(J) *Tcf7*-YFP^{lo} percentages as in (I), over a range of IL-12 concentrations. Mean ± SD. Statistical significance was calculated with an unpaired two-tailed t test; n.s. p = 0.05, ***p < 0.005. Individual data points are from a single experiment representative of 2 independent experiments (I and J).

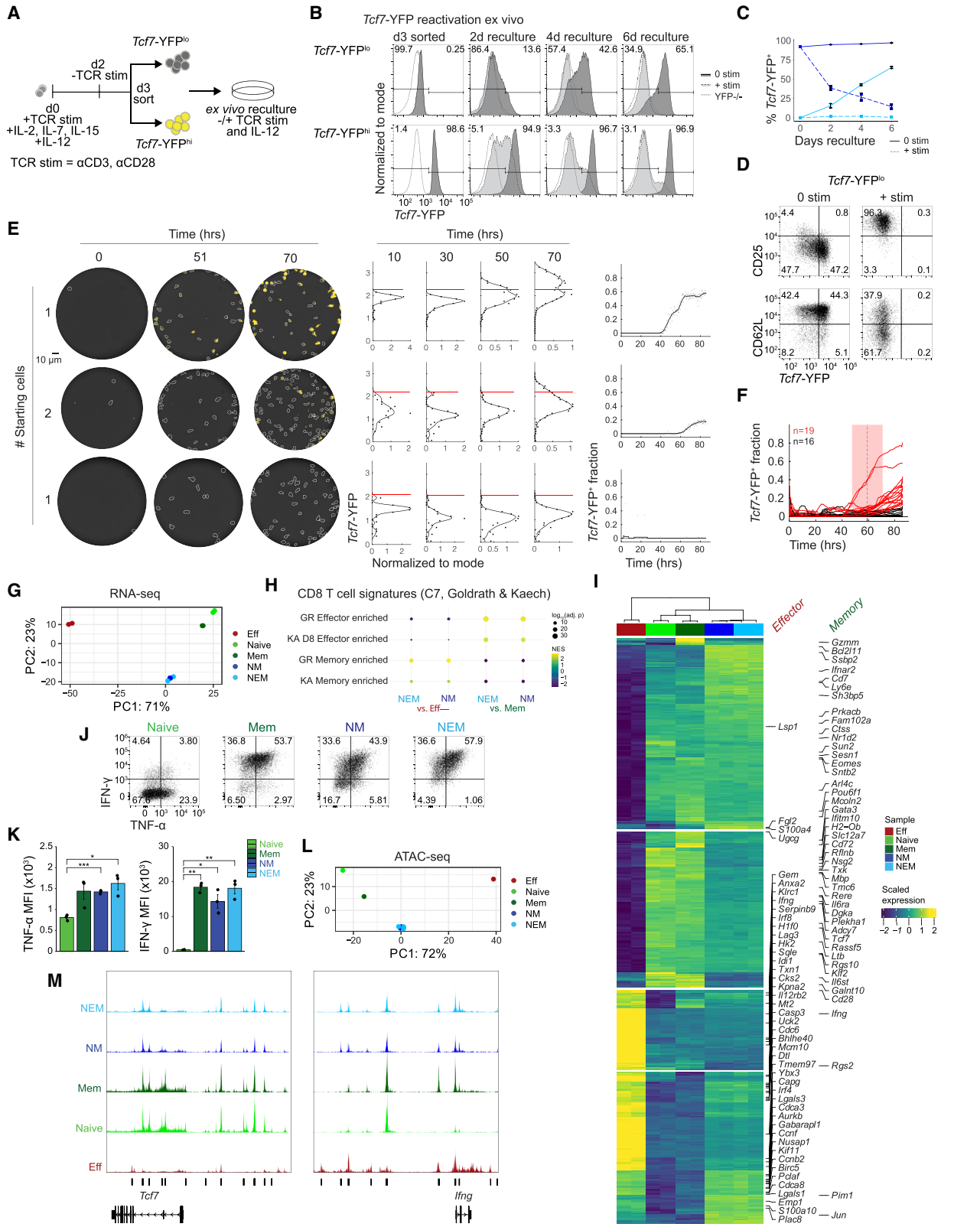


Figure 4. Effector cells reverse *Tcf7* silencing and regain memory potential upon stimulation withdrawal

(A) Naive cells from *Tcf7*-YFP mice were stimulated with αCD3 and αCD28 (TCR stim) for 2 days with IL-2, IL-7, IL-15, and IL-12. At day 2, TCR stimulation was removed but the same cytokine cocktail was used. At day 3, cells were sorted for *Tcf7*-YFP^{lo} and *Tcf7*-YFP^{hi} populations and recultured *ex vivo*, with and without

(legend continued on next page)

response, and accessibility at other memory- and effector-associated loci support this conclusion (Figure S4G).

Although NM and NEM cells were largely similar, differences in gene expression suggest that they may exhibit different functional properties. NEM cells had higher expression and accessibility of some effector-associated genes compared with NM, possibly indicative of enhanced effector capabilities or an effector memory state[39,41] (Figures S4H and S4I). Overall, both NM and NEM decision strategies gave rise to cells with genomic and functional characteristics of memory, suggesting that memory formation may proceed through a flexible decision-making strategy, allowing both for memory and effector divergence during the initial immune challenge and for effector reacquisition of memory potential after the challenge is resolved.

Memory CD8$^+$ T cells form through early and late decisions during acute infection

To determine whether the flexible decision-making strategy we characterized in our *ex vivo* system plays a role in memory formation *in vivo*, we monitored memory and effector differentiation in the lymphocytic choriomeningitis virus (LCMV) acute infection mouse model. First, to determine whether the early decision of naive T cells to silence *Tcf7* expression and enter an effector state also occurs heterogeneously *in vivo*, we transferred CellTrace Violet (CTV)-labeled *Tcf7*-YFP × P14 CD8$^+$ T cells to recipient mice that were then infected with LCMV Armstrong (LCMV-Arm) (Figure 5A), using a high viral dose (2×10^6 plaque-forming units [PFUs]) to ensure that the large starting population is completely activated (1×10^6 cells). After 3 days of infection, transferred CD8$^+$ T cells divided and bifurcated into distinct populations with high and low *Tcf7*-YFP and TCF1 protein levels (Figure 5B), consistent with *ex vivo* observations (Figure 1). PD-1 expression, an indicator of antigen exposure, was uniform across the entire population (Figure 5B), consistent with this early decision to silence *Tcf7* being controlled by stochastic epigenetic mechanisms (Figure 3I).

We next determined whether T cells that silenced *Tcf7* in response to LCMV-Arm challenge had the capacity to reactivate the *Tcf7* locus and re-express TCF1 protein to form memory upon infection clearance. To do so, we sorted *Tcf7*-YFP^{lo}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{hi} cells at 72 h post infection and re-transferred them into secondary infected recipient mice to assay their memory potential (Figure 5C). Analysis shortly after transfer (15 h) demonstrated purity of the *Tcf7*-YFP^{lo} and *Tcf7*-YFP^{hi} transferred populations, with each maintaining similar expression levels to the respective sorted population. The *Tcf7*-YFP^{mid} transferred population bifurcated into *Tcf7* high- and low-expressing cells at 15 h, consistent with it actively transitioning between *Tcf7* high and low states at the time of transfer. Upon infection clearance (day 15), *Tcf7*-YFP^{hi} cells maintained *Tcf7*/TCF1 expression (Figures 5D, 5E, and S5A–S5C). *Tcf7*-YFP^{mid} cells gave rise to an intermediate fraction of *Tcf7*/TCF1-high cells between that of the *Tcf7*-YFP^{hi} and *Tcf7*-YFP^{lo} cells, again consistent with their transitory nature. *Tcf7*-YFP^{lo} cells robustly reactivated *Tcf7*/TCF1 expression, and to similar levels as *Tcf7*-YFP^{hi} cells (Figures 5D and 5E), in line with findings in our *ex vivo* system (Figure 4). The majority of transferred *Tcf7*-YFP^{hi} cells maintained a central memory phenotype (KLRG1$^-$CD127$^+$CD62L$^+$) at day 15 (Figure 5F), as expected. *Tcf7*-YFP^{lo} cells that reactivated *Tcf7* expression also re-expressed CD127 and CD62L to acquire a central memory phenotype, indicating that these *Tcf7*-YFP^{lo} starting cells maintain the potential to form central memory cells. *Tcf7*-YFP^{lo} and *Tcf7*-YFP^{mid} sorted cells also gave rise to KLRG1$^+$CD127$^-$ cells without *Tcf7*/TCF1 expression, consistent with their giving rise to terminal effector cells (Figure 5F). In all transferred populations, PD-1 expression was initially high after transfer (at 15 h), indicating continuing antigen exposure, but dropped markedly at 15 days, consistent with antigen clearance and memory formation occurring over this duration. Taken together, these results show that memory decisions can be made both early and late during acute viral infection, as is the case in our *ex vivo* system.

Tcf7 reactivation and the late memory decision may require cell division or may occur with little or no cell division, potentially via active removal of repressive modifications at the *Tcf7* locus.[8]

continued TCR stimulation and IL-12 in the presence of IL-2, IL-7, and IL-15. Light and dark blue coloring throughout correspond to sorted *Tcf7*-YFP^{lo} and *Tcf7*-YFP^{hi} populations, respectively.

(B and C) *Tcf7*-YFP expression during reculture.

(D) CD25, CD62L, and *Tcf7*-YFP expression in *Tcf7*-YFP^{lo} cells recultured for 6 days.

(E) Representative microwells of *Tcf7*-YFP^{lo} cells recultured without stimulation: snap shots, with scale bar, 10 μm (left), top and bottom wells represent single clones; corresponding histograms (middle) with binned cell data for each time point, with YFP ± gate drawn at 2 standard deviations above the mean YFP intensity from the first 25 h; corresponding YFP$^+$ fractions over time (right).

(F) YFP$^+$ fraction for all wells overlaid. Mean activation time = 59.1 h. (C) Mean ± SD. (B–D) Data are from a single experiment representative of 1 and 3 independent experiments for +stim and 0 stim, respectively.

(G) Principal component analysis (PCA) of RNA-seq profiles (top 500 DE genes) for recultured cells compared with day 3 *ex vivo* activated effector (Eff) and day 0 (isolated from a naive mouse spleen) naive (CD44$^-$CD62L$^+$) and memory (CD44$^+$CD62L$^+$, Mem) controls. NM and NEM cells were sorted as *Tcf7*-YFP^{hi} and *Tcf7*-YFP^{lo} on day 3, respectively.

(H) Gene set enrichment analysis (GSEA) of gene signatures from MSigDB (C7), collections deposited by Goldrath (GR) and Kaech (KA) comparing recultured populations to Eff and Mem controls.

(I) Heatmap displaying top 500 DE genes ($\log_2$FC $\geq$ 1, adj. $p < 0.05$) between recultured populations and Eff, N, and Mem controls. Scale bar indicates row *Z* scores of regularized log-transformed count data. Memory and effector-associated genes from MSigDB Goldrath and Kaech collections are highlighted.

(J and K) Cytokine secretion of recultured cells compared with N and Mem controls after PMA (phorbol 12-myristate 13-acetate)/ionomcyin restimulation.

(L) PCA of assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) counts of top 500 differentially accessible peaks between recultured cells and controls.

(M) ATAC-seq read coverage tracks; vertical bars annotate differentially accessible peaks between recultured cells and controls. (G–I) n = 2 biological replicates for each sample. (K) Mean ± SEM. Statistical significance was calculated with an unpaired two-tailed t test performed between groups. *$p < 0.5$, **$p < 0.01$, ***$p < 0.001$. Data are n = 3 biological replicates from a single experiment. (L and M) n = 1 biological replicate for Eff, N, Mem; n = 2 for NM, n = 3 for NEM.

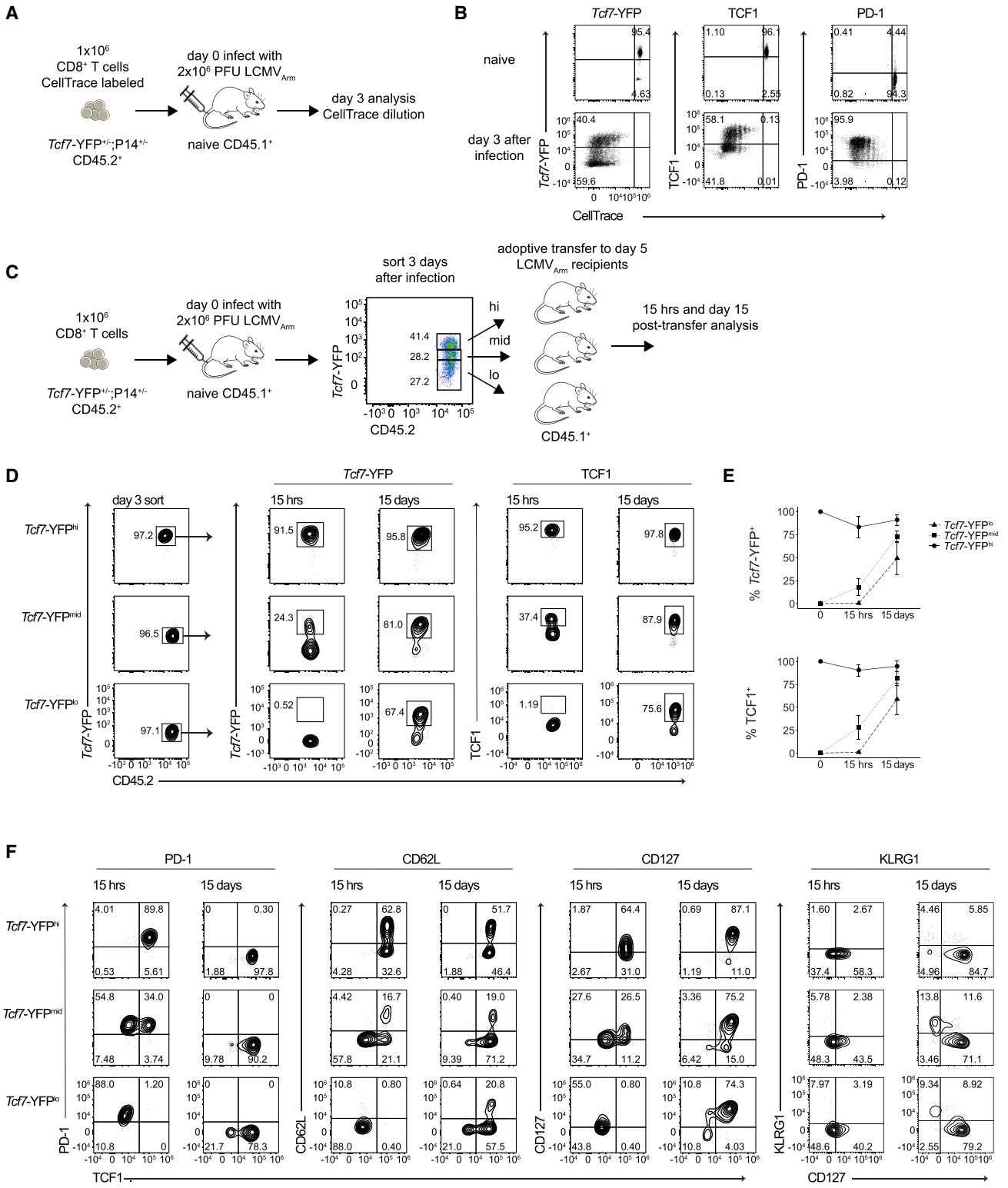


Figure 5. T cells show reversible *Tcf7* silencing and flexible memory pathway choice in acute LCMV infection

(A) Congenically marked *Tcf7*-YFPxP14 cells (1×10^6) were CellTrace Violet (CTV)-labeled and transferred into naive mice and infected with 2×10^6 PFUs of LCMV Armstrong. Cells were isolated from the recipient mice after 3 days.

(B) CTV dilution and expression of *Tcf7*-YFP, TCF1, and PD-1 on the transferred P14 CD8⁺ T cells before infection (top) and 3 days after infection (bottom).

(legend continued on next page)

To distinguish between these possibilities, we repeated these transfer experiments using sorted *Tcf7*-YFP^{lo}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{hi} populations additionally labeled by CTV (Figure 6A). These populations were sorted at day 5 post infection, CellTrace-labeled, and transferred into time-matched secondary infected recipients and analyzed 7 days after transfer. As with cells sorted at 3 days post infection, cells sorted from 5 days post infection also had the ability to reactivate *Tcf7*/TCF1 and reacquire a memory phenotype (KLRG1$^-$CD127$^+$CD62L$^+$) (Figures 6B and 6C), indicating that this flexibility in reactivation potential extends beyond the first few days of infection. *Tcf7* reactivation occurred selectively in a cell population that underwent minimal to no cell division during the 7 days after transfer (Figures 6B, right and S6), with the fast-dividing population more strongly upholding *Tcf7* silencing and effector phenotypes (CD127$^-$KLRG1$^+$) (Figure S6). These results rule out the possibility that this population arose from rare *Tcf7* expressing cells that expanded in number after transfer and additionally show that *Tcf7*/TCF1 reactivation can occur with minimal cell division upon infection clearance.

Multiple paths to memory enable robust encoding of pathogen experience through memory population size

An ability for T cells to make memory lineage decisions at multiple junctures during an infection may be important for robust protective immunity against diverse threats. In particular, it may enable the memory population sizes to scale with infection severity and immune response magnitudes, as observed.[1,42] To test this idea, we used mathematical modeling to evaluate different T cell decision-making strategies in their memory outcomes in response to pathogens of different virulence, modeled as having different rates of replication (see STAR Methods). In our first model, we considered the flexible strategy we observed (Figure 7A). Here, naive T cells (T_n) initially transition to a *Tcf7*-expressing, memory-competent state (MC, T_m) that divides upon exposure to pathogen (v) but stops dividing and persists upon pathogen clearance. These cells can either maintain memory competence upon continuing stimulation or transition to the *Tcf7*-silent effector state (T_e), where they control pathogen growth but are short-lived. Effector cells can reverse *Tcf7* silencing and re-enter the memory-competent state in the absence of pathogen, as observed (Figures 4, 5, and 6). Based on our findings that both *Tcf7* silencing and reactivation occur with probabilities that depend on whether or not stimulatory signals are present (Figures 3 and 4), we modeled transitions between effector and memory states using first order transition rates that vary with pathogen abundance, with the assumption that increasing pathogen levels result in greater antigen and inflammatory stimulation.

Mathematical simulations of this flexible decision model recapitulated the canonical features of the T cell response to acute infection (Figures 7B, S7A, and S7B). T cells expanded rapidly in response to pathogen, reaching a peak 4–8 days after infection onset that consisted mostly of effector cells, followed by a contraction to a stable, lower level of memory-competent cells (T_m). Consistent with known studies,[1,42] the quantity of memory cells was ~5% of the peak cell number. In response to pathogens with varying replication rates, this flexible decision model allowed memory cells to form robustly and scale linearly with peak cell expansion numbers. Increasing effector expansion with faster pathogen replication was accompanied by a proportional increase in memory cells, such that the memory fraction relative to the peak T cell number remained constant (Figures 7B and 7C, top; yellow shading, $\gamma_v > 0.02$/h). This relation is given by:

$$f_{T_m} = \frac{\beta_{e,m}}{\beta_{e,m}+\delta_e}$$

where $\beta_{e,m}$ is the maximum effector to memory conversion rate and δ_e is the effector death rate.

This scaling breaks down when pathogen replication is slow ($\gamma_v < 0.02$/h): reduced antigen encounter decreases the probability of the early effector cell decision, such that the number of memory cells generated converges to the starting naive cell number rather than increasing with pathogen replication rate. This ensures a baseline level of memory amid weak challenges that do not elicit a full effector response.[3]

To ask whether flexibility is necessary for scalable memory encoding, we analyzed two alternative decision models, where memory decisions are made at only one juncture. The early decision model, where naive cells irreversibly commit to the *Tcf7*-silent effector state, generated robust memory upon challenge with slow-dividing pathogens but could not reproduce the linear scaling of the memory population to the peak population in response to faster-replicating pathogens (Figures 7C middle and S7C–S7F; see STAR Methods). Conversely, the late decision model, where naive cells transition obligatorily to the effector state and decide later whether to regain memory competence, generated constant memory fractions upon stronger challenges but attenuated memory populations in response to weaker challenges (Figures 7C bottom, S7G, and S7H). These analyses underscore the importance of flexibility in memory decision making for optimal long-term immunity against variable threats.

DISCUSSION

Our finding that reversible epigenetic silencing of *Tcf7* generates inherent flexibility in the T cell memory decision reconciles two

(C) Congenically distinct (CD45.2) naive *Tcf7*-YFP P14 cells (1 × 10^6) were transferred into CD45.1 naive mice that were then infected with 2 × 10^6 PFUs of LCMV Armstrong. The activated and proliferating P14 cells were sorted 3 days post infection into *Tcf7*-YFP^{hi}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{lo} subsets and transferred into recipient mice that had been infected with 2 × 10^5 PFUs of LCMV Armstrong 5 days earlier.

(D) Expression of *Tcf7*-YFP and TCF1 on the transferred *Tcf7*-YFP^{hi}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{lo} P14 CD8$^+$ T cell subsets isolated from the spleens of recipient mice at 15 h and 15 days post-transfer.

(E) Mean frequencies of the transferred P14 CD8$^+$ T cell populations expressing *Tcf7*-YFP (top) and TCF1 (bottom) prior to transfer and at 15 h and 15 days post-transfer.

(F) TCF1, PD-1, CD62L, CD127, and KLRG1 expression for transferred populations 15 h and day 15 post-transfer. (B) Data are representative of n = 2 naive mice and n = 5 day-3 mice. (D–F) n = 2 mice for 0 and 15 h and n = 6 mice for 15 days after transfer.

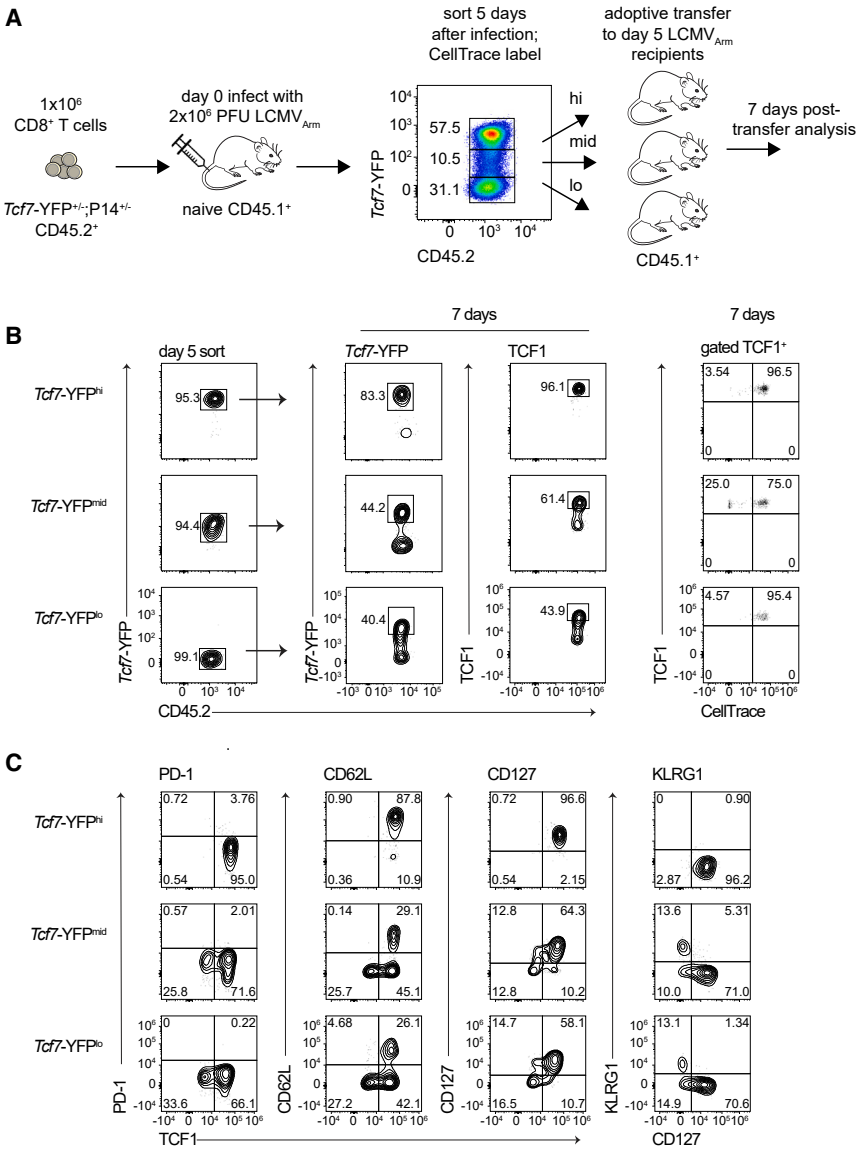


Figure 6. Reversal of *Tcf7* silencing occurs in the absence of division *in vivo*

(A) *Tcf7*-YFP P14 cells (1 × 10^6) were transferred into naive mice and infected with 2 × 10^6 PFUs LCMV Armstrong. After 5 days of infection, P14 cells were sorted into *Tcf7*-YFP^{hi}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{lo} subsets. These subsets were CTV labeled and transferred into time-matched secondary infected recipient mice as in Figure 5.

(B) *Tcf7*-YFP and TCF1 expression on the transferred cells isolated from the spleens of the recipient mice at 7 days post-transfer (middle). CTV dilution on the transferred cells expressing TCF1 (right).

(C) TCF1, PD-1, CD62L, CD127, and KLRG1 expression for transferred *Tcf7*-YFP^{hi}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{lo} cells 7 days post-transfer. (B and C) n = 3 mice per group.

prevailing models for memory development that have often been regarded as mutually opposed. Although there is evidence that memory cells can form both directly from naive cells with little or no effector differentiation and from effector cells that dedifferentiate upon infection clearance,[8,9] both in mice and in humans,[43] no model has explained how both pathways can coexist. In this mechanism, stochastic control of *Tcf7* silencing enables early divergent memory and effector decision making, and its reversibility enables late effector dedifferentiation. Antigen and inflammatory signals tune the decision-making probabilities at both junctures and would thereby influence which pathway would predominate across challenges that differ in signal duration and intensity.[44] This study, together with others,[35] implicates stochastic epigenetic switches as drivers of cellular diversification in the immune system. Through regulatory events that initiate over timescales spanning cell generations, these switches allow multiple cell populations to emerge in defined numbers without strict spatially organized cues,[45] facilitating division of labor for optimal pathogen defense.

Our modeling results lay the groundwork for understanding how the adaptive immune system can encode information about the nature and severity of a pathogen in its memory cell population. In future work, it will be interesting to determine whether other pathogen features, such as antigenicity or latency, may also be encoded quantitatively. Our findings that memory cells emerging from different decision points may differ in their functional and phenotypic properties raise the possibility that flexible decision making could underlie qualitative encoding of pathogen information through the generation of heterogeneous memory subsets.[39,41] In future work, it will be interesting to investigate the extent to which each decision pathway is utilized under various threats *in vivo* and whether cells emerging from different pathways are functionally heterogeneous.[46]

Overall, our study highlights the utility of plasticity in cell fate decision making in biological systems. From a social and cognitive sciences perspective,[47,48] flexibility allows decision-making agents to adapt and mount optimal responses amid uncertain and dynamic environments. For immune cells responding to a pathogen, the flexibility to make the memory fate decision at multiple junctures may enable greater responsiveness as an immune challenge evolves. Observed plasticity in mammalian stem cell fate decision making[49,50] may similarly allow the body to rapidly adapt its regenerative output to changing physiological needs.[51] A fuller consideration of flexibility in cellular decision making, along with its mechanisms and roles, will shed light on

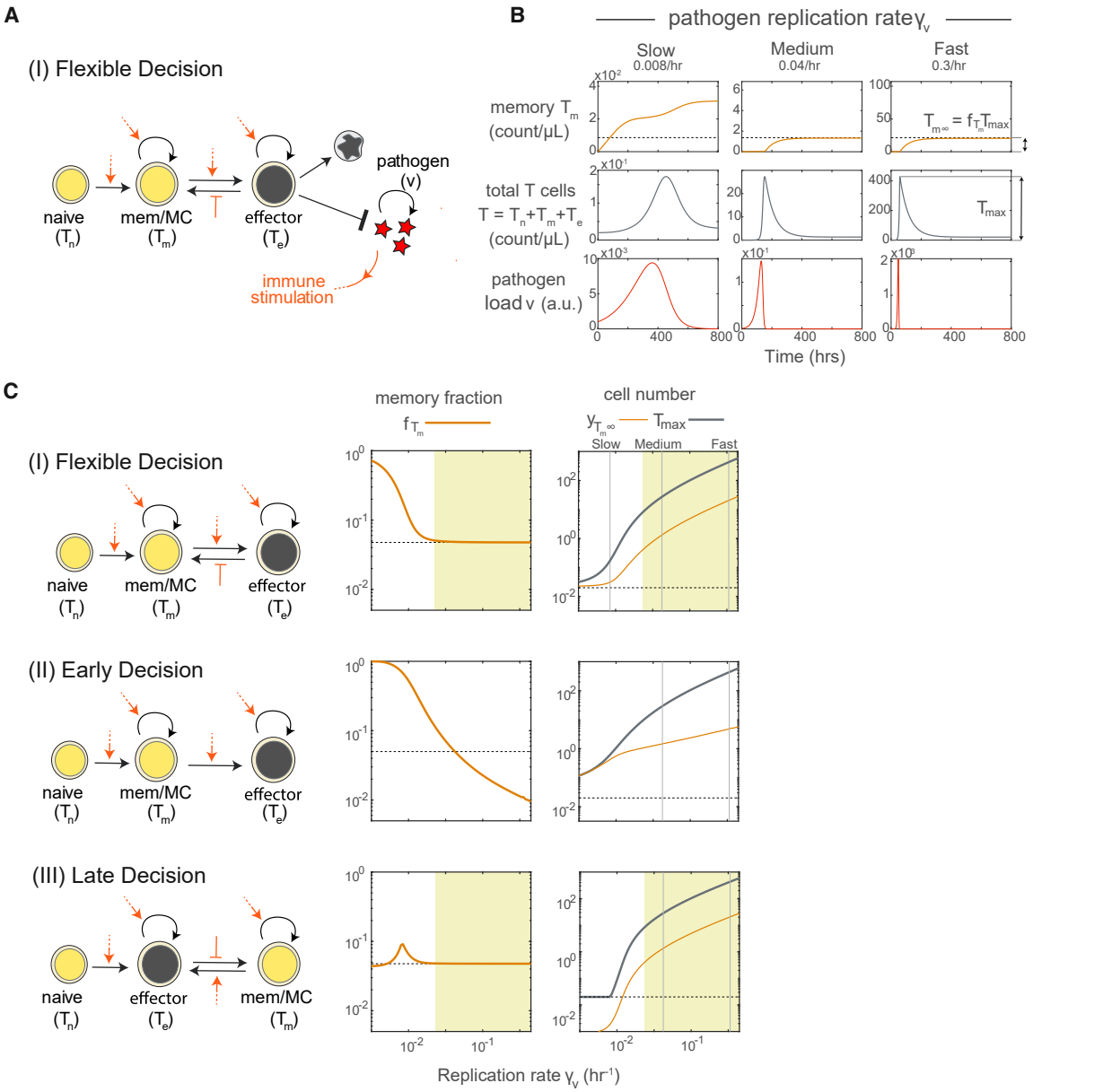


Figure 7. Flexible decision making enables quantitative encoding of pathogen experience during T cell memory formation

(A) Model incorporates pathogen proliferation and T cell memory decision making through reversible epigenetic switching. Orange arrows indicate modulation of T cell state transitions by pathogen load.

(B) Time traces show memory T cell concentrations (top), total T cell concentrations (middle), and pathogen load (bottom), for different rates of pathogen replication (left to right). Dotted line shows the number of memory T cells formed in the case when this number is a defined fraction of the peak total T cell number, f_{Tm}.

(C) Distinct strategies for memory decision making: flexible (top), early (middle), or late (bottom); the fraction of T cells at the response peak that become memory cells f_{Tm}; the peak cell number (black) and memory cell number (orange), both plotted against pathogen replication rate γ_v. The dotted line indicates the number of starting naive cells, and the yellow shading marks scalable memory.

the design principles of these systems and provide valuable insight for harnessing cells as environmentally responsive therapeutic agents.

Limitations of the study

Our *ex vivo* system intentionally minimizes environmental heterogeneity to enable the analysis of cell-intrinsic mechanisms for generating heterogeneity. Although this setup enabled us to identify stochastic epigenetic switching as a contributor to intraclonal heterogeneity, it cannot determine the relative degrees to which this cell-intrinsic mechanism and environmental factors contribute to lineage outcomes. Although our LCMV experiments show that the mechanism we identified is indeed operational during acute infection, further *in vivo* studies will be

needed to distinguish the contributions of myriad factors for heterogeneity in lineage decision making in different settings. Furthermore, future work is needed to elucidate the molecular mechanisms that mediate stochastic and reversible *Tcf7* silencing. These include investigation of TFs and *cis*-regulatory elements that control *Tcf7* silencing probabilities, as well as chromatin regulators important for maintenance of the *Tcf7*-silenced state.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice and viruses
- METHOD DETAILS
  - Naive T cell extraction
  - *Ex vivo* T cell differentiation
  - Flow cytometry analysis
  - Sample processing for sci-fate-seq
  - Computational analysis for sci-fate-seq
  - Time-lapse imaging
  - Computational analysis for time-lapse imaging
  - Analysis of *Tcf7*-YFP negative fractions in homozygous and heterozygous reporter cells
  - Sample processing for RNA-seq
  - Computational analysis for RNA-seq
  - Sample processing for ATAC-seq
  - Computational analysis for ATAC-seq
  - LCMV adoptive transfer experiments
  - Flow cytometry for LCMV adoptive transfer experiments
  - Mathematical modeling of T cell memory decision-making strategies
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.immuni.2023.12.006.

ACKNOWLEDGMENTS

We thank members of the Kueh, Nourmohammad, and Shendure labs, as well as Philip Greenberg, Jinfang Zhu, and Douglas Fowler for stimulating discussions and feedback, and William Noble for advice on statistical analysis methods. We also thank Sam Nguyen for help with brightfield image cell segmentation. We thank Anh Leith and Choli Lee in the UW Genome Sciences Flow Core and Xiaoping Wu, Ngoc-Han Nguyen, and Aurelio Silvestroni in the UW Pathology Flow Core for assistance with sorting. This study was funded by an NIH/NIBIB Trailblazer Award (R21EB027327 to H.Y.K.), an NIH/NHGRI R01 grant (R01HG010632-01 to J.S.), an NIH R01 grant (R01A1030048 to R.A.), NSF Graduate Research Fellowships (K.A. and E.C.C.), NSF CAREER award (grant no. 2045054 to A.N.), and startup funds from the Bioengineering Department at the University of Washington (H.Y.K.), the Department of Physics at the University of Washington (A.N.), and the Rockefeller University (J.C.). J.S. is an investigator at the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

K.A. and H.Y.K. conceived the study. K.A., E.C.C., and H.Y.K. designed the experiments and K.A. and E.C.C. performed the experiments and analyzed the data. R.A. and R.M.V. designed acute LCMV infection experiments and R.M.V. performed these experiments and analyzed the data. J.S. and J.C. contributed to experimental design. J.C. performed the single-cell RNA sequencing (scRNA-seq) experiments and R.M.D. performed bulk ATAC-seq. W.Y. performed initial analysis on the scRNA-seq experiments. J.F. and A.L.W. performed analysis on imaging data. K.K.H.N. helped set up the *ex vivo* T cell activation system. H.Y.K. developed the mathematical models and O.U., A.N., and H.Y.K. analyzed the mathematical models. A.B. provided the *Tcf7*-YFP reporter mice and guidance. K.A. and H.Y.K. wrote the manuscript. E.C.C., R.M.V., and R.A. contributed to the writing of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 9, 2023
Revised: October 9, 2023
Accepted: December 7, 2023
Published: January 31, 2024

REFERENCES

1. Murali-Krishna, K., Altman, J.D., Suresh, M., Sourdive, D.J., Zajac, A.J., Miller, J.D., Slansky, J., and Ahmed, R. (1998). Counting antigen-specific CD8 T cells: a reevaluation of bystander activation during viral infection. Immunity *8*, 177–187.
2. Vijh, S., and Pamer, E.G. (1997). Immunodominant and subdominant CTL responses to Listeria monocytogenes infection. J. Immunol. *158*, 3366–3371.
3. Busch, D.H., Pilip, I.M., Vijh, S., and Pamer, E.G. (1998). Coordinate regulation of complex T cell populations responding to bacterial infection. Immunity *8*, 353–362.
4. Kaech, S.M., Wherry, E.J., and Ahmed, R. (2002). Effector and memory T-cell differentiation: implications for vaccine development. Nat. Rev. Immunol. *2*, 251–262.
5. Lin, W.-H.W., Nish, S.A., Yen, B., Chen, Y.-H., Adams, W.C., Kratchmarov, R., Rothman, N.J., Bhandoola, A., Xue, H.-H., and Reiner, S.L. (2016). CD8 + T lymphocyte self-renewal during effector cell determination. Cell Rep. *17*, 1773–1782.
6. Chang, J.T., Palanivel, V.R., Kinjyo, I., Schambach, F., Intlekofer, A.M., Banerjee, A., Longworth, S.A., Vinup, K.E., Mrass, P., Oliaro, J., et al. (2007). Asymmetric T lymphocyte division in the initiation of adaptive immune responses. Science *315*, 1687–1691.
7. Kakaradov, B., Arsenio, J., Widjaja, C.E., He, Z., Aigner, S., Metz, P.J., Yu, B., Wehrens, E.J., Lopez, J., Kim, S.H., et al. (2017). Early transcriptional and epigenetic regulation of CD8+ T cell differentiation revealed by single-cell RNA sequencing. Nat. Immunol. *18*, 422–432.
8. Youngblood, B., Hale, J.S., Kissick, H.T., Ahn, E., Xu, X., Wieland, A., Araki, K., West, E.E., Ghoneim, H.E., Fan, Y., et al. (2017). Effector CD8 T cells dedifferentiate into long-lived memory cells. Nature *552*, 404–409.
9. Bannard, O., Kraman, M., and Fearon, D.T. (2009). Secondary replicative function of CD8+ T cells that had developed an effector phenotype. Science *323*, 505–509.
10. Jacob, J., and Baltimore, D. (1999). Modelling T-cell memory by genetic marking of memory T cells in vivo. Nature *399*, 593–597.

11. Chung, H.K., McDonald, B., and Kaech, S.M. (2021). The architectural design of CD8+ T cell responses in acute and chronic infection: parallel structures with divergent fates. J. Exp. Med. *218*, e20201730.

12. Merrell, A.J., and Stanger, B.Z. (2016). Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. Nat. Rev. Mol. Cell Biol. *17*, 413–425.

13. Sun, Q., Lee, W., Hu, H., Ogawa, T., De Leon, S., Katehis, I., Lim, C.H., Takeo, M., Cammer, M., Taketo, M.M., et al. (2023). Dedifferentiation maintains melanocyte stem cells in a dynamic niche. Nature *616*, 774–782.

14. Zhao, X., Shan, Q., and Xue, H.-H. (2022). TCF1 in T cell immunity: a broadened frontier. Nat. Rev. Immunol. *22*, 147–157.

15. Jeannet, G., Boudousquié, C., Gardiol, N., Kang, J., Huelsken, J., and Held, W. (2010). Essential role of the Wnt pathway effector Tcf-1 for the establishment of functional CD8 T cell memory. Proc. Natl. Acad. Sci. USA *107*, 9777–9782.

16. Harly, C., Kenney, D., Ren, G., Lai, B., Raabe, T., Yang, Q., Cam, M.C., Xue, H.-H., Zhao, K., and Bhandoola, A. (2019). The transcription factor TCF-1 enforces commitment to the innate lymphoid cell lineage. Nat. Immunol. *20*, 1150–1160.

17. Mescher, M.F., Curtsinger, J.M., Agarwal, P., Casey, K.A., Gerner, M., Hammerbeck, C.D., Popescu, F., and Xiao, Z. (2006). Signals required for programming effector and memory development by CD8+ T cells. Immunol. Rev. *211*, 81–92.

18. Rubinstein, M.P., Lind, N.A., Purton, J.F., Filippou, P., Best, J.A., McGhee, P.A., Surh, C.D., and Goldrath, A.W. (2008). IL-7 and IL-15 differentially regulate CD8+ T-cell subsets during contraction of the immune response. Blood *112*, 3704–3712.

19. Xiao, Z., Casey, K.A., Jameson, S.C., Curtsinger, J.M., and Mescher, M.F. (2009). Programming for CD8 T cell memory development requires IL-12 or Type I IFN. J. Immunol. *182*, 2786–2794.

20. Danilo, M., Chennupati, V., Silva, J.G., Siegert, S., and Held, W. (2018). Suppression of Tcf1 by inflammatory cytokines facilitates effector CD8 T cell differentiation. Cell Rep. *22*, 2107–2117.

21. Joshi, N.S., Cui, W., Chandele, A., Lee, H.K., Urso, D.R., Hagman, J., Gapin, L., and Kaech, S.M. (2007). Inflammation directs memory precursor and short-lived effector CD8+ T cell fates via the graded expression of T-bet transcription factor. Immunity *27*, 281–295.

22. Cao, J., Zhou, W., Steemers, F., Trapnell, C., and Shendure, J. (2020). Sci-fate characterizes the dynamics of gene expression in single cells. Nat. Biotechnol. *38*, 980–988.

23. Erhard, F., Saliba, A.-E., Lusser, A., Toussaint, C., Hennig, T., Prusty, B.K., Kirschenbaum, D., Abadie, K., Miska, E.A., Friedel, C.C., et al. (2022). Time-resolved single-cell RNA-seq using metabolic RNA labelling. Nat. Rev. Methods Primers *2*, 1–18.

24. He, B., Xing, S., Chen, C., Gao, P., Teng, L., Shan, Q., Gullicksrud, J.A., Martin, M.D., Yu, S., Harty, J.T., et al. (2016). CD8 + T cells utilize highly dynamic enhancer repertoires and regulatory circuitry in response to infections. Immunity *45*, 1341–1354.

25. Qiu, Q., Hu, P., Qiu, X., Govek, K.W., Cámara, P.G., and Wu, H. (2020). Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. Nat. Methods *17*, 991–1001.

26. Best, J.A., Blair, D.A., Knell, J., Yang, E., Mayya, V., Doedens, A., Dustin, M.L., and Goldrath, A.W. (2013). Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. Nat. Immunol. *14*, 404–412.

27. Qiu, X., Zhang, Y., Martin-Rufino, J.D., Weng, C., Hosseinzadeh, S., Yang, D., Pogson, A.N., Hein, M.Y., Hoi (Joseph) Min, K., Wang, L., et al. (2022). Mapping transcriptomic vector fields of single cells. Cell *185*, 690–711.e45.

28. Nguyen, P., Chien, S., Dai, J., Jr., Monnat, R.J., Becker, P.S., and Kueh, H.Y. (2021). Unsupervised discovery of dynamic cell phenotypic states from transmitted light movies. PLoS Comput. Biol. *17*, e1009626.

29. Verbist, K.C., Guy, C.S., Milasta, S., Liedmann, S., Kamiński, M.M., Wang, R., and Green, D.R. (2016). Metabolic maintenance of cell asymmetry following division in activated T lymphocytes. Nature *532*, 389–393.

30. Buchholz, V.R., Flossdorf, M., Hensel, I., Kretschmer, L., Weissbrich, B., Gräf, P., Verschoor, A., Schiemann, M., Höfer, T., and Busch, D.H. (2013). Disparate individual fates compose robust CD8+ T cell immunity. Science *340*, 630–635.

31. Duffy, K.R., Wellard, C.J., Markham, J.F., Zhou, J.H.S., Holmberg, R., Hawkins, E.D., Hasbold, J., Dowling, M.R., and Hodgkin, P.D. (2012). Activation-induced B cell fates are selected by intracellular stochastic competition. Science *335*, 338–341.

32. Gerlach, C., Rohr, J.C., Perié, L., van Rooij, N., van Heijst, J.W.J., Velds, A., Urbanus, J., Naik, S.H., Jacobs, H., Beltman, J.B., et al. (2013). Heterogeneous differentiation patterns of individual CD8+ T cells. Science *340*, 635–639.

33. Horton, M.B., Cheon, H., Duffy, K.R., Brown, D., Naik, S.H., Alvarado, C., Groom, J.R., Heinzel, S., and Hodgkin, P.D. (2022). Lineage tracing reveals B cell antibody class switching is stochastic, cell-autonomous, and tuneable. Immunity *55*, 1843–1855.e6.

34. Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell level. Science *351*, 720–724.

35. Pease, N.A., Nguyen, P.H.B., Woodworth, M.A., Ng, K.K.H., Irwin, B., Vaughan, J.C., and Kueh, H.Y. (2021). Tunable, division-independent control of gene activation timing by a polycomb switch. Cell Rep. *34*, 108888.

36. Lövkvist, C., Mikulski, P., Reeck, S., Hartley, M., Dean, C., and Howard, M. (2021). Hybrid protein assembly-histone modification mechanism for PRC2-based epigenetic switching and memory. eLife *10*, e66454.

37. Gray, S.M., Amezquita, R.A., Guan, T., Kleinstein, S.H., and Kaech, S.M. (2017). Polycomb repressive complex 2-mediated chromatin repression guides effector CD8 + T cell terminal differentiation and loss of multipotency. Immunity *46*, 596–608.

38. Ladle, B.H., Li, K.-P., Phillips, M.J., Pucsek, A.B., Haile, A., Powell, J.D., Jaffee, E.M., Hildeman, D.A., and Gamper, C.J. (2016). De novo DNA methylation by DNA methyltransferase 3a controls early effector CD8 + T-cell fate decisions following activation. Proc. Natl. Acad. Sci. USA *113*, 10631–10636.

39. Pais Ferreira, D., Silva, J.G., Wyss, T., Fuertes Marraco, S.A., Scarpellino, L., Charmoy, M., Maas, R., Siddiqui, I., Tang, L., Joyce, J.A., et al. (2020). Central memory CD8+ T cells derive from stem-like Tcf7hi effector cells in the absence of cytotoxic differentiation. Immunity *53*, 985–1000.e11.

40. Silva, J.G., Pais Ferreira, D., Dumez, A., Wyss, T., Veber, R., Danilo, M., Pinschewer, D.D., Charmoy, M., and Held, W. (2023). Emergence and fate of stem cell–like Tcf7+ CD8+ T cells during a primary immune response to viral infection. Sci. Immunol. *8*, eadh3113.

41. Herndler-Brandstetter, D., Ishigame, H., Shinnakasu, R., Plajer, V., Stecher, C., Zhao, J., Lietzenmayer, M., Kroehling, L., Takumi, A., Kometani, K., et al. (2018). KLRG1+ effector CD8+ T cells lose KLRG1, differentiate into all memory T cell lineages, and convey enhanced protective immunity. Immunity *48*, 716–729.e8.

42. Hou, S., Hyland, L., Ryan, K.W., Portner, A., and Doherty, P.C. (1994). Virus-specific CD8+ T-cell memory determined by clonal burst size. Nature *369*, 652–654.

43. Akondy, R.S., Fitch, M., Edupuganti, S., Yang, S., Kissick, H.T., Li, K.W., Youngblood, B.A., Abdelsamed, H.A., McGuire, D.J., Cohen, K.W., et al. (2017). Origin and differentiation of human memory CD8 T cells after vaccination. Nature *552*, 362–367.

44. Plumlee, C.R., Sheridan, B.S., Cicek, B.B., and Lefrançois, L. (2013). Environmental cues dictate the fate of individual CD8+ T cells responding to infection. Immunity *39*, 347–356.

45. Abadie, K., Pease, N.A., Wither, M.J., and Kueh, H.Y. (2019). Order by chance: origins and benefits of stochasticity in immune cell fate control. Curr. Opin. Syst. Biol. *18*, 95–103.

46. Wagner, D.E., and Klein, A.M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. Nat. Rev. Genet. *21*, 410–427.

47. Benjaafar, S., Morin, T.L., and Talavage, J.J. (1995). The strategic value of flexibility in sequential decision making. Eur. J. Oper. Res. *82*, 438–457.

48. Tello-Ramos, M.C., Branch, C.L., Kozlovsky, D.Y., Pitera, A.M., and Pravosudov, V.V. (2019). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. Anim. Behav. *147*, 129–136.

49. Weinreb, C., and Klein, A.M. (2020). Lineage reconstruction from clonal correlations. Proc. Natl. Acad. Sci. USA *117*, 17041–17048.

50. Xin, T., Gonzalez, D., Rompolas, P., and Greco, V. (2018). Flexible fate determination ensures robust differentiation in the hair follicle. Nat. Cell Biol. *20*, 1361–1369.

51. Manz, M.G., and Boettcher, S. (2014). Emergency granulopoiesis. Nat. Rev. Immunol. *14*, 302–314.

52. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

53. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

54. Hastie, T., and Stuetzle, W. (1989). Principal curves. J. Am. Stat. Assoc. *84*, 502–516.

55. Jackson, C.H. (2011). Multi-State Models for Panel Data: the msm Package for R. J. Stat. Softw. *38*, 1–29.

56. Smith, J.P., Corces, M.R., Xu, J., Reuter, V.P., Chang, H.Y., and Sheffield, N.C. (2021). PEPATAC: an optimized pipeline for ATAC-seq data analysis with serial alignments. NAR Genom. Bioinform. *3*, lqab101.

57. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

58. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

59. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

60. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast Gene Set Enrichment Analysis. Preprint at bioRxiv.

61. Zhu, J., Jankovic, D., Oler, A.J., Wei, G., Sharma, S., Hu, G., Guo, L., Yagi, R., Yamane, H., Punkosdy, G., et al. (2012). The transcription factor T-bet is induced by multiple pathways and prevents an endogenous Th2 cell program during Th1 cell responses. Immunity *37*, 660–673.

62. Pircher, H., Bürki, K., Lang, R., Hengartner, H., and Zinkernagel, R.M. (1989). Tolerance induction in double specific T-cell receptor transgenic mice varies with antigen. Nature *342*, 559–561.

63. Wherry, E.J., Blattman, J.N., Murali-Krishna, K., van der Most, R., and Ahmed, R. (2003). Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. J. Virol. *77*, 4911–4927.

64. Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science *357*, 661–667.

65. Cao, J., Cusanovich, D.A., Ramani, V., Aghamirzaie, D., Pliner, H.A., Hill, A.J., Daza, R.M., McFaline-Figueroa, J.L., Packer, J.S., Christiansen, L., et al. (2018). Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science *361*, 1380–1385.

66. Arsenio, J., Kakaradov, B., Metz, P.J., Kim, S.H., Yeo, G.W., and Chang, J.T. (2014). Early specification of CD8+ T lymphocyte fates during adaptive immunity revealed by single-cell gene-expression analyses. Nat. Immunol. *15*, 365–372.

67. Kurd, N.S., He, Z., Louis, T.L., Milner, J.J., Omilusik, K.D., Jin, W., Tsai, M.S., Widjaja, C.E., Kanbar, J.N., Olvera, J.G., et al. (2020). Early precursors and molecular determinants of tissue-resident memory CD8 + T lymphocytes revealed by single-cell RNA sequencing. Sci. Immunol. *5*, eaaz6894.

68. Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427.

69. Rodriguez, A., and Laio, A. (2014). Machine learning. Clustering by fast search and find of density peaks. Science *344*, 1492–1496.

70. Kueh, H.Y., Champhekar, A., Nutt, S.L., Elowitz, M.B., and Rothenberg, E.V. (2013). Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. Science *341*, 670–673.

71. Kueh, H.Y., Yui, M.A., Ng, K.K.H., Pease, S.S., Zhang, J.A., Damle, S.S., Freedman, G., Siu, S., Bernstein, I.D., Elowitz, M.B., et al. (2016). Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. Nat. Immunol. *17*, 956–965.

72. Polonsky, M., Rimer, J., Kern-Perets, A., Zaretsky, I., Miller, S., Bornstein, C., David, E., Kopelman, N.M., Stelzer, G., Porat, Z., et al. (2018). Induction of CD4 T cell memory by local cellular collectivity. Science *360*, eaaj1853.

73. Ng, K.K., Yui, M.A., Mehta, A., Siu, S., Irwin, B., Pease, S., Hirose, S., Elowitz, M.B., Rothenberg, E.V., and Kueh, H.Y. (2018). A stochastic epigenetic switch controls the dynamics of T-cell lineage commitment. eLife *7*, e37851.

74. Ounkomol, C., Seshamani, S., Maleckar, M.M., Collman, F., and Johnson, G.R. (2018). Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. Label. Nat. Methods *15*, 917–920.

75. Lian, H., Thompson, W.A., Thurman, R., Stamatoyannopoulos, J.A., Noble, W.S., and Lawrence, C.E. (2008). Automated mapping of large-scale chromatin structure in ENCODE. Bioinformatics *24*, 1911–1916.

76. Cohen, J. (1960). A coefficient of agreement for nominal scales. Educ. Psychol. Meas. *20*, 37–46.

77. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

78. Luckey, C.J., Bhattacharya, D., Goldrath, A.W., Weissman, I.L., Benoist, C., and Mathis, D. (2006). Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. Proc. Natl. Acad. Sci. USA *103*, 3304–3309.

79. Kaech, S.M., Hemby, S., Kersh, E., and Ahmed, R. (2002). Molecular and functional profiling of memory CD8 T cell differentiation. Cell *111*, 837–851.

80. Saunders, A., Core, L.J., Sutcliffe, C., Lis, J.T., and Ashe, H.L. (2013). Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. Genes Dev. *27*, 1146–1158.

81. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

82. Cusanovich, D.A., Hill, A.J., Aghamirzaie, D., Daza, R.M., Pliner, H.A., Berletch, J.B., Filippova, G.N., Huang, X., Christiansen, L., DeWitt, W.S., et al. (2018). A single-cell atlas of in vivo mammalian chromatin accessibility. Cell *174*, 1309–1324.e18.

83. Plambeck, M., Kazeroonian, A., Loeffler, D., Kretschmer, L., Salinno, C., Schroeder, T., Busch, D.H., Flossdorf, M., and Buchholz, V.R. (2022). Heritable changes in division speed accompany the diversification of single T cell fate. Proc. Natl. Acad. Sci. USA *119*, e2116260119.

84. Fowler, A.C. (2021). Atto-foxes and other minutiae. Bull. Math. Biol. *83*, 104.

85. Alves, R., and Savageau, M.A. (2000). Extending the method of mathematically controlled comparison to include numerical comparisons. Bioinform. Oxf. Engl. *16*, 786–798.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-CD8a	ThermoFisher	Cat# 45-0081-82, RRID:AB_1107004
anti-CD8a	Biolegend	Cat# 100734, RRID:AB_2075239
anti-CD8a	BD	Cat# 563786, RRID:AB_2732919
anti-CD44	ThermoFisher	Cat# 17-0441-82, RRID:AB_469390
anti-CD44	ThermoFisher	Cat# 12-0331-82, RRID:AB_2637179
anti-CD44	BD	Cat# 612799, RRID:AB_2870126
anti-CD62L	ThermoFisher	Cat# 47-0621-82, RRID:AB_1603256
anti-CD62L	Biolegend	Cat# 104438, RRID:AB_2563058
anti-CD25	ThermoFisher	Cat# 17-0251-82, RRID:AB_469366
anti-TCF1	BD	Cat# 564217, RRID:AB_2687845
anti-TCF1	Cell Signaling Technology	Cat# 6444, RRID:AB_2797627
anti-Tbet	Biolegend	Cat# 644809 (also 644810), RRID:AB_2028583
anti-IFN-gamma	Biolegend	Cat# 505849, RRID:AB_2616698
anti-IFN-gamma	Biolegend	Cat# 505808, RRID:AB_315401
anti-CD4	Biolegend	Cat# 100414 (also 100413), RRID:AB_312699
anti-CD19	BD	Cat# 568287, RRID:AB_2916875
anti-CD19	Biolegend	Cat# 115530, RRID:AB_830707
anti-PD-1	Biolegend	Cat# 135225, RRID:AB_2563680
anti-CD45.1	BD	Cat# 748978, RRID:AB_2873378
anti-CD45.2	Biolegend	Cat# 109830, RRID:AB_1186103
anti-CD127	Biolegend	Cat# 135010, RRID:AB_1937251
anti-KLRG1	BD	Cat# 561620, RRID:AB_10895798
anti-KLRG1	BD	Cat# 562897, RRID:AB_2737875
Bacterial and virus strains		
LCMV Armstrong	in-house (Emory)	N/A
Chemicals, peptides, and recombinant proteins		
anti-CD3e antibody (*in vivo* ready)	Cytek (Tonbo)	Clone: 145-2C11
anti-CD28 antibody (*in vivo* ready)	Cytek (Tonbo)	Clone: 37.51
anti-CD11a (ultra-LEAF)	Biolegend	Clone: M17/4
RetroNectin	Takara	Cat# T100B
Recombinant Human IL-2	Peprotech	Cat# 200-02
Recombinant Human IL-7	Peprotech	Cat# 200-07
Recombinant Murine IL-15	Peprotech	Cat# 210-15
Recombinant Murine IL-12 p70	Peprotech	Cat# 210-12
Recombinant Mouse IFN-β1 (carrier-free)	Biolegend	Cat# 581302
CellTrace Violet	ThermoFisher	Cat# C34557
Zombie NIR Viability Stain	Biolegend	Cat# 423117
LiveDead Fixable Aqua	ThermoFisher	Cat# L34965
LiveDead Fixable NIR	ThermoFisher	Cat# L34976
Cell Stimulation Cocktail 500x	ThermoFisher	Cat# 00-4970-93
Protein Transport Inhibitor 500x	ThermoFisher	Cat# 00-4980-93
Cytofix/Cytoperm™ Fixation/Permeabilization Kit	BD	Cat# 554714
Foxp3 / Transcription Factor Staining Buffer Set	ThermoFisher	Cat# 00-5523-00

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Paraformaldehyde	Electron Microscopy Sciences	Cat# 15710
Tazemetostat (EZH2i)	Selleckchem	Cat# S7128
GSK126 (EZH2i)	Cayman Chem	Cat# 15415
4sU	Sigma-Aldrich	Cat# T4509-25MG
SUPERaseIn	ThermoFisher	Cat# AM2694
Iodoacetamide (IAA)	Sigma-Aldrich	Cat# I1149-5G
Trizol	ThermoFisher (Ambion)	Cat# 15596018
Superscript IV reverse transcriptase with 100 mM DTT and buffer	Invitrogen	Cat# 18090200
Magnesium acetate tetrahydrate - Mg(OAc)2	Sigma-Aldrich	Cat# M5661-50G
IGEPAL CA-630	Sigma-Aldrich	Cat# I8896-50ML
1M DTT	Sigma-Aldrich	Cat# 64563-10x.5ML
Protease Inhibitor Cocktail	Sigma-Aldrich	Cat# P8340-5ML
5% Digitonin	Thermo Fisher Scientific	Cat# BN2006
2x TD buffer	Illumina Inc.	Cat# FC-121-1031
TDE1 (Tn5 enzyme)	Illumina Inc.	Cat# FC-121-1031
NEBNext High-Fidelity 2X PCR Master Mix	New England Biolabs	Cat# M0541L
SYBR green	ThermoFisher	Cat # S7567
Critical commercial assays		
CD8a+ T Cell Isolation Kit, mouse	Miltenyi	Cat# 130-104-075
EasySep Mouse CD8+ T Cell Isolation Kit	StemCell Technologies	Cat# 19853
RNeasy micro kit	Qiagen	Cat# 74004
DNA Clean & Concentrator-5	Zymo Research	Cat # D4041
NextSeq 500 High150 cycle kit	Illumina Inc.	Cat # FC-404-2002
Deposited data		
Raw and processed data for time-resolved single-cell RNA seq (sci-fate), bulk RNA-seq, and bulk ATAC-seq	This paper	Gene Expression Omnibus: GSE237830
Experimental models: Organisms/strains		
C57BL/6J (WT)	The Jackson Laboratory	Strain #: 000664
Tcf7-YFP reporter strain	Harly et al.[16]	N/A
B6;D2-Tg(TcrLCMV)327Sdz/JDvsJ (P14 TCR strain)	The Jackson Laboratory	Strain #: 004694
B6.SJL-Ptprca Pepcb/BoyJ (CD45.1)	The Jackson Laboratory	Strain#: 002014
Software and algorithms		
FlowJo	BD	https://www.flowjo.com/
Python	N/A	https://www.python.org/
MATLAB	Mathworks	https://www.mathworks.com/products/matlab.html
R	R Project	https://www.r-project.org/
RStudio	Posit	https://posit.co/download/rstudio-desktop/
Monocle3 v0.2.3.0	Cao et al.[52]	https://cole-trapnell-lab.github.io/monocle3/
Dynamo v0.95.2.dev	Qui et al.[25]	https://dynamo-release.readthedocs.io/en/latest/
scVelo v.0.2.2	La Manno et al.[53]	https://github.com/theislab/scvelo
Princurve	Hastie and Stuetzle[54]	https://github.com/rcannood/princurve
Msm package v1.6.9	Jackson[55]	https://github.com/chjackson/msm
Scipy v1.5.2	N/A	https://scipy.org/
PEPATAC v0.10.3	Smith et al.[56]	https://github.com/databio/pepatac/
Bowtie2	Langmead and Salzberg[57]	https://github.com/BenLangmead/bowtie2
HOMER v4.10	Heinz et al.[58]	http://homer.ucsd.edu/homer/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DESeq2 v1.30.1	Love et al.[59]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
fgsea v1.16.0	Korotkevich et al.[60]	https://bioconductor.org/packages/release/bioc/html/fgsea.html
KaryoploteR v1.14.1	N/A	https://bioconductor.org/packages/release/bioc/html/karyoploteR.html
Rstatix v0.7.0	N/A	https://cran.r-project.org/web/packages/rstatix/index.html
Sci-fate analysis scripts	Cao et al.[22]	https://github.com/JunyueC/sci-fate_analysis
Analysis scripts and mathematical modeling simulations	This paper	Zenodo: https://zenodo.org/records/10558973
Other		
LS columns	Miltenyi	Cat# 130-042-401
HBH (flow staining buffer)	in-house	N/A
Fc Blocking buffer (2.4G2 supernatant)	in-house	N/A
Red blood cell lysis buffer	in-house	N/A
BD FACS Aria III	BD	N/A
BD FACS Aria II	BD	N/A
Attune NxT	ThermoFisher	N/A
Aurora	Cytek	N/A
NovaSeq	Illumina Inc.	N/A
TapeStation	Agilent	N/A
Leica DMi8	Leica	N/A
Glass bottom 96 well plates	MatTek	Cat# PBK96G-1.5-5-F
250 μm micromesh	Microsurfaces	Cat# MMA-0250-100-08-01
Nanodrop 2000c	ThermoFisher	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Hao Yuan Kueh (kueh@uw.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Time-resolved single-cell RNA-seq and bulk RNA-seq and ATAC-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Microscopy data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited on Github and Zenodo and is publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice and viruses

Tcf7-YFP mice have been described.[16] We note that a small number of experiments utilized mice harboring an additional non-perturbing *Tbx21*-CFP BAC transgene reporter allele,[61] though this reporter was not further analyzed for this study. All mice used in experiments were heterozygous for the *Tcf7*-YFP reporter except where specified. WT C57BL/6 mice (Jackson Laboratory) were utilized as reporter negative controls, where applicable. Both male and female mice were used for *ex vivo* experiments, aged 8 to 12 weeks. Polyclonal T cells were used for all *ex vivo* experiments except where specified. For donors for adoptive transfer

experiments, homozygous *Tcf7*-YFP mice were crossed with an LCMV specific TCR transgenic strain[62] (P14) (Jackson Laboratory), and female heterozygous offspring were used.

For LCMV adoptive transfer experiments, LCMV infections were performed as previously described.[63] C57BL/6J and CD45.1 congenic female mice were purchased from the Jackson Laboratory. For LCMV infections, 6- to 8-week-old mice were either injected intravenously with 2×10^6 PFU of LCMV Armstrong or intraperitoneally with 2×10^5 PFU of LCMV Armstrong.

All mice were used in accordance with Institutional Animal Care and Use Committee guidelines for the University of Washington or the Emory University Institutional Animal Care and Use Committee.

METHOD DETAILS

Naive T cell extraction

Spleens were harvested from mice, massaged between rough glass slides to generate a single-cell suspension, and filtered through 40 μm nylon mesh into HBH (HBSS, 10 mM HEPES, 0.5% BSA, pH 7.4). Cells were spun down for 5 min at 300g, resuspended in 3 mL red blood cell (RBC) lysis buffer (150 mM NH_4Cl, 10 mM $NaHCO_3$, 1 mM EDTA) for 3–5 min, and quenched with HBH. Cells were spun down for 5 min at 300g and resuspended in HBH with 2.4G2 Fc blocking solution and incubated for 30 min on ice. Cells were counted, spun down again, and then enriched for $CD8^+$ T cells using a $CD8a^+$ T Cell Isolation Kit, mouse (Miltenyi, #130-104-075), with the volume and amount of antibodies and microbeads used scaled down to 70% of that specified by the manufacturer. One LS column was used per spleen (Miltenyi, # 130-042-401). To obtain a pure population of naive $CD8^+$ T cells, the cell suspension was stained with anti-CD8 (PerCP/Cyanine5.5, eBioscience, # 45-0081-82 or Biolegend, #100734), anti-CD44 (APC or PE, Invitrogen, #17-0441-82, or #12-0331-82), and anti-CD62L (APC/eFluor780, Invitrogen, #47-0621-82) at 1:600 antibody to cell suspension volume ratio in $30x10^6$ cell/mL HBH with Fc block for 15-30 min on ice and then sorted with a BD FACS Aria III (BD Biosciences) with assistance from the University of Washington Pathology Flow Cytometry Core Facility. The naive population was gated as $CD8^+CD44^-CD62L^+$*Tcf7*-YFP^+. Memory cells were gated as $CD8^+CD44^+CD62L^+$*Tcf7*-YFP^+. The cells were sorted into HBH and kept on ice until plating.

Ex vivo T cell differentiation

One day prior to T cell harvest and activation (day -1), plates were prepared by coating with anti-CD3e (Tonbo, #40-0031-U100), anti-CD28 (Tonbo, #40-0281-U100), RetroNectin (Takara, #T100B), and when specified, anti-CD11a (Biolegend, #101117). Unless otherwise specified, each well of a 96-well plate received 0.2 μg anti-CD3, 0.1 μg anti-CD28, 1 μg RetroNectin, and (when specified) 1 μg anti-CD11a in 50 μL of PBS. For differentiation in larger wells, these amounts were scaled up by well surface area. Plates were sealed with parafilm and incubated at 4°C overnight. On day 0, plates were allowed to come to room temperature for at least 30 min and washed 2x with PBS. Purified cells were added to wells in T cell media [85% RPMI 1640 with L-glutamine, 10% Fetal Bovine Serum, Pen-Strep-Glutamine, 20 mM HEPES, 1 mM Sodium Pyruvate, 0.1 mM NEAA, 50 μM BME] with indicated cytokine concentrations, mixed, and spun down for 1 min at 150g to ensure initial contact for all cells with the coated plate surface. Cytokines added to the media were 100 U/mL IL-2 (PeproTech, # 200-02), 0.5 ng/mL IL-7 (PeproTech, # 200-07), 50 ng/mL IL-15 (PeproTech, # 210-15), and 1 ng/mL IL-12 (PeproTech, #210-12) unless otherwise specified. Where specified, IFN-β1 (Biolegend, #581302) was added at 1000 U/mL. Where specified, Tazemetostat (3 μM, Selleckchem) or GSK126 (5 μM, Cayman Chem) were added to the culture along with 100 U/mL IL-2 in the absence of TCR stimulation. The cell seeding concentration was 0.1 - 2.5 million cells / mL unless otherwise indicated. Cells were incubated at 37°C in 5% CO_2 and split every two days by mixing, removing half of the well volume, and topping off the volume with TCM and respective cytokines. Where applicable, prior to seeding, cells were stained with 5 μM CellTrace Violet (CTV) (Invitrogen, #C34557) following the manufacturer's instructions.

Flow cytometry analysis

For timecourse analyses with cell surface protein staining, cells were spun down in round-bottom 96-well plates or 1.5 mL eppendorf tubes, resuspended in 2.4G2 blocking solution for 15-30 min on ice, stained with cell surface antibodies at 1:1200 (anti-CD8: PerCP-Cyanine5.5, eBioscience, # 45-0081-82, or Biolegend, #100734, anti-CD44: APC, Invitrogen, #17-0441-82, anti-CD62L: APC-e780, Invitrogen #47-0621-82, anti-CD25: APC, #17-0251-82), antibody to cell suspension volume ratio for an additional 15-30 min on ice, and spun down again for a final resuspension in HBH prior to acquisition. For samples that required intracellular protein staining, cells were fixed and permeabilized using Cytofix/Cytoperm Fixation and Permeabilization kit (BD #554714) according to manufacturer instructions and incubated with antibody for 30 min on ice. The TCF1 antibody (PE, BD Biosciences, # 564217) and T-bet antibody (PE, Biolegend, #644809) were used at 1:50 and 1:200, respectively. For samples that required intracellular cytokine staining, cells were restimulated for 5 h with PMA (phorbol 12-myristate 13-acetate)/Ionomycin (1x in 100 μL per sample Thermofisher, #00-4970-93) in round-bottom 96-well plates, with a protein transport inhibitor (1x ThermoFisher, #00-4980-93) added after 1 h. For cytokine secretion after sorting (for Naive, Mem, and NM/NEM) cells were stained with Zombie Near IR at a 1:1000 dilution in PBS following the manufacturer's instructions (Biolegend, #423117). Cells were then fixed, permeabilized, and stained with antibodies for cytokine and other intracellular protein antibodies as described above. All cytokine antibodies were used at 1:100 dilution in 1x BD Perm/Wash buffer (anti-IFN-γ (APC/Cyanine7 or PE, Biolegend, #505849, #505808) and anti-TNF-α (BV711 Biolegend, #506349). Data were acquired using an Attune NxT Flow Cytometer (ThermoFisher Scientific) and analyzed using FlowJo (BD) software.

Sample processing for sci-fate-seq

Naive CD8$^+$ T cells were activated *ex vivo*, as described. For this experiment, media was supplemented with 100 U/mL IL-2, 0.5 ng/mL IL-7, 50 ng/mL IL-15, and 0.05 ng/mL IL-12. The moderate level of IL-12, 0.05 ng/mL, was chosen for this experiment to produce a relatively even representation of *Tcf7* high and low cells (see Figure 1D). At days 1, 2, and 4 of activation, two subsequent sci-fate time points were taken as follows: cells were mixed and split into two wells, which had been coated with anti-CD3 and anti-CD28 at day -1 and remained in the incubator with TCM; 4sU was added to one well for a final concentration of 200 μM, and that well was harvested 2 h later. At that time, 4sU was similarly added to the second well, and that well was harvested 2 h later. After each 4sU addition, cells were mixed and spun down at 150g for 1 min. Harvested cells were prepared for sci-RNA-seq as described for the sci-fate protocol.[22] Briefly, cells were fixed with ice-cold 4% PFA for 15 min, washed and flash frozen with PBSR [PBS, pH 7.4, 0.2 mg/mL bovine serum albumin (Fisher), 1% Superasein (Thermofisher) and 10 mM dithiothreitol (DTT)]. PFA-fixed cells were thawed, washed, and treated with iodoacetamide (IAA) to attach a carboxyamidomethyl group to 4sU. Following these steps, a single-cell RNA sequencing library was prepared using the sci-RNA-seq protocol.[64,65] The library was sequenced on the Illumina NovaSeq system.

Computational analysis for sci-fate-seq

Read alignment, downstream processing, and TF module construction

Read alignment and downstream processing, linking of TFs to regulated genes, and construction of TF modules was performed as described in Cao et al.,[22] with minor modifications. Briefly, for each gene, across all cells, the correlation between mRNA levels of each expressed TF and that gene was computed using LASSO (least absolute shrinkage and selection operator) regression. We sought to comprehensively define gene programs with distinct dynamics by doing this correlation separately both using only newly synthesized transcript levels for potential target genes and using overall transcript levels, expecting that target genes with more stable transcripts would be more readily identified using newly synthesized transcripts, while less abundant, more lowly detected target genes would be more readily identified in the overall transcriptome. After filtering out the resultant covariance links with a correlation coefficient less than 0.03, we identified 2,117 putative TF - target gene covariance links using newly synthesized transcriptome levels and 9,927 using overall transcriptome levels, resulting in a total of 10,405 unique links after aggregation. These were further filtered to retain only links supported by ChiP-seq binding, motif enrichment, or predicted enhancer binding,[24] resulting in 1065 links between 51 TFs and 632 genes. Of these 1065 links, 147 were identified using the newly synthesized transcriptome levels, 649 were identified using the overall transcriptome levels, and 269 were identified by both. To calculate TF activity scores in each cell, newly synthesized unique molecular identifier (UMI) counts for all linked target genes were scaled by library size, log transformed, aggregated, and normalized. The absolute correlation coefficient was computed between all TF pairs with respect to their activity across all cells. Pairwise correlations were hierarchically clustered using the ward D2 method to identify TF modules, with the reasoning that co-regulatory TFs must be simultaneously active within the same cell.

Cell ordering, clustering, and differential gene expression analysis between clusters

We initially attempted to resolve T cell differentiation states by performing dimensionality reduction with Uniform Manifold Approximation and Projection (UMAP) on whole or new transcriptomes using all detected genes. This analysis largely separated cells by the time point at which they were sampled (Figure S2B), as previously observed,[66,67] likely a consequence of the host of other temporal changes occurring during activation apart from differentiation, such as cell cycle control and metabolic programming. To characterize T cell differentiation dynamics apart from other regulatory processes, cells were represented in UMAP space using newly synthesized reads for all genes within the T cell differentiation TF module with monocle3 (v.0.2.3.0) (reduction_method = 'UMAP', umap.n_neighbors = 15L, umap.min_dist = 0.001)[52] using the function align_cds[68] to remove effects of cell cycle phase (preprocess_method = 'PCA', alignment_group = 'Phase'). The resultant UMAP was clustered using density peak clustering,[69] which resulted in 5 main clusters (Figure 2C, U and E2(early) combined, E1(early), E1(late), E2(late), and MP(early) and MP(late) combined). To further resolve observed variable T cell differentiation marker expression within two of these clusters, *k*-means clustering was used to further divide U and E2(early) into separate states and MP(early) and MP(late) into separate states ($k = 2$ and 2.5, respectively). Cells in different cell cycle phases were relatively evenly distributed across this UMAP, with S phase representation highest in E1(early) (Figure S2D). Differential gene expression testing was performed between clusters using the monocle3 fit_models function.

RNA velocity analysis

RNA velocity analysis and visualization of velocity streamlines was performed using Dynamo (v.0.95.2.dev)[25,27] using expression matrices from the full and new transcriptome. The dataset was subsetted to include only the T cell differentiation module genes prior to analysis, but the resultant streamlines were similar when the analysis was performed with all genes. The streamline results were also similar when scVelo (v.0.2.2)[53] was used for velocity analysis (data not shown), with the full and new transcriptome used as the spliced/unspliced expression matrices, indicating that the streamline results are consistent between multiple analysis methods. The scVelo results were also similar with or without subsetting to include only the T cell differentiation module genes.

Trajectory analysis

Cells in each putative trajectory (E1, E2, MP) were ordered in pseudotime based on the point position on the principal curve estimated using the princurve package.[54] To align the precursor cells between trajectories, cells in the undifferentiated (U) cluster were set to pseudotime = 0. To identify genes that distinguish the trajectories, differentially expressed genes were identified using the monocle3 fit_models function with the model formula as the trajectory and pseudotime terms. Only resulting DEG associated with the trajectory term were selected.

Time-lapse imaging

Long-term time-lapse imaging of cultured cells, both to track *Tcf7* regulation during initial activation in naive cells and to track *Tcf7* reactivation in sorted *Tcf7*-low cells, was performed as previously described with some modifications.[70,71] Images were acquired with an inverted widefield fluorescence microscope (Leica DMi8) fit with an incubator to maintain a constant humidified environment at 37°C and 5% CO2, using a 40X dry objective. For imaging of the initial 4 days of activation (Figure 3), cells were seeded at low density (2-5k c/well) in wells of a 96-well glass bottom plate (Mattek) coated with anti-CD3, anti-CD28, anti-CD11a, and RetroNectin, as described above. For *Tcf7* reactivation imaging experiments (Figure 4), *Tcf7*-low cells were sorted on day 3 after 2 days of initial culture with anti-CD3 and anti-CD28 in media with IL-2, IL-7, IL-15, and IL-12 and one additional day of culture with anti-CD3 and anti-CD28 removed. These cells were seeded onto PDMS micromesh (250 μm hole diameter, Microsurfaces) mounted on top of a 24-well glass bottom plate (Mattek) to enable clonal tracking, as seeded cells show considerably enhanced motility in the absence of TCR stimulation. To prepare the micromesh for imaging, the surface was first coated with BSA while mounted on top of a 24-well plate overnight at 4°C and then transferred to a new glass well and coated with anti-CD11a and RetroNectin for improved adhesion but without anti-CD3 and anti-CD28. For reactivation experiments, cells were cultured in TCM with IL-2, IL-7, and IL-15, but without IL-12.

To determine if the experimental conditions required for imaging affect differentiation, we systematically compared expression of CD44, CD62L, and *Tcf7*-YFP in cells activated on glass or tissue culture plates, at high or low seeding density, and with or without presence of anti-CD11a (Figure S3M). CD44 expression was comparable across all conditions, confirming that all cells activated in all conditions. In tissue culture plates, CD62L and *Tcf7*-YFP expression were also comparable, though the *Tcf7*-YFP expression was slightly reduced at lower cell density, particularly in the condition without IL-12, consistent with previous findings that memory differentiation occurs less efficiently at lower cell densities.[72] On glass plates, the fraction of CD62L low cells was increased compared to on tissue culture plates. *Tcf7* expression was similarly low for the condition with IL-12, but the combination of low seeding density and presence of anti-CD11a on the glass plate resulted in a lower *Tcf7* distribution in the no IL-12 condition than was otherwise observed. This analysis shows that the specific conditions used for imaging do not affect overall differentiation trends but may underestimate the differences in differentiation between conditions with and without IL-12.

Computational analysis for time-lapse imaging

Image segmentation and tracking

Image pre-processing, cell segmentation, and tracking was performed in MATLAB (Mathworks, Natick, MA) using the ictrack movie analysis pipeline we described previously[71,73] (Figures S3A and S3B), modified to enable segmentation of cells from brightfield movies. To segment cells without additional fluorescent labels besides *Tcf7*-YFP, we first trained a convolutional neural network (CNN) with a U-net architecture[74] to predict fluorescence images of whole cells from brightfield images, using images of CellTrace Violet (CTV) labeled T cells as a training data set.[28] We trained separate CNNs for the images acquired in 96-well plates (Figure 3) and in microwells (Figure 4), as predictions are optimal when images for training and prediction have similar features. For each training dataset, hundreds of images of CTV-stained cells were acquired at multiple timepoints during the process of interest (e.g. initial T cell activation or culture after stimulation removal). Using the trained CNN, we then generated predicted whole-cell fluorescence images from acquired brightfield movies, which were used for cell segmentation (Figure S3B, 1.). Briefly, in the ictrack analysis pipeline, images underwent (1) correction by subtraction of uneven background signal stemming from the bottom of the glass plate or the side of the PDMS microwells (2) Gaussian blur followed by pixel value saturation to fix uneven signal intensity within the nucleus of the cell and (3) Laplacian edge detection algorithm to identify the nucleus boundary. Non-cell objects were excluded via size and shape limit exclusions. To generate clonal lineage trees, cells were tracked automatically between adjacent movie frames using the Munkres assignment algorithm, and the resulting cell tracks were manually checked for errors and to annotate cell divisions (Figure S3B, 2.).

Tcf7 promoter state assignment and analysis

To enable quantitative analysis of *Tcf7* promoter activity in clonal cell lineages, we assembled separate full tracks of total *Tcf7*-YFP fluorescence levels from the starting cell to each ending cell within a lineage tree, for all lineage trees analyzed (Figure S3B, 3.). Fluorescence levels are halved at each cell division; thus, to ensure continuity in *Tcf7*-YFP fluorescence in these tracks, we calculated for each parent-daughter cell pair an offset in *Tcf7*-YFP fluorescence, that we added to the daughter cells and their progeny, as previously implemented.[34] These 'continuized' tracks were then smoothed using MATLAB medfilt1 (N=5) and smooth (span = 80 time points, equivalent to 20 hours, method = lowess), and their first derivatives with respect to time were calculated to generate single-cell tracks of *Tcf7* promoter activity for downstream HMM analysis (Figure S3B, 4.).

Cell tracks were exported from MATLAB to R for downstream processing. *Tcf7* promoter states for each cell and time point were called from tracks of *Tcf7*-YFP level derivatives using a Hidden Markov Model (HMM), implemented with the msm Package for R (v1.6.9).[55] We initially tested four candidate HMMs with either three or four promoter states and variable constraints on the derivative ranges within each state (Figures S3C and S3D). For each model, we constrained the mean and variance in *Tcf7* promoter activities of each state by fitting Gaussian distributions to the *Tcf7*-YFP derivatives at different time windows, to reflect our observations that cells are expected to be mostly in an inactive, active, or attenuated state at different times.

We then compared the performance of these four models by calculating their log-likelihood and corresponding AIC (Akaike information criterion) scores. We also checked the quality of each model's fit to the data by assessing whether residuals of the fit follow a Gaussian distribution[75] (Figure S3E). Based on this analysis, we chose a model in which cells transition between 4 states: off (initial), low active, high active, and off (Figures 3B and S3F), and all start in the off-initial state at the beginning of the track. This four-state

model performed favorably compared to other models, likely because it better accounts for distinct distributions of promoter activity within active cells.

Using this four-state model, we assigned promoter activity states at each time point for each cell, removing potentially spurious transient promoter states by finding all promoter states lasting less than 8 hours and replacing them with the previously assigned promoter state. From these states, we then identified promoter silencing events as those involving a switch from active (high or low) to an inactive (off) state, and activation events as those involving a switch from inactive (off-initial or off) to active (high or low) states. We did not allow transitions back to the starting inactive (off-initial) state, as this state has a distinct *Tcf7* promoter activity distribution from the later silent state (off), likely reflecting the distinct noise characteristics of *Tcf7*-YFP fluorescence at different stages after activation.

For analysis of *Tcf7* silencing between sister cells, we first assigned an ending cell state to all cells in the dataset, representing the final promoter state of the cell prior to division or the end of the cell track. Cells with a tracked duration of less than 3 hours and parents with ending cell state durations of less than 10 hours were also excluded, to ensure the analysis only includes sufficiently tracked cells and durable promoter states. We then collected all division events for which the parent cell was in an ON promoter state prior to division and asked whether the daughter cell tracks ended in an ON or OFF promoter state. We thus calculated the number of division events that lead to no (ON/ON), unequal (ON/OFF), or concordant (OFF/OFF) daughter silencing and then calculated the fractions of each category in the entire dataset and within each generation. We statistically analyzed the degree of discordance in *Tcf7* silencing decisions between sister pairs by modifying Cohen's kappa statistical test for inter-rater reliability as follows: division events were categorized as concordant (ON/ON or OFF/OFF) or discordant (ON/OFF) between sisters. The modified Cohen's kappa coefficient, κ', was calculated as the observed percentage of discordant events minus the percentage of discordant events expected by chance, divided by 1 minus the percentage of discordant events expected by chance[76] (Table S2).

Analysis of *Tcf7*-YFP negative fractions in homozygous and heterozygous reporter cells

For analysis in Figures 3I, 3J, and S3L, YFP distributions were exported from FlowJo as csvs, imported to Python, and represented as histograms. The positive and negative populations were fit simultaneously as two gaussian distributions using the scipy.optimize.least_squares function (scipy v1.5.2), and the gate between YFP positive and negative populations was identified as the intersection between the gaussian curves. The silent fraction was then calculated as the sum of the histogram below the gate divided by the sum of the entire histogram. Two-tailed unpaired t tests between homozygous and heterozygous YFP silent fractions were performed using scipy.stats.

Sample processing for RNA-seq

Cells were centrifuged at 500g for 5 minutes, resuspended in 350 μL of Trizol (Ambion), mixed well, and frozen at -80°C for processing, starting from step 2 of the RNeasy micro kit (Qiagen, #74004) following the manufacturer's instructions. After processing, RNA was resuspended in RNase free water, quantified using a NanoDrop 2000c (Thermo Scientific), and shipped on dry ice to Novogene Corporation Inc. (Sacramento, CA) for library preparation and sequencing.

Computational analysis for RNA-seq

Raw FASTQ files from RNA-seq paired-end sequencing were aligned to the GRCm38/mm10 reference genome using Kallisto (v0.46.1),[77] and the resultant transcript-level abundance estimates were imported to genes by cells matrices using tximport (v1.18.0) for downstream analysis. Transcripts with low counts (<10) were removed. Differentially expressed genes were identified with DESeq2 (v1.30.1).[59] PCA plots were generated using the top 500 differentially expressed genes between NM and NEM samples and naive, memory, and effector controls. Significantly differentially expressed genes were also used for gene set enrichment analysis, performed with fgsea (v1.16.0)[60] and using gene sets from the C7 immunologic or the H Hallmark gene-sets from Molecular Signatures Database deposited by Goldrath and Kaech.[78,79]

Sample processing for ATAC-seq

After sorting, cells were centrifuged at 500g for 5 minutes then supernatant was aspirated without disturbing the pellet. The pellets were resuspended in 100 μL of ATAC freezing buffer[80] (50 mM Tris at pH 8.0, 25% glycerol, 5 mM $Mg(OAc)_2$, 0.1 mM EDTA, 5 mM DTT, 1× protease inhibitor cocktail (Roche-noEDTA tablet), 1:2,500 superasin (Ambion)), flash frozen in liquid nitrogen and stored at -80°C. On the day of processing, samples were thawed, centrifuged at 4°C 500g for 5 minutes, and washed with 100 μL of cold 1X PBS. Cells were again centrifuged and resuspended in 100 μL Omni lysis buffer[73] (RSB with 0.1% NP40, 0.1% Tween 20 and 0.01% Digitonin) and incubated on ice for 3 minutes, then quenched with 500 mL of RSB + 0.1% Tween 20. Nuclei were pelleted at 500g for 5 minutes at 4°C, resuspended in 100 μL cold PBS and counted. 50,000 nuclei were used per reaction, pelleted (500g for 5 min at 4°C), resuspended in tagmentation master mix[81] (50 μL total: 25 μL 2X TD buffer, 16.5 μL 1x DPBS, 0.5 μL 1% Digitonin, 0.5 μL 10% Tween 20, 5 μL water, 2.5 μL Tn5 enzyme), and incubated at 55°C for 30 minutes. Samples were purified using DNA Clean and Concentrate-5 (Zymo Research) and eluted in EB buffer (10 mM Tris) for amplification of tagmented DNA. PCR amplifications were performed using Illumina indexed primers and NEBNext High-Fidelity 2X PCR Master Mix. SYBR green was added to each PCR reaction to monitor amplification before it reached saturation. Samples in this study were amplified between 11-15 cycles using recommended conditions.[82] Unpurified products were run on a 6% TBE gel for quality control. PCR product/library were purified using DNA Clean and Concentrate-5 (Zymo Research) then ran on a tapestation to visualize nucleosome distribution. The libraries

were normalized to 2nM then pooled equimolar for sequencing. Pooled libraries were loaded onto a NextSeq 500 High150 cycle kit at 1.5 pM loading concentration with paired ends sequencing (read 1: 74 cycles, read 2: 74 cycles, index 1: 10 cycles, index 2: 10 cycles).

Computational analysis for ATAC-seq

Raw ATAC-seq FASTQ files from paired-end sequencing were processed and aligned to the mm10 mouse genome using the PEPATAC (v0.10.3)[56] pipeline, which uses bowtie2[57] for alignment. Unmapped, unpaired, and mitochondrial reads were removed. Following alignment, peak calling, merging across all samples, and annotation was performed using HOMER (v4.10).[58] Differentially accessible regions were identified using DESeq2. PCA plots were generated using the top 500 differentially accessible regions between recultured samples and naive, memory, and effector controls. Coverage tracks were generated from bigwig read alignment files using karyoploteR (v1.14.1).

LCMV adoptive transfer experiments

Donor mice: 1 × 10^6 CD45.2$^+$ *Tcf7*-YFP P14 cells were transferred into CD45.1$^+$ mice. These mice were infected with 2 × 10^6 PFU of LCMV Armstrong. For the transfer experiment in Figure 5, splenocytes were isolated 3 days after infection and CD8$^+$ T cells were enriched with the EasySep Mouse CD8$^+$ T Cell Isolation Kit (StemCell). For the CTV experiments in Figure 6, splenocytes were isolated 5 days after infection and subsequently labeled with CTV (ThermoFisher) according to the manufacturer's protocol. Using FACS, three populations of CD8$^+$CD45.2$^+$ T cells were isolated (*Tcf7*-YFP^{hi}, *Tcf7*-YFP^{mid}, and *Tcf7*-YFP^{lo}). Cells were reconstituted in RPMI, and 1x10^6 isolated cells from a single population were transferred i.v. to a single recipient CD45.1$^+$ C57BL/6J mouse, which was infected with 2 × 10^5 PFU LCMV Armstrong 5 days prior to transfer.

Flow cytometry for LCMV adoptive transfer experiments

Surface staining was performed by incubating cells with fluorochrome-conjugated antibodies against CD8 (BUV395 from BD Bioscience), CD4 (clone GK1.5; allophycocyanin (APC)–Cy7 from Biolegend), CD19 (clone 6D5; APC-Cy7 from Biolegend or BUV805 from BD), PD-1 (clone 29F.1A12; BV785 from Biolegend), CD44 (clone IM7; BUV737 from BD Bioscience), CD45.1 (clone A20; BUV563 from BD Bioscience), CD45.2 (clone 104; PE-Cy7 Biolegend), CD127 (clone AR7R34: PE from Biolegend) CD62L (clone MEL-14; BV605 Biolegend) and KLRG1 (clone 2F1; BV421 or APC from BD Bioscience). Cells were incubated on ice for 30 min in PBS + 2% FBS + 0.5 mM EDTA. TCF1 (clone C63D9; AF488 from Cell Signaling Technology) was stained intracellularly with the eBioscience Foxp3/Transcription Factor Fixation/Permeabilization Kit (ThermoFisher Scientific). Cell viability was determined with the LIVE/DEAD Fixable Aqua or near IR. Samples were analyzed with a Cytek-Aurora and data were analyzed with FlowJo software (BD). The TCF1 intracellular stain was analyzed in a separate panel from the *Tcf7*-YFP reporter, as permeabilization reduces the intensity of the reporter.

Mathematical modeling of T cell memory decision-making strategies

We consider a series of mathematical models that describe these T cell memory decision-making strategies in response to an acute infection: (1) memory decisions are made in a flexible manner, occurring either early or late; (2) memory decision arise early, as a result of irreversibility in the decision to become short-lived effector cells (''Early decision model'')[5–7]; and (3) memory decision occur late in effector cells, as a result of a direct conversion of naive cells to effector cells following antigen encounter (''Late decision model'').[8–10] In all models, we account for replication of pathogen, its removal by effector T cells, as well as the differentiation of naive cells into effector and memory cells following different decision-making strategies described above. Additionally, our model explicitly accounts for activity of the innate immune system, which acts both as a first-line responder and as an executor of T cell-directed effector activity. By comparing the performance of these different models using modeling, we aim to gain insight as to why certain decision-making strategies may have been functionally beneficial for pathogen defense by the immune system and hence, selected for during evolution.

Flexible decision model

Here, we describe the reversible model for T cell memory decision making (Figure 7). In this model, naive cells first transition into memory cell precursors in the presence of pathogens. These memory precursors then transition into effector cells with a probability that increases with increasing pathogen load. Both memory precursors and effector cells proliferate with a similar rate that increases with antigen level, as previously observed[83]; however, only effector cells undergo apoptosis as a result of activation-induced cell death, a reflection of their short-lived nature. In addition to T cells, we also consider the innate immune system in both its inactive and active states, and collectively model the innate immune response in these two states using two variables. T cells and innate cells then mediate pathogen killing, both independently from each other and in a cooperative manner. This model is described by the following system of equations:

$$\text{naive T cell}: \frac{dT_n}{dt} = -\alpha_n v T_n$$

$$\text{memory T cell}: \frac{dT_m}{dt} = \alpha_n v T_n + \left(\frac{v}{v+K_m}\right)\gamma_m T_m - \left(\frac{v}{v+K_{m,e}}\right)\alpha_{m,e} T_m + \left(\frac{K_{e,m}}{v+K_{e,m}}\right)\beta_{e,m} T_e$$

$$\text{effector T cell}: \frac{dT_e}{dt} = \left(\frac{v}{v+K_e}\right)\gamma_e T_e + \left(\frac{v}{v+K_{m,e}}\right)\alpha_{m,e} T_m - \left(\frac{K_{e,m}}{v+K_{e,m}}\right)\beta_{e,m} T_e - \delta_e T_e$$

$$\text{pathogen}: \frac{dv}{dt} = \left(\frac{v}{\epsilon+v}\right)^N \gamma_v v - (\delta_{v1} T_e + \delta_{v2} z^* + \delta_{v3} z^* T_e) v$$

$$\text{inactive innate cell}: \frac{dz}{dt} = -\alpha_z v z$$

$$\text{active innate cell}: \frac{dz^*}{dt} = \alpha_z v z + \left(\frac{v}{v+K_{z^*}}\right)\gamma_{z^*} z^* - \delta_{z^*} z^* \qquad \text{(Equation 1)}$$

Here, the subscripts n, m, and e denote naive, memory, and effector types, respectively, v denotes the pathogen population, and z and z^* denote the inactive and active innate immune cells, respectively. The rates $\alpha_{x,y}$ denote differentiation rates from the x to the y cell type (for $x, y \in \{n, m, e\}$), $\beta_{e,m}$ denotes the dedifferentiation rate from effector to memory (i.e., the only reversed differentiation in the model), γ_x denotes proliferation rate of cell type x, and δ_x denotes the death rate of cell type x. The parameters $K_{(\cdot)}$ denote the pathogen load for half maximal rate of a process indicated in the subscript.

A detailed description of the model variables, parameters and initial conditions are given in the two tables below. Parameters have been chosen based on the immune compartment sizes as measured in mice, as well as T cell biological parameters that we and others have measured. The initial conditions have been chosen to reflect the initial onset of an infection by a pathogen for which no prior immunological memory has been developed; specifically, antigen-specific naive cells are present at low amounts, effector and memory cells are absent and the initiating pathogen is introduced at a low initial level:

	Variable description	Initial value
T_n	naive T cell	0.02 cells/μL
T_m	memory T cell	0 cells/μL
T_e	effector T cell	0 cells/μL
v	pathogen load	10^{-3} units/μL
z	inactive innate cell	1 cell/μL
z^*	activated innate cell	0 cells/μL

We point out that this system is fundamentally a Lotka-Volterra model where immune cells are predators and pathogens are prey. However, we have modified this framework to describe the immune response in the following ways: first, we have incorporated saturation terms in the rates of pathogen-induced T cell and innate cell division (K_m, K_e), as well as T cell effector to memory differentiation ($K_{m,e}$) and dedifferentiation ($K_{e,m}$). The values for these saturation terms are further chosen to reflect the biological upper-bounds for these cellular processes. Second, we incorporate a threshold pathogen load, ε, below which the pathogen replication rate drops to zero, reflecting extinction of the pathogen. This pathogen extinction threshold ensures that this deterministic system of equations, when simulated, has a well-defined behavior and is not subject to numerical integration errors at very low pathogen loads.[84] However, we note that the dynamics of simulated response does not generally depend on the exact value of the extinction coefficient chosen.

From numerical simulations, we see that the flexible decision model reproduces the canonical dynamics of the adaptive immune response (Figures 7B, S7A, and S7B). Upon introduction to the system, pathogens increase exponentially in number, giving rise to a subsequent expansion of the T cell numbers from their initial low levels in the naive cell population. This expansion occurs concomitantly with pathogen clearance, and is followed by a decline in T cell numbers to a stable elevated baseline, reflecting the generation of long-lived memory cells that can survive following pathogen clearance. During the course of the immune response, the numbers of activated innate immune cells increases rapidly and decreases steadily for the remainder of simulation (Figures S7A and S7B). This heightened innate immune activity is critical for ensuring that pathogens clear after T cell contraction and do not rebound in number.

	Parameter description	Value
α_n	naive cell activation rate constant	5/(units/μL)/hr
γ_m	maximal memory cell proliferation rate constant	0.25/hr
K_m	pathogen load for half maximal memory cell proliferation	0.1 units/μL
$\alpha_{m,e}$	maximal effector differentiation rate constant	0.25/hr
$K_{m,e}$	pathogen load for half maximal effector differentiation	0.1 units/μL
γ_e	maximal effector cell proliferation rate constant	0.25/hr
K_e	pathogen load for half maximal effector proliferation	0.1 units/μL
$\beta_{e,m}$	maximal rate of effector dedifferentiation	$8x10^{-4}$/hr
$K_{e,m}$	pathogen load for half maximal effector dedifferentiation	$2.5x10^{-3}$ units/μL
δ_e	rate constant for effector cell death	0.016/hr
γ_v	rate of pathogen replication	0.003–0.45/hr
ε	pathogen load for extinction	10^{-4}/μL
N	sharpness of extinction effect for pathogen	100
δ_{v1}	rate constant for T cell pathogen killing	$4.5x10^{-3}$/(cells/μL)/hr
δ_{v2}	rate constant for innate cell pathogen killing	$1.5x10^{-3}$/(cells/μL)/hr
δ_{v3}	rate constant for T-cell assisted innate cell pathogen killing	$1.5x10^{-3}$/(cells/μL)$^{-2}$/hr
α_z	innate cell activation rate constant	2.5/(units/μL)/hr
γ_{z^*}	maximal activated innate cell proliferation rate	0.02/hr
K_{z^*}	pathogen load for half maximal innate cell proliferation	0.01 units/μL
δ_{z^*}	turnover rate for activated innate cell	$2x10^{-4}$/hr

How does the size of the generated memory population depend on the severity of infection in the reversible switching model? In particular, we wish to ascertain whether this system can produce memory cells in numbers that scale linearly with the peak T cell numbers during an infection, as observed experimentally.[1] To ask this question, we performed simulations of the system with different values of pathogen replication rate, γ_v, as a means to vary pathogen virulence. We found that pathogens with different replication rates gave rise to different degrees of T cell expansion and contraction, with faster-replicating pathogens generating a stronger T cell response, as expected (Figure 7B).

However, the fraction f_{T_m} of memory cells to the total number of T cells present at the expansion peak becomes a fixed number (Figures 7B and 7C: shaded area). Specifically, in the regime where there is substantial T cell expansion (> 10^2 fold relative to naive cell numbers) the memory fraction f_{T_m} remains constant for a broad range of viral replication rates γ_v, spanning an order of magnitude (Figure 7C, top, shaded area). On the other hand, for slowly growing pathogens (small γ_v), the memory fraction f_{T_m} increases, with a non-linear inverse dependence on the virulence and the average viral load accumulated during the infection (Figure 7C); in this regime, the number of memory cells depends strongly on the initial number of naive cells present. In summary, these results show that a flexible switching strategy for T cell memory generation allows for the amount of the generated T cell memory to scale with the size of the T cell response, in a way that depends on the severity of the infection. Our analytical results in the following section well recapitulate the behavior of this memory fraction, as indicated in Figure 7.

Analytical results for the flexible decision model

Consider the dynamics of memory T_m and effector T_e populations, given by Equation 1. We begin by identifying the dominant processes in different regimes of accumulated viral load v. Specifically, we compare the viral load with the half maximal loads ($K_{(.)}$'s in Equation 1), necessary for different processes. Below the terms that are relevant for the dynamics of memory and effector populations in the high viral load ($v \gg K_e, K_{m,e}$)and the low viral load ($v \ll K_{e,m}$) regimes are indicated:

$$\text{memory T cell}: \frac{dT_m}{dt} = \underbrace{\alpha_n v T_n + \left(\frac{v}{v+K_m}\right)\cdot\gamma_m T_m - \left(\frac{v}{v+K_{m,e}}\right)\cdot\alpha_{m,e}T_m}_{v \gg K_m, K_{m,e}} + \underbrace{\left(\frac{K_{e,m}}{v+K_{e,m}}\right)\cdot\beta_{e,m}T_e}_{v \ll K_{e,m}} \quad \text{(Equation 2)}$$

$$\text{effector T cell}: \frac{dT_e}{dt} = \underbrace{\left(\frac{v}{v+K_e}\right)\cdot\gamma_e T_e + \left(\frac{v}{v+K_{m,e}}\right)\cdot\alpha_{m,e}T_m}_{v \gg K_e, K_{m,e}} - \underbrace{\left(\frac{K_{e,m}}{v+K_{e,m}}\right)\cdot\beta_{e,m}T_e}_{v \ll K_{e,m}} - \delta_e T_e \quad \text{(Equation 3)}$$

We can formally integrate over the dynamical equation in Equation 2 to find a formal solution for the number of memory cells $T_m(t)$ at time t post infection:

$$\begin{aligned} T_m(t) &= T_m(0)\exp\left[\int_0^t \gamma_m - \alpha_{m,e} - \frac{\gamma_m K_m}{v(s)+K_m} + \frac{\alpha_m K_{m,e}}{v(s)+K_{m,e}}\mathrm{d}s\right] \\ &+ \left[\int_0^t \exp\left[\int_s^t \gamma_m - \alpha_{m,e} - \frac{\gamma_m K_m}{v(s)+K_m} + \frac{\alpha_m K_{m,e}}{v(s)+K_{m,e}}\mathrm{d}s\right]\right. \\ &\left.\times\left(\alpha_n v(s) T_n(s) + \frac{K_{e,m}}{v(s)+K_{e,m}}\beta_{e,m} T_e(s)\right)\mathrm{d}s\right] \end{aligned} \quad \text{(Equation 4)}$$

$$\begin{aligned} &\approx T_m(0)\exp[(\gamma_m - \alpha_{m,e})\min(t,\tau_{v\approx 0})] + \int_0^{\min(t,\tau_{v\approx 0})} \exp[(\gamma_m - \alpha_{m,e})(\min(t,\tau_{v\approx 0}) - s)]\alpha_n v(s) T_n(s)\mathrm{d}s \\ &+ H(t-\tau_{v\approx 0})\int_{\tau_{v\approx 0}}^t \beta_{e,m} T_e(s)\mathrm{d}s \end{aligned} \quad \text{(Equation 5)}$$

where $\tau_{v\approx 0}$ is the time to effectively clear the infection, and $H(t-\tau_{v\approx 0})$ is a Heaviside step function that takes value 1 for $t > \tau_{v\approx 0}$, and 0, otherwise. In arriving at Equation 5, we assumed that the typical viral load over the course of the infection is much higher than the differentiation thresholds K_m, $K_{m,e}$, K_e, and $K_{e,m}$, and thus, we approximated these processes by their maximal rates in Equation 4.

The following terms are important in determining the size of the memory pool:

1. $T_m(0)$: the initial memory size;
2. $\gamma_m - \alpha_{m,e}$: the effective growth rate of the memory pool;
3. $\tau_{v\approx 0}$: time to effectively clear the infection;
4. $\beta_{e,m}$: the transition rate from effector to memory.

Here, we are interested in an immune response to a primary infection, and therefore, we can assume that $T_m(0) = 0$. Moreover, given parameters used for the model, we can neglect the effective growth of the memory pool, i.e., $\gamma_m - \alpha_{m,e} \approx 0$. With these assumptions, the size of the memory pool from Equation 5 follows:

$$T_m(t) \approx \int_0^{\min(t,\tau_{v\approx 0})} \alpha_n v(s) T_n(s)\mathrm{d}s + H(t-\tau_{v\approx 0})\int_{\tau_{v\approx 0}}^t \beta_{e,m} T_e(s)\mathrm{d}s$$

Our goal is to estimate the asymptotic (long-term) fraction of memory to the total number of T cells (primarily effector cells) present at the expansion peak $f_{T_m} = T_m(\infty)/{T_e}^{\max}$. The asymptotic amount of memory follows:

$$T_m(t\to\infty) \approx \int_0^{\tau_{v\approx 0}} \alpha_n v(s) T_n(s)\mathrm{d}s + \int_{\tau_{v\approx 0}}^{\infty} \beta_{e,m} T_e(s)\mathrm{d}s$$

then assuming $\tau_{v\approx 0} \approx {\tau_e}^{\max}$:

$$T_m(\infty) \approx \int_0^{{\tau_e}^{\max}} \alpha_n v(s) T_n(s)\mathrm{d}s + \int_{{\tau_e}^{\max}}^{\infty} \beta_{e,m} T_e(s)\mathrm{d}s$$

$$\approx \int_0^{{\tau_e}^{\max}} \alpha_n v(s) T_n(s)\mathrm{d}s + \int_{{\tau_e}^{\max}}^{\infty} \beta_{e,m} {T_e}^{\max} e^{-(\beta_{e,m}+\delta_e)s}\mathrm{d}s$$

$$\approx \int_0^{{\tau_e}^{\max}} \alpha_n v(s) T_n(s)\mathrm{d}s + \beta_{e,m}\frac{{T_e}^{\max}}{\beta_{e,m}+\delta_e} \quad \text{(Equation 6)}$$

where we used the relation $T_e(t > {\tau_e}^{\max}) = T_e e^{-(\beta_{e,m}+\delta_e)s}$, indicating an exponential decay of effector cells after the peak of the response ($v \ll K_{e,m}$), from Equation 3.

From Equation 1, we can also formally express the size of the naive pool as,

$$T_n(t) = T_n(0)\exp\left[-\alpha_n\int_0^t v(s)\mathrm{d}s\right]$$

Therefore, the first term in the solution of Equation 6 follows,

$$\alpha_n \int_0^{\tau_e^{\max}} v(s)T_n(s)\mathrm{d}s = T_n(0)\left[1 - e^{-\alpha_n \tilde{V}}\right], \text{with } \tilde{V} = \int_0^{\tau_e^{\max}} v(r)\mathrm{d}r$$

Here, $\tilde{V}$ reflects the total amount of pathogens accumulated during the infection. Putting it all together, we find

$$T_m(\infty) \approx T_n(0)\left[1 - e^{-\alpha_n \tilde{V}}\right] + \beta_{e,m}\frac{T_e^{\max}}{\beta_{e,m}+\delta_e} \quad \text{(Equation 7)}$$

resulting in the following memory fraction,

$$f_{T_m} = \frac{T_m(\infty)}{T_e^{\max}} \approx \frac{x(0)}{T_e^{\max}}\left[1 - e^{-\alpha_n \tilde{V}}\right] + \frac{\beta_{e,m}}{\beta_{e,m}+\delta_e} \quad \text{(Equation 8)}$$

When $T_e^{\max} \gg T_n(0)$, we recover the constant memory fraction $f_{T_m} \approx \frac{\beta_{e,m}}{\beta_{e,m}+\delta_e}$.

So far we have assumed that $\gamma_m - \alpha_{m,e} \approx 0$. When this assumption does not hold, the first term in our expression for f_{T_m} becomes:

$$C_0 = \frac{1}{T_e^{\max}} \int_0^{\tau_e^{\max}} \exp[(\gamma_m - \alpha_{m,e})(\tau_e^{\max} - s)]\alpha_n v(s)T_n(s)\mathrm{d}s$$

$$= \frac{\alpha_n T_n(0)}{T_e^{\max}} \int_0^{\tau_e^{\max}} \exp\left[(\gamma_m - \alpha_{m,e})(\tau_e^{\max} - s) - \alpha_n \int_0^s v(r)\mathrm{d}r\right] v(s)\mathrm{d}s$$

$$\leq \frac{\alpha_n T_n(0)}{T_e^{\max}} \int_0^{\tau_e^{\max}} \exp[(\gamma_m - \alpha_{m,e})(\tau_e^{\max} - s)]v(s)\mathrm{d}s \quad \text{(Equation 9)}$$

with a strong dependence on $\gamma_m - \alpha_{m,e}$.

Note that the innate immune dynamics do not explicitly determine the memory fraction f_{T_m} , however, they influence the magnitude of $\tilde{V}$ and so are expected to be important in the low viral replication γ_v regime.

Early decision model

From experimental studies, it has been proposed that memory cells originate primarily from cells that have undergone little or no effector differentiation, and that memory precursors, upon silencing the memory regulator TCF1 and differentiating, are committed to becoming short-lived effectors.[5,39] Using the mathematical modeling framework developed above, we evaluate whether this irreversible effector decision strategy could also enable the asymptotic (long-term) memory T cell numbers $T_m(\infty)$ to scale linearly with the peak T cell number (primary effector cells) $T_e^{\max}$. To do so, we performed simulations of the above model, rendering effector differentiation irreversible by setting the rate of effector dedifferentiation $\beta_{e,m}$ to zero.

From our simulations, we found that the number of memory T cells emerging depends on the balance between effector differentiation and memory precursor proliferation, as determined by the rate constants $\alpha_{m,e}$and γ_m respectively. When these two rate constants are equal, the number of generated memory cells cannot exceed the initial number of naive cells (Figures S7C–S7F). This is because the memory cell population, upon emerging from the naive cell pool, cannot further change in number as proliferation is balanced exactly by differentiation. We note that this regime captures the dynamics of obligate asymmetric division, where the division of each memory precursor necessarily gives rise to a precursor and a differentiated progeny. When the rate of effector differentiation $\alpha_{m,e}$ is smaller than that of memory precursor proliferation γ_m, the number of memory cells can exceed the initial naive cell number, due to a net proliferation of this population; in this regime, the size of the memory pool grows with increasing γ_v.

To perform a mathematically comparable comparison of the flexible and irreversible switching models,[85] we chose a rate of effector differentiation for the latter model to be $\alpha_{m,e}$ = 0.12/hr, such that the fraction of memory cells generated under conditions of moderate pathogen virulence (γ_m = 0.04/hr) were equivalent for the two models with f_{T_m} = 0.05. All other parameters were kept constant. These simulations show that the irreversible switching model is unable to generate a constant fraction of memory cells amid changes in pathogen replication rates (Figure 7C middle). The memory fraction was upheld at a similar value f_{T_m} = 0.05 at moderate pathogen replication rates γ_v= 0.04/hr; however, this fraction decreased steadily with increasing γ_v, eventually approaching less than 0.01 at high pathogen replication rates (Figures 7C, middle and S7C–S7F). Indeed, from an approximate analytical solution of this system, we found that the memory cell fraction has an inherently inverse dependence on the peak T cell population size, as follows:

$$f_{T_m} \approx \frac{T_n(0)}{T_e^{\max}}\left[1 - e^{-\alpha_n \tilde{V}}\right] \quad \text{(Equation 10)}$$

This dependence cannot be offset when the fraction is much smaller than unity, which typically holds in the regime where the majority of cells generated at the height of an acute infection are those with effector function. Thus, from these results, we conclude that

the T cell decision-making strategy, where the memory precursors switch irreversibly into becoming short-lived effectors, cannot produce memory T cell numbers that scale proportionally with the peak T cell population sizes.

Late decision model (obligate reversible decision model)

An alternate strategy for memory differentiation is that, upon activation, all naive cells must first pass through an effector stage prior to the decision as whether to differentiate into memory (Figure 7, bottom). This view is supported by the evidence that cells with a history of effector gene expression can become memory cells,[9] and that cells on the road to forming memory retain chromatin signatures of the effector state, while harboring the ability to reactivate memory genes that are silenced during effector differentiation[8]; this model can be thought as an obligate reversible decision model.

To evaluate such a decision-making strategy, we alter our model above, such that naive cells, upon activation, directly transition to an effector state instead of a memory precursor state (Figures 7, bottom, S7G, and S7H). Following the model in Equation 1, the ordinary differential equations describing the T cell populations in the late decision model are modified as follows:

$$\text{naive T cell}: \frac{dT_n}{dt} = -\alpha_n v T_n$$

$$\text{memory T cell}: \frac{dT_m}{dt} = \left(\frac{v}{v+K_m}\right)\cdot\gamma_m T_m - \left(\frac{v}{v+K_{m,e}}\right)\cdot\alpha_{m,e} T_m + \left(\frac{K_e}{v+K_e}\right)\cdot\beta_{e,m} T_e$$

$$\text{effector T cell}: \frac{dT_e}{dt} = \alpha_n v T_n + \left(\frac{v}{v+K_e}\right)\cdot\gamma_e T_e + \left(\frac{v}{v+K_{m,e}}\right)\cdot\alpha_{m,e} T_m - \left(\frac{K_{e,m}}{v+K_{e,m}}\right)\cdot\beta_{e,m} T_e - \delta_e T_e$$

For direct comparison of this decision-making strategy to the model with flexible switching, we keep all the parameters unchanged. From simulations, we find this obligate reversible switching strategy can generate constant fractions of memory cells over a range of pathogen proliferation rates, but fails to generate any substantial memory when pathogens replicate slowly (small γ_v) and the ensuing immune responses are mild. When pathogens replicate rapidly and give rise to a substantial T cell expansion, memory cells form robustly at defined fraction and number, similar to the flexible switching model (Figures 7C, bottom, S7G, and S7H, center, right); however, when pathogens proliferate very slowly, such that there is minimal amount of T cell expansion, the number of formed memory cells constitutes only a small fraction of the starting naive cells. Consequently, in this regime, the ability of the immune system to respond to a secondary challenge is no longer heightened, and is likely compromised.

Analysis of late decision model

Following the analytical analyses for the flexible switching model, we can again identify the dominant terms in different regimes of viral load. By assuming that the viral load triggered by the infection is much larger than the cellular differentiation thresholds K_m, $K_{m,e}$, K_e, and $K_{e,m}$ we arrive at the following approximate expression for the size of the memory pool $T_m(t)$ at time t in the late decision model.

$$T_m(t) \approx T_m(0)\exp[(\gamma_m - \alpha_{m,e})\min(t, \tau_{v\approx 0})] + H(t - \tau_{v\approx 0})\int_{\tau_{v\approx 0}}^{t} \beta_{e,m} T_e(s)ds$$

Then for primary immune response $(T_m(0) = 0)$ we have:

$$f_{T_m} = \frac{T_m(\infty)}{T_e^{\max}} \approx \frac{\beta_{e,m}}{\beta_{e,m}+\delta_e} \qquad \text{(Equation 11)}$$

We note that this expression does not hold in the regime of slow/inefficient virus dynamics, consistent with results from simulations (Figure 7).

QUANTIFICATION AND STATISTICAL ANALYSIS

All analyses and p or adjusted p value significance are listed with each figure caption. Statistics were performed in R using the rstatix package (v0.7.0) or Python using scipy (v1.5.2).