Mixed Lineage Leukemia (Mll1) establishes global histone H3K4 trimethylation and stem like memory Cd8 T cell formation

Authors: Adam J. Getzler[1,2], Megan A. Frederick[1#], Huitian Diao[1,2$], Justin J. Milner[3%], Simon Bélanger[3‖], Shanel M. Tsuda[1,2], Dominic Albao[1,2] Jihye Kim[1], Shane Crotty[4], Ananda Goldrath[3] and Matthew E. Pipkin[1]*

Affiliations:

[1] Department of Immunology and Microbiology, UF-Scripps Biomedical Research, Jupiter, FL.

[2] The Skaggs Graduate Program in Chemical and Biological Sciences at Scripps Research, Jupiter, FL

[3] Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA

[4] Division of Vaccine Discovery, The La Jolla Institute of Allergy and Immunology, La Jolla, CA

[#]Current address: Department of Cell and Developmental Biology, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA

[$]Current address: BioNTech US, Cambridge, MA

[%]Current address: Lineberger Cancer Center, University of North Carolina, Chapel Hill, NC

[‖]Current address: Vir Biotechnology, San Francisco, CA

*Correspondence to: m.pipkin@ufl.EdU

Abstract: Memory Cd8+ T cells with stem cell-like properties (T_{SCM}) sustain long-lived adaptive immunity to intracellular pathogens and tumors. However, the chromatin regulatory factors (CRFs) and transcriptional dynamics that program their differentiation are unclear. Using an RNA interference screen of all CRFs we discovered that Mll1 was essential during initial T cell receptor (TCR) stimulation of naive Cd8+ T cells to establish long-lived memory and progenitor T_{SCM}-like cells during acute and during chronic viral infections and tumors. We demonstrate that primary TCR stimulation drives *de novo* histone H3 lysine 4 trimethylation (H3K4me3) deposition and RNA polymerase II pausing at genes whose RNA velocities explain single-cell trajectories that diversify nascent effector and memory cell states early during viral infection. Mll1 was essential for induced H3K4me3 at half of all TCR-activated transcription start sites (> 7,000 regions), especially those specific to memory cells. Our results suggest that Mll1-dependent H3K4me3 and regulation of RNA Pol II pausing govern dynamic transcriptional states that patterns T_{MEM} cell diversity.

Main Text:

Introduction

Naïve Cd8 T cells become activated in response to infections or cancer and develop into diverse effector (T_{EFF}), memory (T_{MEM}) or exhausted (T_{EX}) cell phylogenies(*1, 2*). Successful adaptive immunity forms following acutely resolved infections and is maintained long-term by memory T stem cells (T_{SCM}) that self-renew under homeostatic

conditions, and that regenerate both T_{EFF} and T_{MEM} cell populations following antigen re-exposure (*3, 4*). An analogous subset of progenitor T_{SCM}-like cells develops during chronic infections and cancer which self-renew and differentiate into transitional T_{EFF}-like cells that curb pathogen or tumor loads, but progressively lose function during persistent antigen stimulation and become "exhausted"(*5-7*). These exhausted progenitor T_{SCM}-like (T_{EX}^{prog}) cells respond to therapeutic antibodies that block co-inhibitory receptors expressed on T_{EX} cells, which can reinvigorate T cell responses and promote pathogen or tumor clearance in some patients(*6, 8*). The ontological basis of how T_{SCM} or T_{EX}^{prog} cells initially form from naïve Cd8 T cells is unknown, although initial T cell receptor (TCR) stimulation of naive Cd8 T cells induces transcriptionally heterogeneous cell states that rapidly develop differential potentials for developing T_{SCM} states, during either acute or chronic viral infection(*9-11*). Moreover, the chromatin regulatory factors (CRFs) and transcriptional mechanics that establish the early transcriptional variability of TCR-primed cells and stabilizes their future differentiated states are undefined.

Results

Identification of CRFs that coordinately regulate expression of Pd-1, LAG3, TIM3 and Id3.

In the first 3 days of lymphocytic choriomeningitis virus (LCMV) infection, activated Cd8 T cells that more highly express the transcriptional regulators Id3, Tcf1 (encoded by *Tcf7*) and Bach2 have an enhanced capacity to develop into T_{SCM} or T_{EX}^{prog} cells, during acutely resolved or chronic infections, respectively (*9, 10, 12*). In response to either infection, TCR activated naive Cd8 T cells initially induce uniform cell surface expression of the co-inhibitory receptors Pd-1 (*Pdcd1*), Lag3 (*Lag3*) and Tim3 (*Havcr2*) (*13*). Their expression is transient during acute viral infection and might be important for establishing protective T_{MEM} cell populations, but their expression is sustained during chronic infection, and contributes to development of T_{EX} cells (*9{Ahn, 2018 #400, 10, 12*). We hypothesized that CRFs which coordinately regulated expression of the T_{SCM}- and T_{EX}^{prog}-marker Id3, and the T_{EX} cell markers Pd-1, Lag3 and Tim3 during initial naive Cd8 T cell stimulation, might also more generally control the initial formation of T_{SCM} and T_{EX}^{prog} cells.

To screen CRFs that control expression of these markers, a library of retrovirally expressed microRNA30-embedded short-hairpin RNAs (shRNAmirs) targeting all murine CRFs was constructed and TCR-activated naive LCMV-specific (P14) Cd8 T cells were transduced with individual shRNAmirs in separate wells(*14*)(Fig 1A) . The expression of nine cell surface receptors and a GFP reporter allele in the endogenous *Id3* locus (Id3-GFP) was monitored in shRNAmir-transduced cells, and the average loss-of-function effects on expression of these T cell differentiation markers derived from multiple, unique shRNAmirs targeting each CRF were summarized (RNAi-effects) and ranked (Fig S1A). We performed two orthogonal screens: The first examined surface receptor expression on transduced P14 cells 6 days after primary TCR-activation and culture with interleukin-2 (IL-2) under conditions that promote key aspects of differentiation into effector (IL-2^{hi}) or memory (IL-2^{lo}) cells (Fig 1B, Fig S1A); the second analyzed P14 *Id3*-GFP Cd8 T cells 3 days after primary TCR stimulation (two days after transduction) and focused on expression of *Id3*-GFP, Pd-1 and Lag3 (Fig 1B, Fig S1A-B).

Multiple genes encoding CRF subunits that operate in the same complexes exhibited analogous RNAi-effects, such that suppressing each factor either increased (positive RNAi-effect) or decreased (negative RNAi-effect) expression of Pd-1, Lag3, Tim3 or *Id3*-GFP (Fig S1C). Genes encoding the COMPASS-family of histone methyltransferase complexes, MOZ/MOF acetyltransferases, NuRD and polycomb group (PcG) proteins were significantly enriched among all CRFs with positive Pd-1 RNAi-effects, whereas genes encoding BAF complex subunits were enriched among CRFs with negative Pd-1 RNAi-effects (Fig S1B-C). In addition, CRFs with positive Pd-1 RNAi-effects positively correlated with those that exhibited positive Lag3 or Tim3 RNAi-effects, and modestly correlated with those that had negative RNAi-effects on *Id3*-GFP (Fig S1D). Thus, multiple distinct CRFs that can assemble in shared complexes might cooperatively regulate the coordinated, and reciprocal, expression of co-inhibitory receptors and *Id3*-GFP upon naïve Cd8 T cell activation.

To prioritize individual CRFs, we considered those whose depletion had both the strongest effects and coordinately increased or decreased expression of more than one co-inhibitory receptor (Fig 1B, S1B), top/bottom 15% of RNAi-effects). Depletion of Mll1 (encoded by *Kmt2a*) the catalytic subunit of Mll-COMPASS-like complexes increased expression of Pd-1 and Lag3 in IL-2^{lo} cultures, as well as, Pd-1, Lag3 and Tim3 in IL-2_{hi} cultures (Fig 1B-E and Fig S1B, S2A). Depletion of Menin (*Men1*), a key component of Mll1, and Mll2 (*Kmt2d*) -COMPASS-like complexes exhibited similar effects (Fig 1B, Fig S2A). In addition, suppression of the catalytic acetyltransferases MOZ (Myst3/ *Kat6a*) and MORF (Myst4/*Kat6b*), which each associate with Mll1-COMPASS-like complexes also increased both Pd-1 and Lag3 expression. However, depletion of Mll2, the closest paralog of Mll1, exhibited distinct effects (Fig 1B-D, Fig S1B, S2A), implying Mll1 and Mll2 each have specific functions in Cd8 T cells, perhaps reminiscent of their distinct requirements in other cell types(*15, 16*). Transduction of single-guide RNAs (sgRNAs) specific for *Mll1* into P14 Cas9-transgenic Cd8 T cells also caused upregulation of Pd-1 and Tim3 (Fig S2B). Immunoblotting confirmed the on-target activity of all Mll1-specific shRNAmirs and sgRNA/Cas9-mediated disruption of *Mll1* (Fig S2C). Thus, an Mll1-specific COMPASS-like complex might be required for controlling co-inhibitory receptor expression during TCR activation.

In the second screen, Mll1 RNAi-effects were also in the top 15% of CRFs whose depletion caused upregulation of both Pd-1 and Lag3, and that impaired Id3-GFP expression (Fig 1B, Fig S1B). In this setting, depletion of Mll2 (*Kmt2d*), Mll3 (*Kmt2c*) as well as Setd1a (*Setd1a*) or Setdb1 (*Setdb1*) each also impaired upregulation of *Id3*-GFP (Fig 1B, 1D, S1B), which suggests that multiple Mll-COMPASS-like or Set1-COMPASS complex activities might concertedly or redundantly promote *Id3* expression during initial TCR stimulation. Expression of the TF Tcf1 (encoded by *Tcf7*) correlates with and promotes development of cells that express Id3 (*12, 17*). Depletion of Mll1 in P14 *Tcf7*-GFP cells also impaired Tcf7-GFP expression compared to control-transduced cells, which suggests that reduced Id3 expression in the absence of Mll1 could involve reduced Tcf1 expression (Fig S2D). However, Mll1-deficient cells were not generically defective, because depletion of Mll1 did not impair cell division history or overall accumulation during the first week of cell culture, or the ability to induce expression of key cytokines expressed by T_{EFF} and T_{MEM} cells upon TCR-like restimulation (Fig S2E, Data not shown). Thus, Mll1 is required during TCR stimulation of naive Cd8 T cells *in vitro* to ensure expression of genes encoding transcriptional regulators that promote the formation of long-lived T_{MEM} cells during acute infection, and progenitor T_{SCM}-like cells during chronic infection and cancer, while also restraining expression of Pd-1, Lag3 and Tim3.

Mll1 is essential for development of T cell states that sustain responses in tumors and during chronic viral infection

The requirement of Mll1 for gene expression during TCR stimulation suggested it could be important for establishing antigen experienced Cd8 T cell subsets that provide the most durable responses to infection or tumors. Consistent with this, analysis of previously published bulk gene expression data from *ex vivo* isolated Cd8 T cells demonstrated that *Kmt2a* expression was high in long-lived T_{MEM} (Fig 1F, *left*) and T_{EX}^{prog} cells (Fig 1F, *right*) and clustered with the T_{SCM}-signature genes *Id3, Tcf7, Slamf6* and *Cxcr5,* which were highly expressed in these cell types. Conversely, *Kmt2a* expression was lower during the effector phase of acute infection in T_{EFF} cells (Fig 1F, *left*), and in both T_{EFF}^{trans} and T_{EX}^{term} cells during chronic infection, and therefore clustered apart from *Pdcd1* (Pd-1), *Havcr2* (Tim3) and *Lag3* (Lag3), which were highly expressed in these cell types (Fig 1F, *right*). Thus, *Mll1* expression is dynamic during T cell differentiation and positively correlates with formation of long-lived, protective T cell states during both acute and chronic infection.

To begin defining the *in vivo* requirements of Mll1, the effects of its depletion in Cd8 T cells responding to solid tumors were analyzed in mice. P14 TCR transgenic Cd8 T cells, which recognize the GP_{33-41} epitope of LCMV (GP33), were transduced with shRNAmirs specific for Mll1 (shMll1) or negative control mRNAs (shCon) that are not expressed in Cd8 T cells (shCd4 or shCd19), and were co-transferred into wild type mice engrafted with engineered B16 melanoma or E0771 triple negative breast tumors expressing GP33 fused to GFP ($B16^{GP33}$ or

E0771^{GP33}), and their parental controls expressing GFP-only (B16^{GFP} or E0771^{GFP}) (Fig S3A). In these models, P14 cells accumulated systemically post adoptive cell transfer (pACT), but preferentially expanded and upregulated Pd-1 in B16^{GP33} or E0771^{GP33} tumors, compared to spleen or E0771^{GFP} tumors (Fig S3B, S3C). Thus, the phenotypes of transferred P14 cells *in vivo* were induced specifically by antigen display in the tumor. Compared to shCon-transduced P14 cells in E0771^{GP33} and B16^{GP33} tumors, shMll1-transduced P14 cells formed significantly greater fractions of Pd-1^{hi} Tim3^{hi} cells (Fig 1G, p=0.026; Fig S3D, p=0.0081), exhibited higher Pd-1 expression among gated Tim3^{+} cells in E0771^{GP33} tumors (Fig S3E, p=0.0004), and accumulated in E0771^{GP33+} and B16^{GP33+} tumors as well as, or better, than shCon-transduced cells by day 3 pACT (Fig S3F). In addition, shMll1-transduced P14 cells initially delayed E0771^{GP33+} tumor growth compared to transfer of shCon-transduced P14 cells, but the tumors ultimately regressed analogous to recipients of shCon-transduced P14 cells (Fig S3G). Thus, Mll1-deficient cells preferentially adopt T_{EFF} cell-like states and are transiently functional in the tumor microenvironment, but do not sustain tumor control.

Normally, Tim3^{lo} tumor-specific Cd8 T cells highly express Tcf1 and Slamf6, and comprise a progenitor T_{SCM}-like population that sustains the T cell response (*18*). However, shMll1-transduced P14$^{Tcf7\text{-}GFP}$ donor cells in E0771^{GP33+} tumors on day 3 pACT that were Tim3^{lo} expressed less *Tcf7*-GFP compared to control-transduced P14$^{Tcf7\text{-}GFP}$ cells, as well as less *Tcf7*-GFP overall. (Fig S3H, p=0.0732, p=0.0212). Mll1 deficent cells also comprised significantly smaller fractions and numbers of Tim3^{lo} Slamf6^{hi} cells by day 8 pACT (Fig 1H, p<0.0001), compared to shCon-transduced cells. Altogether, these results indicate that Mll1 is necessary for the early formation and maintenance of progenitor T_{SCM}-like cells, and that Cd8 T cells lacking Mll1 preferentially develop into T_{EFF}-like cells that inefficiently persist in the tumor microenvironment.

The nature of antigen-presentation in the tumor microenvironment could differentially affect the development of Mll1-deficient cells. Therefore, Mll1-depleted P14 cells were analyzed after transfer into wildtype hosts followed by infection with LCMV Clone 13 ($LCMV_{Cl13}$), which causes systemic chronic infection, persistent antigen presentation and development of a continuum of exhausted progenitor stem-like cells (T_{EX}^{prog}), transitory effector/exhausted intermediates (T_{EX}^{trans}/T_{EX}^{int}) and terminally exhausted (T_{EX}^{term}) cells that curtail but do not eliminate the infection(*19-22*). The developmental relationships and regulatory factors that initially establish this pattern during chronic infection are still unclear. On day 5 post infection (pi), both shCon-transduced and shMll1-transduced P14 cells equivalently polarized into Tcf7^{hi} Tim3^{lo} (T_{EX}-precursor) and Tcf7^{lo} Tim3^{hi} (T_{EFF}-precursor) cells (Fig 1I, top)(*12*). However, shCon-transduced T_{EX} precursors comprised equivalent fractions of Slamf6^{hi} Cx3cr1^{lo} or Slamf6^{hi} Cx3cr1^{hi} cells, indicating that T_{EX} precursors are heterogeneous and harbor cells biased toward both T_{EX}^{trans}/T_{EX}^{int} (Cx3cr1^{hi}) and T_{EX}^{prog} (Cx3cr1^{lo}) phenotypes at early times. In contrast, shMll1-transduced Tcf7^{hi} Tim3^{lo} cells were predominantly Slamf6$^{int/lo}$ Cx3cr1^{hi}, which suggests Mll1-deficiency biases development toward the T_{EX}^{trans}/T_{EX}^{int} state before the first 8 cell divisions (Fig 1I, bottom, p=0.0392). Thus, Mll1 appeared to be essential for maintaining the T_{EX}^{prog}-like state among Tcf7^{hi} cells at the initiation of the response.

Longitudinal analysis of P14 cells also demonstrated Mll1-depleted cells could not maintain a progenitor-like state. Their numbers and fraction of total Cd8s rapidly declined as early as day 8 pi (Fig S3I). Consistent with this, shMll1-transduced cells comprised reduced fractions of Tcf7-GFP^{hi} Tim3^{lo} cells on day 15 pi compared to shCon-transduced cells, which normally constitute T_{EX}^{prog1} cells that sustain proliferation during chronic infection (Fig 1J, top, p=0.0571)(*19*). Moreover, of the shMll1-transduced cells that were Tcf7-GFP^{hi} at this time, fewer were also CD69^{hi} (Fig 1J, bottom, p=0.0004). Thus, residual Mll1-deficient Tcf7-GFP^{hi} cells were not *bona fide* T_{EX}^{prog1} cells.

In addition, shMll1-transduced cells expressed much less Slamf6 overall (Fig S3J), suggesting they preferentially developed into either T_{EX}^{trans} or T_{EX}^{int} cells. However, these appeared to be a non-canonical T_{EX}^{trans}-like state as they comprised a reduced fraction of Tcf7-GFP^{lo} Tim3^{hi} cells on day 15 pi compared to shCon-transduced cells (Fig 1J, p = 0.0001), and were uniformly CX3CR1^{hi} (Fig S3K, p=0.0230). Furthermore, shMll1-transduced cells that were Slamf6$^{int/lo}$ did not upregulate CD69, whereas half or more shCon-transduced cells that were Slamf6$^{int/lo}$

were CD69^{hi} (Fig S3J, p =0.0052). These results suggest that wild type T_{EX}^{int} cells (Slamf6$^{int/lo}$ CD69^{lo}) normally transition into T_{EX}^{term} (Slamf6$^{int/lo}$ CD69^{hi}) state as early as day 5 pi during chronic infection, and that this requires Mll1. Thus, Mll1 is essential for establishing *bona fide* T_{EX}^{prog} cells and propelling the normal trajectory of developmental states that sustain Cd8 T cells responding to chronic infection.

Mll1 is essential for establishing the normal pattern of circulating T_{MEM} cells during acute viral infection

Persistent TCR stimulation during chronic infection could exacerbate or modify the defects in Mll1-deficient cells. To examine if Mll1 was important for the formation of protective T_{MEM} cells, which normally develop when infections are resolved rapidly and antigen is completely cleared(*23, 24*), naive P14 Cd8 T cells transduced with *Mll1*-specific (shMll1) or negative control (shCd19 or shCd4) shRNAmirs were co-transferred to wildtype host mice subjected to LCMV Armstrong ($LCMV_{Arm}$) infection, which acutely resolves (*22, 25*). In this setting, Mll1-deficient cells initially accumulated, but formed 2-4x fewer T_{EFF} cells at the normal peak of the response compared to control-transduced cells (Fig 2A). Subsequently, the number of Mll1-depleted T_{EFF} cells decreased more sharply than control cells and resulted in ~10-fold fewer total T_{MEM} cells in the spleen by day 45 pi, and became undetectable in some recipients by day 90 pi (Fig 2A). Thus, Mll1-deficiency modestly reduced T_{EFF} cell numbers, but severely impaired the number of circulating T_{MEM} cells that developed after acute LCMV infection resolves.

Before acute infections resolve, the T_{EFF} cell population becomes heterogeneous and only some cells develop potential to convert into T_{MEM} cells (*3, 11, 12, 26-29*), but how this heterogeneity is initially established and contributes to T_{MEM} cell potential is still unclear. Analysis of P14 cells labelled with cell trace violet (CTV) showed that Mll1 deficent cells underwent fewer cell divisions, less efficiently sustained high Tcf7-GFP expression (p =0.0286) and expressed lower Cd25 compared to control-transduced cells on day 3 pi (Fig 2B, p= 0.0069). This indicated that the typical inverse relationship between Cd25 and Tcf7 expression did not polarize in the absence of Mll1. Furthermore, on day 5 pi, P14 cells transduced with shMll1 developed reduced fractions of Cd27^{hi} Cx3cr1$^{lo/int}$ and Cd27^{int} Cx3cr1^{int} cells, which are normally precursors of central memory (T_{CM}) cells (Fig 2C, *p<0.0001,p=0.0022*) (*30*). Instead, Mll1-deficient P14 cells were mainly Klrg1^{hi} (Fig S4A, p<0.0001) and harbored a much larger fraction of Cd27^{lo} Cx3cr1^{hi} cells (Fig 2C, p<0.0001), which normally comprise precursors of T_{EM} cells and terminally differentiated effector (TE) cells(*29-31*). Notably, shMll1-transduced cells were as abundant as shCon-transduced P14 cells in the spleen on day 5 pi (Fig 2A), but incorporated less EdU, demonstrating they were proliferating less at this time (Fig S4B, top, p=0.0004), which likely explains their reduced accumulation by the peak of the response (Fig 2A). By day 8 pi, shMll1-transduced P14 cells formed lower frequencies of Klrg1^{lo} Cd127^{hi} classical memory precursor (MP) cells (*29*)compared to shCon-transduced cells, and were virtually devoid of Klrg1^{lo} Cd127^{lo} "early effector" (EE) cells (Fig 2D, p=0.0078, p<0.0001). These phenotypes were confirmed using P14-Cas9 transgenic cells transduced with Mll1-specific single-guide RNAs (sgMll1, Fig S4C, *right, p=0.2507,p=0.0216*). Together, these results suggest that Mll1 is necessary for the early polarization of distinct precursor populations that normally give rise to recirculating T_{CM} and T_{EM} cells, and preferentially restrains formation of TE-like cells before the peak response to $LCMV_{Arm}$ infection.

Subsequently, Mll1-deficiency distorted conversion of T_{EFF} cells into distinct T_{MEM} subsets as the infection resolved. By day 15 pi, shMll1-deficient cells generated a modestly increased fraction of Cd27$^{int/lo}$ Cx3cr1^{hi} T_{EM} phenotypic cells (Fig 2E, p=0.0313), and reduced fractions of Cd27^{hi} Cx3cr1^{lo} or Cd27^{hi} Cx3cr1^{int} cells, which both convert into T_{CM} cells(*30*) (Fig 2E, p=0.0309, p=0.0075). However, shMll1-transduced P14 cells that were Klrg1^{hi} on day 8 pi incorporated less EdU (Fig S4D, p=0.0146), indicating they were less proliferative than shCon transduced cells as they entered the typical period (day 7 to day 30 pi) when T_{EFF} cells die or convert into T_{MEM} cells. This suggests Mll1 is essential for generating or sustaining an underlying population of T_{EFF} cells that normally become effector-like T_{EM} cells that predominate during the early memory phase (*32*)despite the modest skewing of shMll1-deficient cells toward the Cd27$^{int/lo}$ Cx3cr1^{hi} T_{EM} phenotype on day 15 pi(Fig 2E, p=0.0313). In line with this observation, over time, the number and frequency of shMll1-transduced P14 cells that exhibited the effector-like terminal T_{EM} (t-T_{EM}) cell (Cd62l^{lo} Cd127^{lo}) (Fig 2F and Fig S4), or T_{EM} cell (Cd62l^{lo} Cd127^{hi})

phenotypes decreased, while the frequency of those with a central memory (T_{CM}) phenotype (Cd62l^{hi} Cd127^{hi}) increased, even though the overall numbers of each T_{MEM} subpopulation were 10-fold or more reduced compared to shCon-transduced cells (Fig 2E-F, S2E-F). Thus, Mll1 is required to create the normal pattern of T_{EM} and T_{CM} cell formation.

Normally, high *Tcf7*/Tcf1 expression underlies the fitness of long-lived T_{CM} cells and their ability to undergo extensive proliferation following antigen re-exposure during secondary infection(*33*). Analysis of shMll1-transduced P14$^{Tcf7\text{-}GFP}$ cells demonstrated that both T_{EM} and T_{CM} phenotypic populations on day 15 pi expressed less Tcf7-GFP compared to shCon-transduced cells (Fig 2G). Thus, reduced formation of both T_{EM} and T_{CM} phenotypic cells derived from Mll1-deficient T_{EFF} cells correlated with reduced Tcf7 expression. To determine whether residual T_{CM}-phenotype cells derived from shMll1-transduced P14 cells were functional, Cd62l^{hi} and Cd62l^{o} shCon-transduced and shMll1-transduced P14 cells were isolated 90 days after primary LCMV$_{Arm}$ infection, and 1:1 mixtures of control and Mll1-depleted Cd62l^{hi} or Cd62l^{o} cells were transferred to new naive hosts and "recalled" by infection with *Listeria monocytogenes*-GP33 (LM$_{GP33}$) (Fig 2H and Fig S2F). On day 5 after LM$_{GP33}$ infection, Cd62l^{hi} donor P14 cells that were shMll1-transduced accumulated 10-fold less than shCon-transduced cells and formed a reduced fraction of new T_{CM} precursors (Fig 2G, $p<0.0001$). Notably, these effects were greater than the effect of Mll1-depletion in naive cells during the primary response. In contrast, Cd62l^{lo} derived donor cells that were either shMll1-transduced or shCon-transduced accumulated similarly, but much less than the Cd62l^{hi} transferred cells (Fig 2H and Fig S2F). Thus, Mll1-deficiency preferentially affected the function of T_{CM}-like cells. Taken altogether, these results indicate that Mll1 promotes the early formation of precursors that generate *bona fide* T_{CM} cells, which are thought to be T_{SCM} cells(*3*).

Mll1 coordinately induces T_{MEM} gene expression and represses TE gene expression

To begin examining how Mll1 controls gene expression before T_{CM} cells definitively form, adoptively transferred shMll1-transduced or shCon-transduced P14 cells that were either Klrg1^{lo} or Klrg1^{hi} 5 days after LCMV$_{Arm}$ infection were sorted from spleens and analyzed using RNA-seq (Fig S5A). All cells had undergone at least 8 cell divisions and had strongly upregulated T_{EFF} cell gene expression compared to naive cells by this time (*23*). However, the gene expression signature of *Tcf7*^{hi} effector cells and those induced by the TFs Tcf1 (*Tcf7*), Tox and Klf2 were positively enriched within Klrg1^{lo} compared to KRLG1^{hi} shCon-transduced cells, whereas KRLG1^{hi} shCon-transduced cells were positively enriched with multiple signatures of *Tcf7*^{low} cells compared to compared to KRLG1^{lo} shCon-transduced cells (*29, 34*)(Fig S5B). Thus, Klrg1^{lo} cells might be T_{SCM}-biased (*3, 23*). Compared to shCon-transduced Klrg1^{lo} cells, shMll1-transduced Klrg1^{lo} cells failed to upregulate *Tcf7*^{hi} effector and T_{Prog} cell signatures (Fig 3A left) and expressed less *Bcl6, Id3, Tcf7l2, Jun, Junb, Jund* and *Il6ra*, which encode factors that promote cMP and T_{SCM}/T_{CM} cell formation (Fig 3A, *top right*). More genes in shMll1-transduced cells were less highly expressed than were overexpressed compared to shCon-transduced cells, indicating that Mll1 functions to activate genes associated with T_{CM} precursor formation during *in vivo* antigen stimulation. In addition, the signatures of genes that are normally upregulated in TE cells were positively enriched in both shMll1-transduced Klrg1^{lo} or Klrg1^{hi} cells (Fig 3A, *left)*, and a larger fraction of signature *Tcf7*^{low} cell genes were overexpressed in shMll1-transduced compared to shCon-transduced Klrg1^{hi} cells, including the key positive regulators of TE cells encoded by *Zeb2, Bhlhe40, Ccl5* and *Cx3cr1* (Fig 3A, bottom right). Thus, Mll1 contemporaneously activates T_{CM} cell gene expression and represses TE cell gene expression, prior to when these definitive cell subsets form.

Mll1 organizes transcriptional dynamics that establish the developmental trajectory of T_{SCM} precursors

he developmental order in which T_{MEM} cells initially arise relative to T_{EFF} cells (*35*)and whether T_{SCM} cells develop as a separate lineage that bypasses T_{EFF} cell differentiation is still unresolved(*3, 12, 23*). We anticipated that the effect of Mll1-deficiency on early single-cell trajectories might clarify how T_{SCM} cells initially form during acute viral infection. To this end, single-cell RNA-sequencing (scRNA-seq) was used to define gene expression in shCon-transduced and shMll1-transduced P14 cells 5 days after LCMV$_{Arm}$ infection, which were analyzed in context of cells derived from separate transfers of small numbers of naive P14 cells isolated on days

3 and 8 pi by implementing a mutual nearest neighbor batch correction (*9*, *36*). Analysis of cells from all time points identified 13 Leiden clusters (Fig 3B, *top*), which were labelled based on latent time (Fig 3B, *middle*) a measure of developmental time(*37*) that positively correlated with velocity pseudotime (Fig S6A), and described a potential developmental order. This arrangement placed clusters populated by transduced cells from day 5 pi intermediate between clusters formed by cells derived from days 3 and 8 pi (Fig 3B, *bottom*), suggesting they were transitional states. Single-cell RNA velocity-derived transition probabilities between each cluster were used to infer the future states of wildtype, shCon-transduced and shMll1-transduced cells, and enrichment of known gene expression signatures was used to identify the biological nature of cells in each cluster (Fig 3C-E). Cluster T1 was earliest in the trajectory and was positively enriched with gene expression induced by TCR and co-stimulatory signals, the T_{CM}-specific signature and the signatures of cMP and *Tcf7*^{hi} cells, but was negatively enriched with those of naive cells and other day 8 T_{EFF} cell subsets (Fig 3C,E, Fig S6B). T1 cells subsequently transitioned into either cluster T2, which repressed signatures of T_{CM}, T_{Prog} and *Tcf7*^{hi}-effector cells, or cluster T3, which maintained and upregulated these signatures (Fig 3C,E, Fig S6B). These results suggest activated cells develop into T_{SCM} precursors before T_{EFF} cells initially develop.

Mll1-deficiency severely impaired formation of cells that activate T_{SCM} gene expression. Wildtype and shCon-transduced cells in clusters T2 and T3 both branched into cluster T5, which was positively enriched with multiple T_{EFF} and T_{MEM} signatures, and cluster T6, which was positively enriched with the signature of double-positive effector (DP) cells (which are known precursors of T_{RM} cells(*38*)) and the mature T_{RM}-specific signature (*39*)(Fig 3C,E, Fig S6B). However, shMll1-transduced cells were severely depleted from cluster T3, and were significantly underrepresented in clusters T5 and T6 compared to shCon-transduced cells (Fig 3D). This suggested that Mll1 was essential for both the development or preservation of the initial T_{CM}/T_{SCM}-like state (T3), and for their ability to differentiate further. Supporting this conclusion, residual shMll1-transduced T3 cells lacked transition potential into clusters T5 and T6 compared to shCon-transduced cells (Fig 3E, Fig S6B). Thus, Mll1 was essential for establishing T_{SCM} gene expression in activated precursors before they developed into more heterogeneous T_{EFF} subtypes.

On day 5 pi, clusters T5 and T6 were transitioning into clusters T7-T9, which appeared to have developed lineage-bias, because cluster T7 was positively enriched with signatures of T_{CM} and TE cells, whereas cluster T8 was somewhat positively enriched with signatures of DP effector and TE cells, and cluster T9 was strongly enriched with the T_{RM}-specific and DP effector cell signatures. However, all three clusters were positively enriched with signatures of T_{Prog} and Tcf7^{hi}-effector cells and transitioned toward cluster T11, which most strongly expressed the Tcf7^{hi}-effector and T_{Prog} signatures compared to the entire trajectory (Fig 3C,E, Fig S6B). Nevertheless, Mll1-insufficient cells in clusters T7-T9 and T11 less strongly expressed the T_{Prog} and Tcf7^{hi} signature compared to shCon-transduced cells (Fig S6C), and specifically expressed less *Tcf7*, *Tox*, *Slamf6*, *Cd27*, *Ctla4* and *Sell*, among other positive regulators of T_{SCM}/T_{CM} cells in each of these clusters (Fig 3F). These data, and the fact that Mll1-deficient cells do not form *bona fide* T_{EX}^{prog} or T_{CM} cells (Figs 1 and 2), implies that activated cells with T_{SCM} cell gene expression early in the response are common precursors of multiple T_{MEM} cell states, which universally require Mll1.

Clusters T7-T9 and T11 harbored larger frequencies of shMll1-transduced cells compared to shCon-transduced cells, suggesting that Mll1-deficient cells accumulated in these states but were developmentally "arrested" (Fig 3D). Consistent with this possibility, shMll1-transduced cells from clusters T7 and T8 exhibited "retrograde" transition potentials into clusters T5, T6 and T9, whereas shCon-transduced cells exhibited nearly uniform "forward" transition potential into cluster T11 (Fig 3E, Fig S6B). In addition, while shCon-transduced cells in cluster T9 had transition potential into cluster T8, which feeds into cluster T11, shMll1-transduced cells did not and instead exhibited weak direct transition potential into cluster T11, unlike shCon-transduced T9 cells (Fig 3E, Fig S6B). These data demonstrate that Mll1-deficient cells establish less "forward" transition potential and regress developmentally.

Normally, cells in cluster T11 bridged development from day 5 pi to a linear array of states (clusters T10-T13) from day 8 pi that sequentially repressed the Tcf7^{hi}-effector and T_{Prog} signatures and then induced gene expression specific to T_{EM} and TE cells (Fig 3C-F, Fig S6B). Although shMll1-transduced T11 cells overexpressed multiple genes that normally correlate or promote differentiation of T_{EM} and TE cells, including *Klrg1*, *Cx3cr1*, *Tbx21*, *Prdm1* and *Zeb2* (*29*, *30*, *40-42*) compared to shCon-transduced T11 cells (Fig 3F), they exhibited reduced transition potential into cluster T12, which normally flowed into cluster T13, compared to

shCon-transduced T11 cells (Fig 3E, Fig S6B]). Thus, Mll1-deficient cells are less poised to undergo T_{EFF} cell maturation typically observed on day 8 pi, even though they over-express multiple key factors that drive formation of T_{EM} and TE cells. Consistent with this conclusion, *Kmt2a* was re-expressed in wildtype day 8 T_{EFF} populations (Fig 3F) and Mll1-deficiency mainly affected the number of T_{EFF} cells that formed between day 8 and day 15 pi (Fig 2). Thus, Mll1 might act in multiple stages during the initial effector response to establish normal T_{EFF} and T_{MEM} cell populations.

To investigate how Mll1-deficiency perturbed transcriptional states defining the trajectory we identified driver genes whose RNA-velocity contributed most to transition potentials between clusters. Mll1-deficient cells in cluster T3 lacked strong transition probability into any state, while genes driving the transition probabilities of shMll1-transduced cells in cluster T5 cells were distinct from those of shCon-transduced cells. Those unique to shMll1-transduced T5 cells were more highly expressed in cluster T9 relative to cluster T5, and in shMll1-transduced cells compared to shCon-transduced cells (Fig S6D). In contrast, those unique to shCon-transduced T5 cells, or common to both shCon-transduced and shMll1-transuced cells, were more highly expressed in cluster T5 compared to T9 (Fig S6D). These results suggest that Mll1 regulates transcriptional dynamics to define the initial trajectories of activated T cells. The transitions between clusters T3, T5 and T6 were likely governed by Mll1 in cells responding to TCR stimulation, because these clusters were positively enriched with gene expression signatures normally induced within the first two days of infection (Best clusters 1-3), a time coinciding with TCR stimulation *in vivo* (Fig 3C, *bottom*), and shMll1-transduced cells less strongly activated these signatures and failed to upregulate genes induced by TCR-stimulation compared to shCon-transduced cells (Fig S6C). Notably, dynamic genes specific to shCon-transduced T5 cells were highly TCR-induced and required Mll1 for upregulation in response to TCR signals (Fig 3G). Thus, Mll1 appears to establish the trajectories of developing cells by activating gene expression in response to TCR signals.

Mll1 pioneers widespread H3K4me3 deposition at TSSs during TCR stimulation

Histone H3 lysine 4 tri-methylation (H3K4me3) deposition at transcription start sites (TSSs) defines poised and active promoters(*43*). Many genes that lack H3K4me3 in naive CD8 T cells acquire this modification in mature T_{EFF} and T_{MEM} cells(*44-46*), which could be a key step that establishes transcriptional capacitance of antigen-experienced cells. However, the point during differentiation when this occurs and how it relates to transcriptional dynamics that orchestrate early developmental trajectories of activated cells is unknown. Mll1 is an archetype H3K4-trimethylase (*47, 48*) but its genetic disruption has demonstrated limited effects on H3K4me3 in mammalians, and other Set-domain family methyltransferases are thought to establish and maintain the majority of H3K4me3(*15, 16, 49, 50*) . To investigate the role of Mll1 and determine how naive CD8 T cells initially acquire H3K4me3 and establish the transcriptional competence of T_{MEM} cells, we examined the regulation of H3K4me3 and RNA Polymerase II (Pol II) during initial TCR stimulation (Fig 4). The global expression of H3K4me3, Pol II hyperphosphorylated on serine-5 (Pol II-Ser5) or serine-2 (Pol II-Ser2) C-terminal domain (CTD) residues, and the catalytic component of MLL-COMPASS-like complexes (Mll1-C) was upregulated within 48 hours of TCR stimulation compared to naive cells (Fig S7A). Remarkably, depletion of Mll1 using shRNAmirs or sgRNAs and Cas9 during TCR-activation impaired induction of overall H3K4me3 and Pol II CTD-hyperphosphorylation on Ser5 and Ser2, but did not impair H3K4me1 or H2BK120 ubiquitylation (Fig 4A). Thus, TCR stimulated cells undergo widespread installation of H3K4me3 and transcriptional activation, which appears to require Mll1-C expression.

Mll1-C bound hyper accessible chromatin at TSSs of the most transcriptionally active genes, which were marked with H3K4me3 in *ex vivo* T_{EFF} cells, much more than distal hyper accessible regions marked with H3K4me1 (Fig S7B). Regions of H3K4me3 ("peaks") identified in naive, TCR-stimulated and *ex vivo* T_{MEM} ($Klrg1^{lo}$,$Il7ra^{hi}$) (*46*) cells exhibited substantial overlap, consistent with a core CD8 T cell associated H3K4me3 landscape in naive and antigen-experienced cells (*44-46*) (Fig 4B). Differential analysis revealed that naive CD8 T cells transiently extinguished and then regained some H3K4me3 peaks during TCR-stimulation, and progressively acquired a larger number of *de novo* H3K4me3 peaks during stimulation that increasingly overlapped those present specifically in *ex vivo* T_{MEM} cells compared to naive CD8 T cells (Fig 4C). Thus, H3K4me3 deposition is dynamically remodeled by TCR-signals and approximately 20% of mature T_{MEM}-specific

peaks were initially acquired in the first 72 hours of TCR stimulation, suggesting they were maintained "epigenetically".

To define how H3K4me3 installation correlated with T_{EFF} and T_{MEM} cell development, we examined how genes that underwent H3K4me3 changes following TCR stimulation were expressed in distinct T_{EFF} and T_{MEM} states (Fig 4D). Genes marked with H3K4me3 specifically in naive cells were more highly expressed by Klrg1^{lo} cells on day 5 pi, Tcf7^{hi} cells on days 8 and 30 pi, and in T_{CM} cells relative to other T_{MEM} subsets on day 30 pi (Fig 4D, *top row*). Conversely, genes that acquired H3K4me3 24 hours after TCR stimulation were more highly expressed in Klrg1^{hi} cells on day 5 pi and Tcf7^{lo} cells on day 8 pi, but paradoxically were also more highly expressed in Tcf7^{hi} cells and T_{CM} cells relative to other T_{EM} cells on day 30 pi (Fig 4D, *2nd row*). Moreover, genes marked with H3K4me3 in T_{MEM} cells were preferentially expressed in Klrg1^{lo} cells on day 5 pi and Tcf7^{lo} cells on days 8 and 30 pi, but were also upregulated in T_{CM} cells (Fig 4D, *3rd row*). Thus, genes marked with H3K4me3 in mature T_{MEM} cells and those marked earliest after TCR stimulation were expressed in both T_{EFF} cells and T_{CM} cells. In contrast, genes that acquired H3K4me3 72 hours after TCR stimulation were more highly expressed in Klrg1^{lo} cells on day 5 pi, Tcf7^{lo} cells on days 8 and 30 pi, and in terminal-T_{EM} cells relative to Tcf7^{hi} or T_{CM} cells on day 30 pi (Fig 4C, *bottom row*). These results demonstrate that early TCR-dependent deposition of H3K4me3 correlates with *in vivo* gene expression that underlies formation of Tcf7^{hi} memory cells.

Genes that acquired H3K4me3 after TCR stimulation were differentially expressed in distinct T_{EFF} and T_{MEM} states, which suggests H3K4me3 enables dynamic regulation rather than enforces transcription *per se*. Consistent with this, changes in expression of chromatin-associated transcripts (nascent RNA) were positively correlated with increased H3K4me3 (Fig 4D), although nascent transcripts from many genes that acquired H3K4me3 upon TCR stimulation were not significantly induced compared to naive cells (Fig 4E, Fig S7C). This indicated that H3K4me3 acquisition during TCR stimulation "primes", but does not commit genes to immediate transcriptional activation. Notably, genes whose expression increases at late T_{EFF} time points and in the T_{MEM} phase following infection exhibited increased H3K4me3 deposition upon TCR stimulation, but were not induced transcriptionally (Fig 4E, Fig S7C), whereas genes upregulated during the first 2 days of infection exhibited both increased H3K4me3 deposition and transcription upon TCR stimulation (Fig 4E, Fig S7C). This suggested that H3K4me3 deposition by early TCR-signals might "prime" genes for subsequent activation by additional signals. Consistent with this hypothesis, ChIP-seq analyses of total RNA Pol II demonstrated that genes whose nascent transcription was induced upon TCR stimulation, and those marked specifically with H3K4me3 in T_{MEM} cells exhibited significantly increased Pol II pausing index in TCR-stimulated cells compared to in naive cells (Fig S7D, $p= 0.011$, $p=0.002$). These data suggest that TCR stimulation and H3K4me3 deposition positively regulates transcriptional potential and T_{MEM} cell formation by promoting transcriptional pausing.

To examine the role of Mll1 in H3K4me3 deposition during TCR stimulation, naive CD8 T cells were transduced 24 hours after initial TCR stimulation with either shMll1 or shCon retroviruses and H3K4me3 was mapped 48 hours later (i.e., 72 hours after TCR-stimulation). Of ~15,000 peaks identified in shCon-transduced cells after TCR-stimulation, more than 8,000 exhibited significantly reduced H3K4me3 signal in shMll1-transduced cells while less than 20 had increased H3K4me3 signal in shMll1-transduced cells, compared to shCon-transduced cells (Fig 4F). Mll1-dependent regions corresponded to ~70%, and 50% of de *novo* H3K4me3 peaks induced 24 hours, or 72 hours after TCR stimulation, respectively, as well as about 70% of the H3K4me3 peaks present in T_{MEM} cells but not naive CD8 T cells (Fig 4G). Thus, Mll1-deficiency impaired H3K4me3 deposition at more than half of all peaks in TCR-stimulated cells, and broadly reduced *de novo* H3K4me3 installation that is induced acutely by TCR stimulation and that develops in mature T_{MEM} cells. All Mll1-dependent peaks were associated with genes that are strongly expressed in Tcf7^{hi} effector and T_{MEM} cells (Fig 4H, *top row*). Those normally induced 72 hours after TCR stimulation and in T_{MEM} cells were annotated to genes that are upregulated in Klrg1^{lo} cells on day 5 pi, in Tcf7^{lo} cells on days 8 and 30 pi, and in T_{EM} cells compared to T_{CM} cells (Fig 4H, *top row*), whereas those induced 24 hours after TCR stimulation were preferentially expressed in Tcf7^{hi} effector and Tcf7^{hi} memory cells, and in T_{CM} cells. These results indicate Mll1 drives TCR stimulated H3K4me3 at a large number of genes that establish transcriptional competence of T_{SCM} cells.

Conclusion:

Our results show that Mll1 appears to be the predominate methylase that deposits H3K4me3 during TCR stimulation, is essential for transcription that establishes T_{SCM} precursors, and it drives maturation of T_{EFF} cells.

Although the expression of multiple genes encoding factors that drive terminal T_{EFF} differentiation became hyperactivated in the absence of Mll1, the cells failed to mature. Thus, our data indicate that an inaugural step following TCR-mediated activation of naive CD8 T cells is Mll1-dependent differentiation into T_{CM} precursors that are primed for differentiation along multiple pathways. TCR stimulation initiates this process via Mll1 by installing *de novo* H3K4me3 and inducing recruitment of Pol II, which appeared to be paused at the TSSs of genes that are the most dynamic in T_{EFF} cells undergoing diversification early during viral infection. This implies that factors which govern the transition of Pol II from paused to elongating states at specific genes might account for transcriptional dynamics that initiate divergence between primed precursors before definitive states have developed. Our previous studies demonstrated that depletion of *Ccnt1* or *Cdk9*, which encode P-TEFb (positive transcription elongation factor b) prevents terminal T_{EFF} cell differentiation, and induces formation of cMP cells (*14*), which suggests that transcriptional pausing favors differentiation of cMP and T_{MEM} cells. Thus, we propose that Mll1 might be central for establishing transcriptionally competent Pol II complexes in activated CD8 T cells that are paused, to both preserve T_{MEM} formation and to facilitate T_{EFF} cell maturation.

Figure Legend

Fig 1. An RNAi screen of chromatin regulators reveals Mll1controls co-inhibitory gene expression and is expressed in regenerative T cell subsets

A. Screen Schematic: In brief, retroviral packaging cells (PLAT-E) were transfected in a 96 well format with individual shRNAmirs targeting murine chromatin regulators and control genes. Simultaneously, naïve P14 TCRα^{KO} mice Cd8^{+} T cells were isolated and activated with αCD3 and αCD28 antibodies. 16 hours post activation, cells were transduced with shRNAmirs (*51, 52*). For the day 3 screen (d3^{TCR}), cells were left on TCR stimulation for an additional two days and FACs analyzed. For the day 6 screen, cells at day 2 were cultured in either high or low doses of IL-2 (IL-2^{hi} and IL-2^{lo} respectively) and then analyzed on day 6(*53*).

B. Histogram of screen results: Histograms show the effect of RNAi on d3^{TCR} *Id3*-GFP (left) and IL-2^{hi} Pd-1 (right). Each bar represents the average RNAi effect of targeting a gene on marker expression. Dotted lines show top and bottom 15th percentile. Blue bars show members of the COMPASS complex and red, BAF.

C. Effects of individual RNAi effects: Bar chart shows raw geometric mean of d3^{TCR} *Id3*-GFP (top) and IL-2^{hi} Pd-1 (bottom). Individual points represent distinct RNAi clones. P values in light blue show significance compared to the distribution of all RNAis, those in grey show in comparison to shCon (shCd4 or shCd19) targeting RNAi. P values computed by Welch's t-test

D. Heatmap of RNAi effects: Euclidean clustering was performed on all RNAi scores of COMPASS (blue) and BAF (red) complex associated RNAis. RNAis were arranged by d3^{TCR} (top), IL-2^{hi} (middle) and IL-2^{lo}(bottom).

E. Flow cytometry histograms of individual RNAi effects: Histograms shows modal expression of d3^{TCR} *Id3*-GFP (left), d3^{TCR} Pd-1 (middle) and IL-2^{hi} Pd-1 (right) in control (grey) or Mll1 RNAi (blue) transduced cells. Each row of histograms shows unique shMll1 targeting RNAi. Numbers show geometric mean in each case.

F. Normalized expression data of COMPASS, COMPASS-like, T_{SCM}, and T_{DYS} related genes: Heatmaps show normalized microarray and RNA-Seq transcript counts from cells responding to Listeria-ova or LCMV$_{CL13}$ on the left and right respectively (*54*) (*20*). Intensity of color represents highest expression of an individual gene across columns.

G-H: P14 Cd8 T cells were transduced with either control (shCd19/shCd4) or Mll1 (shMll1) targeting RNAis and transferred into mice bearing E0771^{GP33} tumors. Data shown is from TIL. Flows plots show concatenated individual mice (n=3). Individual points in bar plots show individual mice. Bar height shows mean. P values were calculated via paired t-test.

G: Day 3 pACT co-inhibitory gene expression: Left shows co-expression of Pd-1 and Tim3, right shows expression of Pd-1 within Tim3+ cells. Numbers in left show percent of population, numbers in right show geometric mean.

H: Day pACT T_{EX}^{prog} cells: Middle shows percentage of population, right shows normalized transduced P14 cell numbers. Numbers show percentage of population's

I-J. P14 Cd8 T cells harboring a Tcf7 reporter construct were transduced with either control (shCd19/shCd4) or Mll1 (shMll1) targeting RNAis and transferred into mice that were given chronic LCMV infections (LCMV$_{CL13}$). Data shown is from spleens. Flows plots show concatenated individuals (I, n=3; J, n=4). Numbers show

percentage of population Cells within the Tcf7^{hi}, Tim3^{lo} population (top) were selected and used for the bottom panel. Individual points in bar plots show individual mice. Bar height shows mean. P values were calculated via Welch's t-test, except for J (top), which uses the Mann-Whitney test.

I. Day 5 PI

J. Day 15 PI

Fig. 2. **Mll1is essential for the early establishment, maintenance, and function of regenerative T cell subsets**

A-F: Wild type or Tcf7 P14 Cd8 T cells were activated, transduced with either control (shCd19/shCd4) or Mll1(shMll1) targeting RNAis and transferred into mice. Mice were given and LCMV$_{ARM}$ infection. Flows plots show concatenated individuals (number in legend). Individual points in bar plots show individual mice. Bar height shows mean. P values reported from specified test.

A. T cell accumulation in presence or absence of Mll1: Lines show normalized transduced P14 cell numbers of control (grey) or shMll1 (blue) cells at given day. Shaded plot shows normalized percent transduced P14 percent of all Cd8 T cells.

B. Day 3: Concatenated data of four individuals. Left- Prior to activation, P14 cells were labeled with CTV. Black bar set to median number of divisions, numbers on plot show percent of population below median. Middle and Right- Numbers show geometric mean of the marker positive fraction of cells. Paired t-test used for statistics.

C. Day 5: Concatenated data of five individuals. Numbers show percentage of population. Paired t-test used for statistics.

D. Day 8: Concatenated data of three individuals. Numbers show percentage of population. Paired t-test used for EE, DP, and TE data; Wilcoxon used for MP.

E. Day 15: Concatenated data of three individuals. Numbers show percentage of population. Paired t-test used for Cd27^{hi}, Cx3cr1^{lo} and Cd27^{hi}, Cx3cr1^{int}; Wilcoxon used for Cd27^{int}, Cx3cr1^{hi}.

F. Day 90: Left- Concatenated data of three individuals. Numbers show percentage of population. Mice with less than 10 shCon or shMll1 cells collected were excluded from statistical analysis of percent population. Paired t-tests were used for p value calculations. Right- Normalized number of transduced T$_{CM}$ (Cd62l^{hi}, Cd127^{hi}) P14 cells at given day.

G. Day 15: Concatenated data of three individuals. Numbers show geometric mean of TCF7-EGFP within labeled groups. Paired t-test used for statistics

H. Expansion and phenotype of Mll1 deficent T$_{CM}$ and T$_{EM}$ cells after secondary infection. Top-Five days post infection cell numbers were quantified and adjusted for input ratios. F.C represents the fold change between control and Mll1 deficient cells, F.E represents fold enrichment from expected #s of engrafted cells. P values were calculated using paired t-tests. Bottom- Effect of Mll1 on recall co-expression of Cd27 and Cx3cr1 in T$_{CM}$. Data shown is concatenated from four individuals. Paired t t-tests used for p value calculations

Fig 3: Mll1 is necessary for the transcriptional programing, maintenance, and transitions of T$_{SCM}$:

A. Analysis of bulk RNA-seq data: Day 5 post $LCMV_{ARM}$ cells were sorted by $Klrg1^{hi}$ and $Klrg1^{lo}$ and transduction with shMll1 or shCon. Left: GSEA of early activation, Day 8, and memory sets within shMll1 vs. shCd19 $Klrg1^{lo}$ and $Klrg1^{hi}$ RNA-seq data. Right: Volcano plot of shMll1 vs. shCd19 $Klrg1^{lo}$ and $Klrg1^{hi}$ RNA-seq data. Genes associated with $TCF7^{hi}$ and $TCF7^{lo}$ cells that were significantly regulated were marked in red or blue accordingly. Genes of interest below p<=0.01 were labeled.

B. Force assisted graphing of scRNA-Seq: In brief, P14 Cd8 T cells from D3 and D8 that were previously analyzed(*9*) were merged with D5 shCon and shMll1 cells. Leiden clustering was then used to arrange cells by cluster and clusters were displayed via PAGA initialized force assisted analysis. Top: Position of original Leiden clusters with time dependent reordering, Middle: Latent time of each cell, Bottom: Identify of cells origin.

C. Cluster dependent gene set enrichment analysis: Gene set enrichment scores of marked subsets within different clusters. Each cluster's Wilcoxon rank score vs. all other clusters was used to tabulate enrichment. Top: Phenotypic based enrichments derived from previous RNA-Seq(*3, 9, 10, 39, 51*). Bottom: Enrichments from time-resolved responses to LM_{OVA}(*54*).

D. Cluster specific metadata: Top: Percent of each cluster's cells origin. Bottom: Percent of each cell's origin per cluster.

E. Cluster specific transition probabilities: Using only D5 cells from control (left) or shMll1 (right) cells transitions from one cluster to another were calculated. Shown is a stick and ball graphic, where circle size represents cluster size and arrow direction transition probability.

F. Heatmap of immunological gene expression across pseudotime: Heatmap shows the log 2-fold changes of a subset of genes that were differently regulated between all shMll1 and shCon cells. Heatmaps are broken up as thematically labeled/ Euclidean clustering was performed on the shMll1 vs. shCon comparison (right). Left: Expression of genes in cluster versus all other clusters using all cells. Bottom: Expression in cluster comparing shMll1 vs. shCon cells.

G: Transcriptional differences between dynamic genes in the absence of Mll1 and during TCR stimulation: The top 10th percentile of dynamic genes in shCon cells in cluster T5 that were not dynamic in shMll1 cells and significantly regulated 48H after TCR stimulation (padj<=0.05) were selected. The log 2FC of expression was plotted and clustered by Euclidean distance. Time resolved columns are from nascent RNA-Seq, shMll1 vs. shCd19 from all cells in the scRNA-seq experiment.

Fig 4: H3K4me3 deposition and RNApol2 accumulation are coordinated after TCR stimulation and are regulated by Mll1

A. Western blot showing the effect of Mll1 depletion on protein: Left: Expression changes in nucleosome components and post-translational modifications. Right: Changes in RNApol2 phosphorylation.

B. Overlaps of all H3K4me3 peaks: Venn diagram shows the overlaps of all H3K4me3 peaks at the given time points.

C. Differential H3K4me3 peak enrichment: Left: The number of differential peaks (pvalue<=0.05) between the day marked and naïve cells. Black represents enrichment in naïve and red enrichment in the day post activation. Right: Venn diagrams of the overlaps of differential peaks (red from left side).

D. Gene set enrichment: Peaks in red from Figure A were assigned to nearest genes and the fraction of these genes that constitute the indicated gene set was calculated as the enrichment score. Naive enrichment score was calculated as the intersect of genes who had differential peaks across naïve cells in all conditions (black from Fig 4C).

E. Comparison of H3K4me3 binding and nascent RNA-Seq at similar time points: The log2 fold changes for H3K4me3 binding and nascent transcription were calculated on a per gene basis and plotted. Genes previously identified as differentially expressed in Tcf7$^{hi/lo}$ cells were marked with blue or red accordingly(*3*). Left: D1 H3K4me3 and D1 nascent RNA-seq. Right: D3 H3K4me3 and D2 nascent RNA-seq.

F. Venn diagram of Ml11 dependent H3K4me3 peaks: Differential peaks between shCon vs. shMll1 cells were marked in grey if their was no difference, blue if they were enriched in control cells, and pink if in enriched in Mll1 deficent.

G. Enrichment of Mll1 dependent H3K4me3: Peaks that were significantly enriched in Mll1 sufficient (blue Fig 4G) samples versus deficient were overlapped with all peaks at D1, D3 or D60 to determine what fraction of peaks are Mll1 dependent. Blue shows number of Mll1 dependent, red all other peaks.

H. GSEA of Mll1 dependent H3K4me3 in memory T cell subsets: The Mll1 dependent (Fig 4G) and Mll1 overlapping dependent (Fig 4H) genes were used as sets for GSEA.

SupFig. 1. **Inhibition of families of chromatin regulators have both distinct and convergent effects on co-inhibitory gene expression**

A. Screen analysis strategy: In brief, the geometric mean from each marker was calculated per shRNAmir. Control RNAis served were used as the population reference to calculate individual Z scores. From these Z scores a per shRNAmir percentile shift was calculated and averaged across all shRNAmirs targeting the same gene. These average shifts were used to calculated Z-scores, pvalues, and RNAi effects.

B. Histogram of screen results: Histograms show the effect of RNAi on labeled marker genes. Each bar represents the average RNAi effect of targeting a gene on marker expression. Dotted lines show top and bottom 15th percentile. Blue bars show members of the COMPASS complex and red, BAF.

C. GSEA of chromatin regulator complexes within screen: Color represents normalized enrichment score values (NES) whereas size is the -log10 of the nominal pvalue.

D. Correlations of RNAi effect across culture conditions: Each box shows the correlation of RNAi scores between different markers in different culture conditions. Size and color of the circles in the box represent intensity of correlation.

SupFig. 2. Disrupted Mll1 increases co-inhibitory gene expression and decreases Tcf7 expression

A . Effects of individual RNAi depletion on marker gene expression. Top: Bar chart shows raw geometric mean of marker expression. Individual points represent distinct RNAi clones. Individual points represent distinct RNAi clones. P values in light blue show significance compared to the distribution of all RNAi, those in grey show in comparison to shCon (shCd4 or shCd19) targeting RNAi. P values computed by Welch's t-tests. Bottom: Histograms shows modal expression of marker genes in control (grey) or Mll1 RNAi (blue) transduced cells. Each row of histograms shows unique shMll1 targeting RNAi. Numbers show geometric mean in each case.

B. Effects of CRISPR deletion on marker gene expression. Top: Bar chart shows raw geometric mean of marker expression. Individual points represent distinct replicates. P values were calculated by Welch's t-tests. Bottom: Expression in control (grey) or Mll1 targeting sRNA (blue) transduced cells. Numbers represent geometric mean of marker.

C. Western blot confirmation of Mll1depletion by RNAi (left) or CRISPR (right).

D. Tcf7-eGFP expression *in vitro*. In brief naïve P14 T cells bearing a Tcf7-eGFP reporter were activated, transduced with RNAi, and cultured to day 6 in 100U (IL-2^{hi}) or 10U (IL-2^{lo}) of IL-2. Top: Representative FACS plot at given day and culture condition. P values represent significant difference of geometric mean based on paired t-test. Bottom: Data shows average gMean of Tcf7-eGFP + populations at given date post activation.

E. CTV dilution *in vitro*. In brief naïve P14 T cells were pulsed with CTV before activation, transduced with RNAi, and to day 6 in 100U (IL-2^{hi}) or 10U (IL-2^{lo}) of IL-2. Top: Top: Representative FACS plot at given day and culture condition. P values represent significant difference full CTV dilution at given timepoint based on paired-test. Bottom: Data show average percent of the populations that had completely diluted out CTV at the day indicated.

SupFig. 3. Mll1regulates co-inhibitory protein expression and the formation of T_{Prog} cells in TIL

A. Tumor model schematic: Top: EO771 tumors were stably transduced with a pMigR vector expressing the GP33-41 LCMV peptide linked to GFP. $1*10^6$ $E0771_{GP33}$ cells engrafted orthotopically into wild type, female, naïve, C57Bl/6J mice in a 1:1 V/V ratio with matrigel. After approx.. two weeks, tumors were palpable. At this time P14 Cd8 T cells transduced with either control or Mll1 targeting RNAi were transferred into mice. Bottom: B16 tumors were stably transduced with the same vector as above. $5*10^5$ cells were implanted and after 8 days, P14 Cd8 T-cells crossed to CD45.1 mice were transduced with either control or Mll1 targeting RNAi and transferred.

B/C. Selective enrichment of P14 cells in antigen specific tumors: B-Mice were contralaterally implanted with $E0771_{GP33}$ and $E0771_{GFP}$ tumors and allowed to grow for ~2 weeks. P14 cells were adoptively transferred, and tumors and spleens were taken at day 5 pACT. Top: Numbers show percent of displayed Thy1.1, Cd44 positive cells. Bottom: Percent of total Cd8s and geometric mean of Pd-1 expression. P value was calculated via paired t-test C- $B16_{GP33}$ tumors were implanted and 8 days later, P14 cells were adoptively transferred, and tumors and spleens were taken at day 3 pACT. Top: Flow plot from the concatenate of four individuals. Numbers show percent of displayed Cd45.1, Cd8 positive cells. Bottom: Percent of total Cd8s and geometric mean of Pd-1 expression. P values were calculated with paired t- tests

D. Day 3 pACT co-inhibitory gene expression in $B16_{GP33}$ tumors: Numbers show percentage of population. Data shown is the concatenate of three individuals. P values were calculated via paired t-test.

E. Pd-1 expression in $Tim3^{lo}$ cells at D3 pACT: The $Tim3^{lo}$ transduced P14 cells within $B16_{GP33}$ tumors were gated and the gMean of Pd-1^{hi} expression within that population was analyzed. Numbers show geometric mean. Data shown is the concatenate of three individuals. P values were calculated via paired t-test.

F. Accumulation of Mll1 sufficient and deficent cells within the TME: Top: Normalized accumulation of transduced P14 cells within $E0771_{GP33}$ at day 3 (left) or day 8 (right). Wilcoxon test used to calculate significance. Bottom: Accumulation as percent of Cd8 T cells within $B16_{GP33}$ at day 3 (left) or day 8 (right).

G. Control of TNBC by Mll1 sufficient and deficent tumors: In brief, $1*10^6$ shCon transduced cells, or $1*10^6$ shMll1 transduced cells were transferred into tumor bearing mice and tumor growth was measured every other day. Top: Growth of TNBC over time. Each line shows the average growth of tumors overtime, smoothed over five days. Black line shows control mice; grey, those transferred with shCon; and blue, shMll1. Bottom: Kaplan Myer curve of mice reaching 1,500mm^3 tumor size. P values calculated using Mantel-Cox test.

H. Expression of Tcf7-eGFP within TIL: Numbers shown are the geometric mean of Tcf7^{hi} cells within each group. Data shown is the concatenate of three individuals. Paired t-tests used for p value calculation. Top: expression within Tim3^{lo} cells. Bottom: Expression of Tcf7-GFP in transduced P14 Cd8 T cells.

SupFig. 4. Mll1 regulates the formation of proliferating cells during primary infection

A. Klrg1 expression at day 5 PI LCMV$_{ARM}$: Concatenated data of five individuals. Numbers show percentage of population. Paired t-test used for statistics.

B. EdU incorporation at day 5 PI LCMV$_{ARM}$: Concatenated data of two individuals. Numbers show percentage of population. Paired t-test used for statistics.

C. Cd8 T-cell of phenotype after disruption of Mll1 via targeting sgRNAs at day 8 PI LCMV$_{ARM}$: Concatenated data of three individuals. Numbers show percentage of population. Paired t-test used for statistics.

D. EdU incorporation at day 8 PI LCMV$_{ARM}$: Concatenated data of three individuals. Numbers show percentage of population. Paired t-test used for statistics.

E. Accumulation of memory cells overtime: Normalized numbers of transduced P14 cells at day 15 or day 90. Mann-Whitney test used for p value calculation. Left-Accumulation of T_{EM}; Right- Accumulation of t-T_{EM}.

F. Schematic of secondary infection: In brief, at day 90, control and Mll1 deficent cells were sorted based on Cd62l expression and ~ 2,500 control and Mll1 deficient cells were transferred into host mice which were given LM$_{GP33}$ infections. Five days later, mice were sacrificed and analyzed.

SupFig 5. Bulk RNA-Seq shows distinct polarization of Klrg1^{lo} vs. Klrg1^{hi} cells at day 5 PI

A. Sequencing schematic: In brief, P14 Cd8 T cells were activated, transduced with shRNAmirs targeting either Mll1 or control and transferred into a mouse. One hour later, mice were infected with LCMV$_{ARM}$. Five days post infection, after gating on live, singlet, transferred cells, transduced (shMll1-Amatrine, shControl-YFP) cells were sorted by Klrg1^{hi} or Klrg1^{lo} expression, respectively. RNA was then extracted for bulk sequencing.

B. Transcriptional differences of Klrg1^{hi} or Klrg1^{lo} cells at day 5 PI: Left: Volcano plot of Klrg1^{hi} vs Klrg1^{lo} cells. Genes associated with TCF7^{hi} and TCF7^{lo} cells that were significantly regulated were marked in red or blue accordingly. Genes of interest below $p<=0.01$ were labeled. Right: GSEA of Klrg1^{hi} vs Klrg1^{lo} cells.

SupFig 6. Mll1 effects core cluster signatures, cell fate trajectories, and dynamic genes early during LCMV$_{ARM}$ responses

A. Pseudotime analysis of scRNA-seq: Pseudotime analysis was performed on all analyzed cells via SCANPY. A random cell in T1 was used as the starting node.

B. Cluster specific transition probabilities: Using only D5 cells from control (left) or shMll1 (right) cells transitions from one cluster to another were calculated.

C. Cluster dependent gene set enrichment analysis: Gene set enrichment scores of marked subsets within different clusters. Each cluster's Wilcoxon rank score of shCon vs. shMll1 was as used to tabulate enrichment. Top: Phenotypic based enrichments derived from previous RNA-Seq(*3, 9, 10, 39, 51*). Middle: Enrichments from time-resolved responses to LM_{OVA}(*54*). Bottom: Enrichments from transcription factor dependent regulation (*55*)

D. Venn diagram of dynamic gene overlaps: Venn diagram shows the overlaps of the dynamic genes from clusters T5 and T8 in shCon and shMll1 cell clusters

E. Transcriptional differences between dynamic genes in the absence of Mll1: The union of top 10th percentile of dynamic genes in clusters 3 and 1 with and without Mll1 were select and (top) their presence in either control, shMll1, or both were determined. Bottom: The expression between clusters (top) and within clusters between shMll1 vs shCon (middle/bottom) was plotted.

SupFig 7. Activation dependent changes in transcriptional machinery

A. Western blot showing the effect of Mll1 depletion on protein: Left: Expression changes in nucleosome components and post-translational modifications. Right: Changes in RNApol2 phosphorylation.

B. Heatmaps of ChIP-Seq binding: Heatmaps show ChIP-exo (Mll1), ChIP-seq (H3K4me1 and me3) and ATAC-seq signals overlapping promoters, and centered on accessible TSSs (left), and distal ATAC-seq regions centered on H3K4me1 peaks (right**)**

C. Comparison of H3K4me3 binding and nascent RNA-Seq at similar time points: The log2 fold changes for H3K4me3 binding and nascent transcription were calculated on a per gene basis and plotted. Genes previously identified as differentially expressed as involved in initial activation (red), naïve/memory cells (blue), or effector/memory cells (gold), were marked accordingly(*2, 26*). Left: D1 H3K4me3 and D1 nascent RNA-seq. Right: D3 H3K4me3 and D2 nascent RNA-seq

E. Pausing rate of RNApol2: The ratio of counts in promoter (TSS +/- 300bp) to counts in matched gene body was calculated and plotted as a cumulative distribution function for significantly enriched nascent naïve or 48H post TCR activation transcripts (top) or H3K4me3 enriched in naïve vs. memory cells (bottom). Significance difference of the distributions was calculated by Kolmogorov Smirnov test.

Materials and Methods:

Vectors, shRNAmirs, cDNA constructs and retroviral packaging

The MSCV-based retroviral vectors MigR1, LMPd, and LsgA were used to transduce activated $Cd8^+$ T cells. MigR1 encoding IRES-GFP, was used to deliver *RUNX3* (long form) cDNAs(*51*). LMPd, encoding Pgk-Ametrine1.1 or Venus was used to express shRNAmirs, whose sequence was predicted by the shERWOOD algorithm and were purchased from TRANSOMICS(*14, 52, 56*). The LsgA vector was previously described and the Chop-Chop program was used to design guides(*57*). Retroviral supernatants were generated in Platinum-E packaging cells as previously described(*14*).

Mice

C57BL/6J mice used in this paper were obtained from the Jackson Laboratory. P14 $Thy1.1^+$ mice were a gift from R. Ahmed (Emory University). $Thy1.1^+$,$Thy1.2^+$ positive mice were generated by crossing the P14 mice to wildtype BL/6J mice. The Tcf7 and Blimp-1 reporter mice were purchased from Jackson and crossed to P14 mice to create P14-Tcf7 or P14-Blimp-1 mice respectively(*58, 59*).The P14 $TCRa^{-/-}$ were initially purchased from Taconic(*53, 60, 61*). All mice were maintained in specific-pathogen free facilities and used according to protocols approved by the Institutional Animal Care and Use Committee of TSRI-FL.

Naïve T-cell culture and transduction:

$Cd8^+$ T cells were isolated from naïve mice using negative selection magnetic beads (StemCell). For primary TCR-mediated activation, purified $Cd8^+$ T cells were plated at a density of $4x10^5$ cells/cm^2 at a concentration of $1x10^6$/mL and on plates pre-coated with goat anti-hamster capture antibody and media supplemented with hamster anti-CD3 (2C11) and anti-CD28 (37.51) (1ug/mL each). For retroviral transductions, 16 hours after initial plating, the T cell conditioned media was taken off, saved and replaced with retroviral supernatants containing 10ug/mL polybrene, and the cultures were immediately centrifuged at ~ 500xg for 60 min at ~ 37 degrees C, followed by incubation for an additional 3 hours at 37 degrees C at 10% CO_2 incubator. Media containing retrovirus was aspirated and replaced with the saved conditions media. Cells were either transferred into mice (see infections) or continued to be cultured. For *in vitro* culturing, after 48 hours of total TCR stimulation, the activated cells were directly resuspended, counted and re-cultured by diluting the cells into fresh media containing 10 or 100 U/mL rhIL-2 (NCI). Cells were re-cultured as $5x10^5$ cells/mL, every 24 hours following removal from the TCR stimulus(*53*).

For experiments where cells were CTV labeled, after isolation, cells were pulsed with CTV for 10 minutes prior to activation.

Screening Analysis:

The geometric means from the live singlets transduced with shRNAmirs were measured via FACs. Control RNAis were used as the population reference to calculate individual Z scores at day 6 while the full population was used at day 3. From these Z scores a per shRNAmir percentile shift was calculated and averaged across all shRNAmirs targeting the same gene. These average shifts were used to calculated Z-scores, pvalues, and RNAi effects, which was calculated as $z * (1 - pvalue)$.

GSEA was computed on the RNAi effect values using cluster profiler(*62*) (Subramanian, 2005 #485}. Chromatin regulator family gene sets were custom built for this paper

Pathogens, infections, and colony forming assays:

LCMV-Armstrong and LCMV-Clone13 were produced as previously described (*63. For experiments using in vitro transduced P14 cells, 50,000 activated cells were adoptively transferred into naïve C57/Bl6J mice r.o and an hour later infected with either 2×105 PFU of LCMV-Arm i.p. or 2x106 PFU r.o. For Listeria monocytogenes-GP33 (LM-GP33) was grown as previously described on brain heart infusion media{Olson, 2013 #463*). Mice were infected with 10,000 CFU for recall experiments.

Tumor transplant, ACT, and harvesting:

E0771 cells containing either a control MigR1-GFP or MigR1-GP33-GFP were cultured per providers instructions. For transplant, $1x10^6$ tumors cells per mouse were resuspended in 100µL of media and mixed 1:1 with 100uL of high concentration growth factor reduced (HC-GFR) matrigel (Corning) on ice and injected s.c. into mice mammary fat pads. 2 weeks later, $1x10^6$ P14 cells transduced with shRNAmirs were transferred into tumor bearing mice.

For harvesting, tumors were excised and dissociated into ~1cm pieces. Tumor were digested for 10 minutes in 10uL/mL DNAseI and 1uL/mL Collagenase and strained over a 70µM filter and spun for 10 minutes at 300g. Supernatant was then aspirated, and pellet was resuspended in 1mL of ACK lysis buffer for 5 minutes with occasional agitation. After 5 minutes, lysis buffer was diluted with 5 mL of cTCM and strained over a 45 µM filter and spun for 10 minutes at 300g. Pellets were then resuspended and spun in a 60/40 Percol gradient to specifically enrich lymphocytes.

B16 ACT experiments were performed as previously described(*64, 65*).

Ex-vivo isolation from tissues-spleen

For EdU experiments, mice were pulsed for either four hours (day 5) or overnight (day 8) with Edu before sacrifice.

T cells from mice infected with LCMV before day 6 were digested as previously reported(*66*). In brief, spleens were harvested into T-cell optimized media (cTCM)(*53*) and cut into ~1cm pieces. Spleens were digested for 10 minutes in 10uL/mL DNAseI and 1uL/mL Collagenase (get **precise** units). Spleens were then treated as all other splenic experiments.

Spleens were strained over a 70µM filter and spun for 10 minutes at 300g. Supernatant was then aspirated, and pellet was resuspended in 1mL of ACK lysis buffer for 5 minutes with occasional agitation. After 5 minutes, lysis buffer was diluted with 5 mL of cTCM and strained over a 45 µM filter and spun for 10 minutes at 300g. For analytic flow cytometry, pellet was resuspended in 1mL of FACS buffer (1xPBS, 2% FBS, 10mM EDTA).

For sorting, cells were negatively isolated using a modified StemCell $Cd8^+$ kit. In brief, biotinylated αCD19, αCD4,αB220,and αTER-20 antibodies were added to cell pellet in 1mL of cTCM and 50µL of rat serum. After 10 minutes, the manufactures instructions and reagents were followed and used. Cells were then transferred into FACS buffer for staining at ~30-40 million cells per mL (See staining protocol). After staining, cells were sorted via a BD Aria Fusion sorter.

Flow cytometry:

After tissue isolation, cell numbers were counted via an Acuri C6 or MACSQuant 10. Samples were transferred into FACS buffer and stained for either 20 minutes at RT or 40 minutes at 4C. Samples were then washed and fixed for 10 minutes at 4 C in 4% PFA. Samples were washed again and placed into FACS buffer at 4 C until acquisition. Samples were acquired on a BD LSR2 or a Cytek Aurora

Generation of bulk RNA-seq libraries and sequencing

For bulk RNA-seq analysis of *in vivo* CTL, distinct phenotypic subsets were isolated by FACS and sorted into FBS coated DNA LoBind Eppendorf 1.5mL with 200uL of 1xPBS, 1mM DTT, and 5U of Recombinant RNAse inhibitors (Takara) , and then were washed in 1x PBS. Full-length cDNAs were generated directly from ~10,000 cells per sample using oligo(dT) priming and the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech), and 150 pg of cDNA was used to prepare strand-specific paired-end sequencing libraries (Nextera XT, Illumina). Libraries were quantified using Qubit HS DNA assay (Invitrogen) and their quality was assessed on Agilent Technology 2100 Bioanalyzer using a High Sensitivity DNA chip (Agilent) and dual-indexed pooled libraries were sequenced. Reads were aligned with salmon using validateMappings(*67*). Aligned reads were passed to DESEQ2 for differential gene analysis(*68*). All gene sets used were derived from previously published papers: EMP, naïve, and 48H(*9*); TE,MP,EE, and DP(*51*) ; T_{RM}, T_{CM}, and T_{EM} (*39*) Tcf7(*3*); and T_{prog}(*10*).

Nascent RNA-seq library preparation, sequencing, and data analysis were previously described(*9*)

Single cell clustering, cluster specific gene expression, and pseudotime analysis:

Single cell library preparation, sequencing, and alignment were previously described(*9*).

Single cell data was manipulated and analyzed first in the scanpy environment (*69*). Cells from D5 shCon, D5 shMll1, and D8 were down sampled to 800 cells and those matched to D3 to 400 cells. Mutual nearest neighbor batch correction was used to remove batch effect from the two separate sequencing experiments (mnn_correct) (*36*). Using the merged experiments, principal components (tl.pca) and nearest neighbors were calculated (pp.neighbors, neighbors=100, pcs=25) and then used to create Leiden clusters (tl.leiden, resolution =1). These clusters were then used to create a PAGA graph to estimate and simplify the connectivity of the clusters (tl.draw_graph) (*70*).

Data from these clusters was used for all single cell differential expression calculations using the Wilcoxon method (tl.rank_genes_groups). For GSEAs the score for each given analysis was used as the rank ordered input. To calculate pseudotime, an arbitrary D3 cell from Leiden cluster 7 (T1 in the text) was selected as the initial cell (tl.dpt).

RNA velocity, latent time, transitions, and dynamic genes:

To calculate RNA velocity, scvelo was used(*37*).First, using the pre-pseudotime PAGA analysis, neighbors, moments and initial velocity were calculated (pp.neighbors neighbors =100, pcs=50, pp.moments, and tl.velocity respectively). Next dynamic genes were recovered and used to compute the dynamical model (tl.recover_dynamics, tl.velocity). Using these dynamic genes both latent times (tl.latent_time) and ranked dynamic genes (tl.rank_dynamical_genes) were calculated. To calculate individual dynamic genes for either control or Mll1 deficent cells, cells originating from either Mll1 deficent or control samples were removed prior to scVelo analysis. These dynamic genes were then used to calculate transition probabilities.

Chromatin I.P: T cells were fixed in PBS containing a 1:10 ratio of fixation buffer (11.1% V/V MeOH free formaldehyde, 100mM NaCl, 1mM EDTA, and 50 mM Hepes pH7.5) for 15 minutes on ice. Fixation was

quenched with a 1:20 addition of 2.5M glycine for an additional 15 minutes. Cells were then washed with PBS and nucli were extracted (10mM Tris pH7.5, 10mM EDTA, 150 mM NaCl, and 1% TritonX-100). Nucli were resuspended in lysis buffer (50mM Tris pH 8.0, 150mM NaCl, 1mM EDTA, 1% Nonidet P-40, 0.1% SDS, and 0.5% Deoxycholate, supplemented with protease and phosphate inhibitor tablets (Roche)). To make mono and di nucleosome size fragments for I.P, sonication was performed in covaris mini tubes for 30-45 minutes with following settings: peak power-160, duty factor -10, cycles/burst=200.

Before the addition of antibody, chromatin was brought up to 1mL total volume in lysis buffer and pre-cleared via the addition of 50uL of Protein G Dynabeads per 10^7 nucli equivalents at 4C for 2-3H. 5% of each reaction was then saved for making the input control. The equivalents of 10-20 million nucli were subdivided for each I.P per sample and brought up to 1mL total volume. Antibodies were then added to the chromatin and I.P'd for ~16-20H at 4C. Each sample was then conjugated to fresh Protein G Dynabeads (5uL beads per 1ug of antibody) at RT for 15 minutes. After conjugation cells were washed twice with lysis buffer and once with TE. Decrosslinking than occurred via 30 minute treatment with RNAseA (100ug/mL) at 37C, followed by overnight digestion of protein via proteinase K (1mg/mL) at 65C.

Library preparation and sequencing:

DNA from chromatin preps was extracted via phenol: chloroform extraction and then purified and concentrated via EtOH cleanup. Resulting DNA was quantified via Qubit. 1 to 5 ng of DNA were used to create sequencing libraries via NEB Ultra II DNA Kits. After amplification, resulting libraries were checked for size (~300bp) via either Bioanalyzer 2100 or TapeStation 4200. Samples were then pooled at equal molar amounts of library and concentrated/ size selected via AMPure beads. Library was then run on NEXT-Seq 500.

ChIP-Seq read processing and peak calling:

Reads from sequencing were trimmed using trim_galore (--paired --length 24 --stringency 3) and aligned to GRCm39 using bowtie2 (--no-mixed –no discordant)(*71*). Aligned reads were then trimmed for duplicates and quality using samtools (samtools view -f 2 -q 20 -h 473-1.sam | samtools sort -n -o $QualFilt -O bam; samtools fixmate -m $QualFilt $fixmate; samtools sort -O bam -o $fixsorted $fixmate; samtools markdup -r $fixsorted $dupremoved)(*72*). Peaks were then called both per replicate and for all replicates via macs 2(callpeak -t $duprembam -c $controlbam --broad -f BAMPE -g mm -q 0.05 -B)(*73*)

ChIP-Seq data processing:

To determine significant differences between given peaks, the r package DiffBind was used(*74*). In the context of replicate controlled samples (shCon/shMll1 H3K4me3 and RNApol2) an additional replicate factor was added as part of the design. Peaks with a pvalue <=0.05 were considered significant for this analysis.

To determine gene enrichment, peaks from the given sample were then annotated to the nearest gene using ChIPseeker (*75*). The number of these genes that overlapped with previously described gene sets were counted and the simple fraction of the whole of the gene set was visualized using ComplexHeatmap (*3, 9, 10, 42, 51, 54, 55, 65, 76*).

To compute per gene log2FCs of H3K4me3 binding, the reported peaks from Diffbind were collapsed to gene and log2fcs from multiple peaks per gene were averaged. (*74*). Using these and the Log2FCs computed via DESEQ2 for the previously published naRNA-seq, data values were plotted on a two way dot plot(*9, 68*). Genes were marked by their presence in previously described gene clusters that were differential across time(*54*).

To compute the enrichment of one set of peaks within another, bedtools was used to determine the number intersecting peaks. The fraction of intersecting peaks over total number was then computed.

To calculate the pausing rate we adapted a previously published strategy (*77*). Annotated genes were separated into their promoters (TSS +/ 300b.p) and gene bodies (the gene minus the promoter region) and the number of reads in either region were calculated. These were converted to transcripts per million and the sum of TPMs to promoter or gene body were calculated and averaged across replicates. The per gene promoter sum and the per gene body sum were then divided to form the pausing index. Differences between indices were calculated via KS test

Bibliography:

1. S. Tsuda, M. E. Pipkin, Transcriptional Control of Cell Fate Determination in Antigen-Experienced CD8 T Cells. *Cold Spring Harbor perspectives in biology*, (2021).
2. H. Diao, M. Pipkin, Stability and flexibility in chromatin structure and transcription underlies memory CD8 T-cell differentiation [version 1; peer review: 2 approved]. *F1000Research* **8**, (2019).
3. D. Pais Ferreira *et al.*, Central memory CD8(+) T cells derive from stem-like Tcf7(hi) effector cells in the absence of cytotoxic differentiation. *Immunity* **53**, 985-1000.e1011 (2020).
4. J. B. Johnnidis *et al.*, Inhibitory signaling sustains a distinct early memory CD8(+) T cell precursor that is resistant to DNA damage. *Sci Immunol* **6**, (2021).
5. D. T. Utzschneider *et al.*, T Cell Factor 1-Expressing Memory-like CD8+ T Cells Sustain the Immune Response to Chronic Viral Infections. *Immunity* **45**, 415-427 (2016).
6. S. J. Im *et al.*, Defining CD8+ T cells that provide the proliferative burst after PD-1 therapy. *Nature* **537**, 417-421 (2016).
7. R. He *et al.*, Follicular CXCR5-expressing CD8+ T cells curtail chronic viral infection. *Nature* **537**, 412-416 (2016).
8. Q. Huang *et al.*, The primordial differentiation of tumor-specific memory CD8⁺ T cells as bona fide responders to PD-1/PD-L1 blockade in draining lymph nodes. *Cell* **185**, 4049-4066.e4025 (2022).
9. H. Diao *et al.*, Single-cell lineage trajectories and chromatin regulators that initialize antiviral CD8 T cell ontogeny. *bioRxiv*, 2021.2008.2011.456014 (2021).
10. C. Yao *et al.*, Single-cell RNA-seq reveals TOX as a key regulator of CD8+ T cell persistence in chronic infection. *Nature immunology* **20**, 890-901 (2019).
11. B. Kakaradov *et al.*, Early transcriptional and epigenetic regulation of CD8+ T cell differentiation revealed by single-cell RNA sequencing. *Nature immunology* **18**, 422-432 (2017).
12. D. T. Utzschneider *et al.*, Early precursor T cells establish and propagate T cell exhaustion in chronic infection. *Nature immunology* **21**, 1256-1266 (2020).
13. E. Ahn *et al.*, Role of PD-1 during effector CD8 T cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 4749-4754 (2018).
14. R. Chen *et al.*, In vivo RNA interference screens identify regulators of antiviral CD4(+) and CD8(+) T cell differentiation. *Immunity* **41**, 325-338 (2014).
15. S. Denissov *et al.*, Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **141**, 526-537 (2014).
16. Y. Chen *et al.*, MLL2, Not MLL1, Plays a Major Role in Sustaining MLL-Rearranged Acute Myeloid Leukemia. *Cancer cell* **31**, 755-770.e756 (2017).
17. Y. Ji *et al.*, Repression of the DNA-binding inhibitor Id3 by Blimp-1 limits the formation of memory CD8+ T cells. *Nature immunology* **12**, 1230-1237 (2011).
18. B. C. Miller *et al.*, Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade. *Nature immunology* **20**, 326-336 (2019).
19. J.-C. Beltra *et al.*, Developmental Relationships of Four Exhausted CD8+ T Cell Subsets Reveals Underlying Transcriptional and Epigenetic Landscape Control Mechanisms. *Immunity* **52**, 825-841.e828 (2020).
20. W. H. Hudson *et al.*, Proliferating Transitory T Cells with an Effector-like Transcriptional Signature Emerge from PD-1+ Stem-like CD8+ T Cells during Chronic Infection. *Immunity* **51**, 1043-1058.e1044 (2019).
21. R. Zander *et al.*, CD4(+) T Cell Help Is Required for the Formation of a Cytolytic CD8(+) T Cell Subset that Protects against Chronic Infection and Cancer. *Immunity* **51**, 1028-1042.e1024 (2019).
22. E. J. Wherry, J. N. Blattman, K. Murali-Krishna, R. van der Most, R. Ahmed, Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *Journal of virology* **77**, 4911-4927 (2003).
23. R. S. Akondy *et al.*, Origin and differentiation of human memory CD8 T cells after vaccination. *Nature* **552**, 362-367 (2017).
24. S. M. Kaech, S. Hemby, E. Kersh, R. Ahmed, Molecular and functional profiling of memory CD8 T cell differentiation. *Cell* **111**, 837-851 (2002).
25. E. J. Wherry, R. Ahmed, Memory CD8 T-cell differentiation during viral infection. *Journal of virology* **78**, 5535-5545 (2004).
26. C. Y. Yang *et al.*, The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. *Nature immunology* **12**, 1221-1229 (2011).
27. W.-Hsuan W. Lin *et al.*, CD8+ T Lymphocyte Self-Renewal during Effector Cell Determination. *Cell Reports* **17**, 1773-1782 (2016).
28. V. Kalia *et al.*, Prolonged interleukin-2Ralpha expression on virus-specific CD8+ T cells favors terminal-effector differentiation in vivo. *Immunity* **32**, 91-103 (2010).
29. N. S. Joshi *et al.*, 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 (2007).
30. C. Gerlach *et al.*, The chemokine receptor CX3CR1 defines three antigen-experienced CD8 T cell subsets with distinct roles in immune surveillance and homeostasis. *Immunity* **45**, 1270-1284 (2016).

31. J. J. Obar, K. M. Khanna, L. Lefrançois, Endogenous naive CD8+ T cell precursor frequency regulates primary and memory responses to infection. *Immunity* **28**, 859-869 (2008).
32. J. A. Olson, C. McDonald-Hyman, S. C. Jameson, S. E. Hamilton, Effector-like CD8(+) T cells in the memory population mediate potent protective immunity. *Immunity* **38**, 1250-1260 (2013).
33. X. Zhou *et al.*, Differentiation and Persistence of Memory CD8+ T Cells Depend on T Cell Factor 1. *Immunity* **33**, 229-240 (2010).
34. S. M. Gray, R. A. Amezquita, T. Guan, S. H. Kleinstein, S. M. Kaech, Polycomb Repressive Complex 2-Mediated Chromatin Repression Guides Effector CD8+ T Cell Terminal Differentiation and Loss of Multipotency. *Immunity* **46**, 596-608 (2017).
35. V. R. Buchholz *et al.*, Disparate Individual Fates Compose Robust CD8⁺ T Cell Immunity. *Science* **340**, 630-635 (2013).
36. L. Haghverdi, A. T. L. Lun, M. D. Morgan, J. C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology* **36**, 421-427 (2018).
37. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology* **38**, 1408-1414 (2020).
38. D. Herndler-Brandstetter *et al.*, KLRG1(+) Effector CD8(+) T Cells Lose KLRG1, Differentiate into All Memory T Cell Lineages, and Convey Enhanced Protective Immunity. *Immunity* **48**, 716-729.e718 (2018).
39. L. K. Mackay *et al.*, Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science* **352**, 459-463 (2016).
40. B. M. Sullivan, A. Juedes, S. J. Szabo, M. von Herrath, L. H. Glimcher, Antigen-driven effector CD8 T cell function regulated by T-bet. *Proceedings of the National Academy of Sciences* **100**, 15818-15823 (2003).
41. R. L. Rutishauser *et al.*, Transcriptional repressor Blimp-1 promotes CD8(+) T cell terminal differentiation and represses the acquisition of central memory T cell properties. *Immunity* **31**, 296-308 (2009).
42. C. X. Dominguez *et al.*, The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *The Journal of experimental medicine* **212**, 2041-2056 (2015).
43. M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, R. A. Young, A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell* **130**, 77-88 (2007).
44. B. E. Russ *et al.*, Regulation of H3K4me3 at Transcriptional Enhancers Characterizes Acquisition of Virus-Specific CD8⁺ T Cell-Lineage-Specific Function. *Cell Reports* **21**, 3624-3636 (2017).
45. B. E. Russ *et al.*, Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8(+) T cell differentiation. *Immunity* **41**, 853-865 (2014).
46. B. Yu *et al.*, Epigenetic landscapes reveal transcription factors that regulate CD8(+) T cell differentiation. *Nature immunology* **18**, 573-582 (2017).
47. T. A. Milne *et al.*, MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters. *Molecular Cell* **10**, 1107-1117 (2002).
48. T. Nakamura *et al.*, ALL-1 Is a Histone Methyltransferase that Assembles a Supercomplex of Proteins Involved in Transcriptional Regulation. *Molecular Cell* **10**, 1119-1128 (2002).
49. T. A. Milne *et al.*, Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Molecular cell* **38**, 853-863 (2010).
50. B. P. Mishra *et al.*, The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. *Cell reports* **7**, 1239-1247 (2014).
51. D. Wang *et al.*, The Transcription Factor Runx3 Establishes Chromatin Accessibility of cis-Regulatory Landscapes that Drive Memory Cytotoxic T Lymphocyte Formation. *Immunity* **48**, 659-674.e656 (2018).
52. Simon R. V. Knott *et al.*, A Computational Algorithm to Predict shRNA Potency. *Molecular Cell* **56**, 796-807.
53. M. E. Pipkin *et al.*, Interleukin-2 and inflammation induce distinct transcriptional programs that promote the differentiation of effector cytolytic T cells. *Immunity* **32**, 79-90 (2010).
54. J. A. Best *et al.*, Transcriptional insights into the CD8(+) T cell response to infection and memory T cell formation. *Nature immunology* **14**, 404-412 (2013).
55. A. Xin *et al.*, A molecular threshold for effector CD8(+) T cell differentiation controlled by transcription factors Blimp-1 and T-bet. *Nature immunology* **17**, 422-432 (2016).
56. C. Fellmann *et al.*, An Optimized microRNA Backbone for Effective Single-Copy RNAi. *Cell Reports* **5**, 1704-1713 (2013).
57. K. Labun *et al.*, CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research* **47**, W171-W174 (2019).
58. Q. Yang *et al.*, TCF-1 upregulation identifies early innate lymphoid progenitors in the bone marrow. *Nature immunology* **16**, 1044-1050 (2015).
59. R. L. Rutishauser *et al.*, Transcriptional Repressor Blimp-1 Promotes CD8+ T Cell Terminal Differentiation and Represses the Acquisition of Central Memory T Cell Properties. *Immunity* **31**, 296-308 (2009).
60. H. Pircher, K. Burki, R. Lang, H. Hengartner, R. M. Zinkernagel, Tolerance induction in double specific T-cell receptor transgenic mice varies with antigen. *Nature* **342**, 559-561 (1989).
61. P. Mombaerts *et al.*, Mutations in T-cell antigen receptor genes α and β block thymocyte development at different stages. *Nature* **360**, 225-231 (1992).
62. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16**, 284-287 (2012).

63. R. Ahmed, A. Salmi, L. D. Butler, J. M. Chiller, M. B. Oldstone, Selection of genetic variants of lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence. *The Journal of experimental medicine* **160**, 521-540 (1984).
64. J. J. Milner *et al.*, Bromodomain protein BRD4 directs and sustains CD8 T cell differentiation during infection. *The Journal of experimental medicine* **218**, (2021).
65. J. J. Milner *et al.*, Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumours. *Nature*, (2017).
66. L. B. Rodda *et al.*, Single-Cell RNA Sequencing of Lymph Node Stromal Cells Reveals Niche-Associated Heterogeneity. *Immunity* **48**, 1014-1028.e1016 (2018).
67. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nature methods* **14**, 417-419 (2017).
68. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
69. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15 (2018).
70. F. A. Wolf *et al.*, PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biology* **20**, 59 (2019).
71. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-359 (2012).
72. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, (2021).
73. Y. Zhang *et al.*, Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137 (2008).
74. C. S. Ross-Innes *et al.*, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393 (2012).
75. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).
76. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
77. P. B. Rahl *et al.*, c-Myc regulates transcriptional pause release. *Cell* **141**, 432-445 (2010).

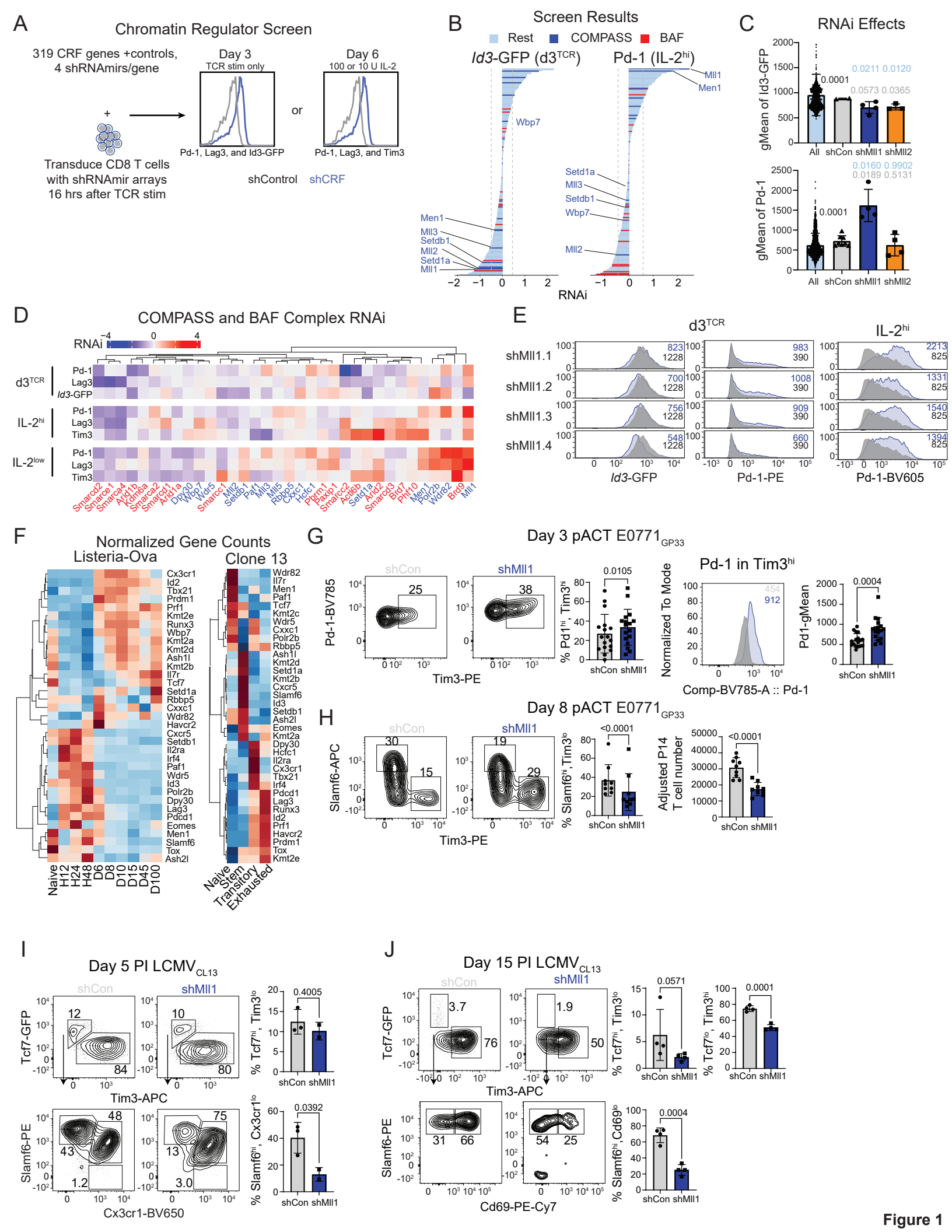

A
Chromatin Regulator Screen
319 CRF genes +controls,
4 shRNAmirs/gene
Transduce CD8 T cells
with shRNAmir arrays
16 hrs after TCR stim
Day 3
TCR stim only
Pd-1, Lag3, and Id3-GFP
or
Day 6
100 or 10 U IL-2
Pd-1, Lag3, and Tim3
shControl
shCRF
B
Screen Results
Rest
COMPASS
BAF
Id3-GFP (d3TCR)
Pd-1 (IL-2hi)
Wbp7
Men1
Mll3
Setdb1
Mll2
Setd1a
Mll1
RNAi
C
RNAi Effects
gMean of Id3-GFP
gMean of Pd-1
All
shCon
shMll1
shMll2
0.0001
0.0211
0.0120
0.0573
0.0365
0.0160
0.9902
0.0189
0.5131
D
COMPASS and BAF Complex RNAi
RNAi
d3TCR
IL-2hi
IL-2low
Pd-1
Lag3
Id3-GFP
Tim3
Smarcd2
Smarce1
Smarca4
Arid1b
Kdm6a
Smarca2
Smarcd1
Arid1a
Dpy30
Wbp7
Wdr5
Smarcc1
Mll2
Setdb1
Paf1
Mll3
Mll5
Rbbp5
Cxxc1
Hcfc1
Prrm1
Paxip1
Smarcc2
Actl6b
Setd1a
Arid2
Smarcd3
Brd7
Phf10
Men1
Polr2b
Wdr82
Brd9
Mll1
E
d3TCR
IL-2hi
shMll1.1
shMll1.2
shMll1.3
shMll1.4
Id3-GFP
Pd-1-PE
Pd-1-BV605
F
Normalized Gene Counts
Listeria-Ova
Clone 13
Naive
H12
H24
H48
D6
D8
D10
D15
D45
D100
Stem
Transitory
Exhausted
G
Day 3 pACT E0771GP33
shCon
shMll1
Pd-1-BV785
Tim3-PE
% Pd1hi, Tim3hi
Pd-1 in Tim3hi
Normalized To Mode
Comp-BV785-A :: Pd-1
Pd1-gMean
H
Day 8 pACT E0771GP33
Slamf6-APC
% Slamf6hi, Tim3lo
Adjusted P14
T cell number
I
Day 5 PI LCMVCL13
Tcf7-GFP
Tim3-APC
% Tcf7hi, Tim3lo
Slamf6-PE
Cx3cr1-BV650
% Slamf6hi, Cx3cr1lo
J
Day 15 PI LCMVCL13
% Tcf7hi, Tim3lo
% Tcf7lo, Tim3hi
Cd69-PE-Cy7
% Slamf6hi, Cd69lo

Figure 1

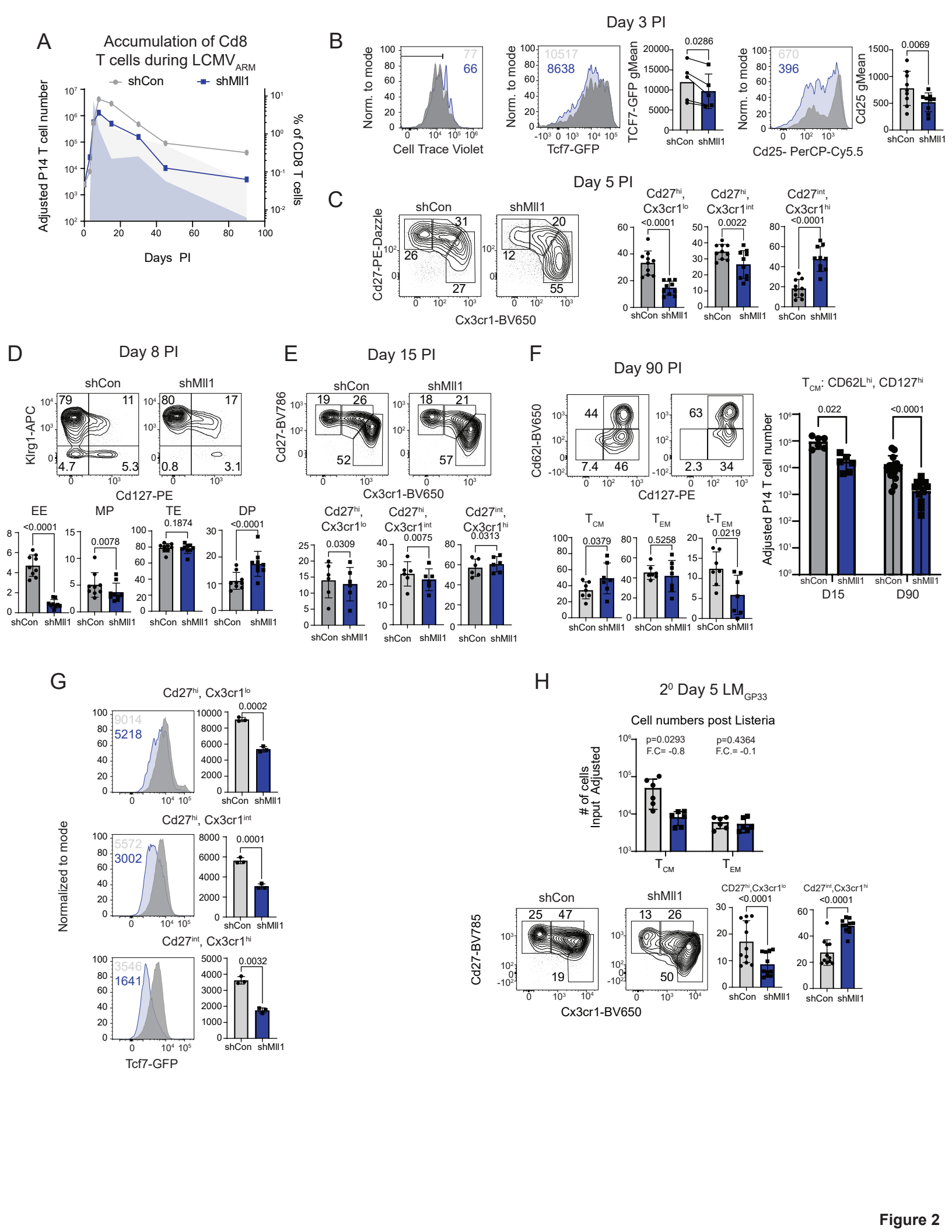

A
Accumulation of Cd8 T cells during LCMV ARM
shCon
shMll1
Adjusted P14 T cell number
% of CD8 T cells
Days PI
B
Day 3 PI
Norm. to mode
Cell Trace Violet
Tcf7-GFP
TCF7-GFP gMean
Cd25- PerCP-Cy5.5
Cd25 gMean
0.0286
0.0069
C
Day 5 PI
Cd27-PE-Dazzle
Cx3cr1-BV650
Cd27hi, Cx3cr1lo
Cd27hi, Cx3cr1int
Cd27int, Cx3cr1hi
<0.0001
0.0022
D
Day 8 PI
Klrg1-APC
Cd127-PE
EE
MP
TE
DP
0.1874
0.0078
E
Day 15 PI
Cd27-BV786
Cx3cr1-BV650
0.0309
0.0075
0.0313
F
Day 90 PI
Cd62l-BV650
Cd127-PE
TCM
TEM
t-TEM
0.0379
0.5258
0.0219
TCM: CD62Lhi, CD127hi
Adjusted P14 T cell number
0.022
D15
D90
G
Cd27hi, Cx3cr1lo
Cd27hi, Cx3cr1int
Cd27int, Cx3cr1hi
Normalized to mode
Tcf7-GFP
0.0002
0.0001
0.0032
H
2° Day 5 LMGP33
Cell numbers post Listeria
# of cells Input Adjusted
p=0.0293
F.C= -0.8
p=0.4364
F.C.= -0.1
Cd27-BV785
Cx3cr1-BV650

Figure 2

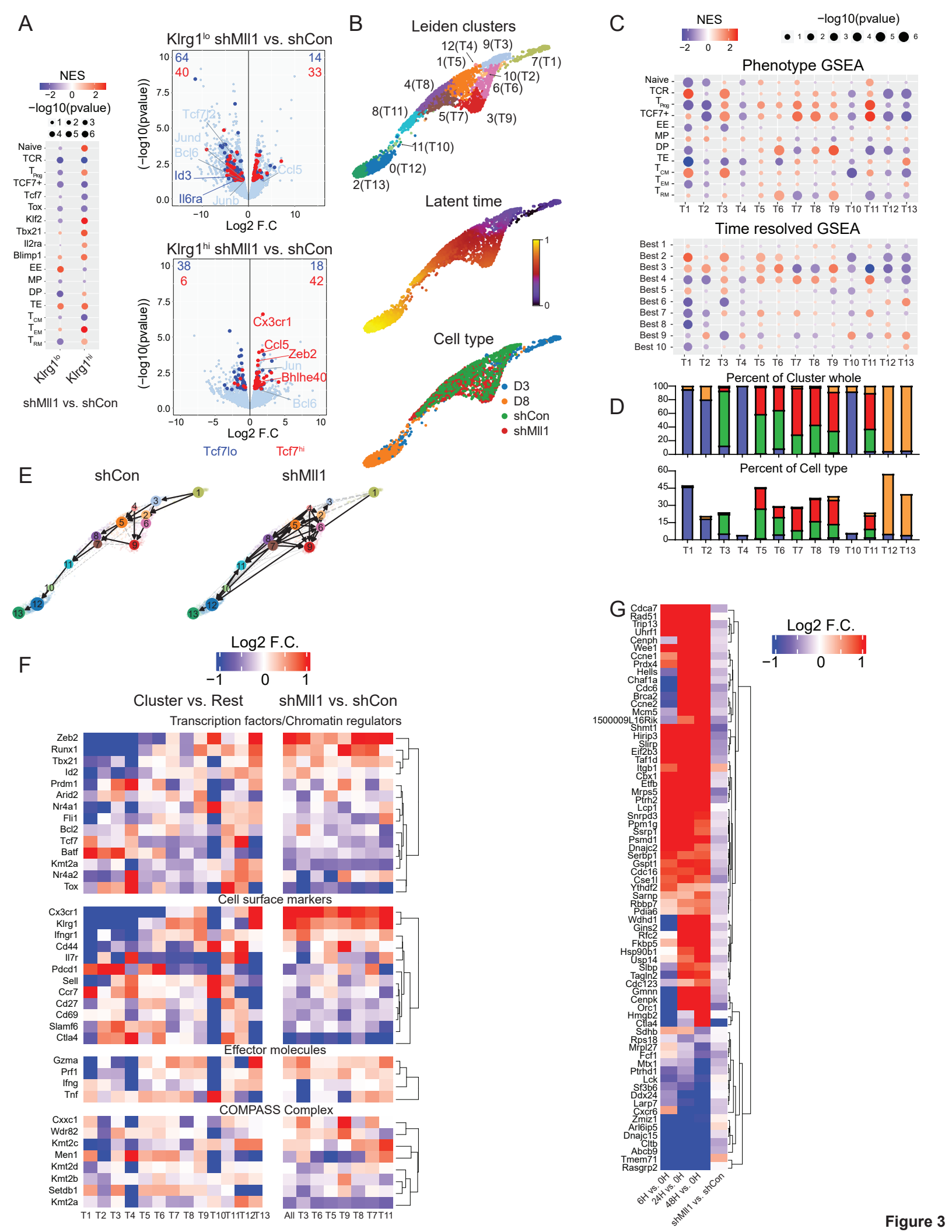
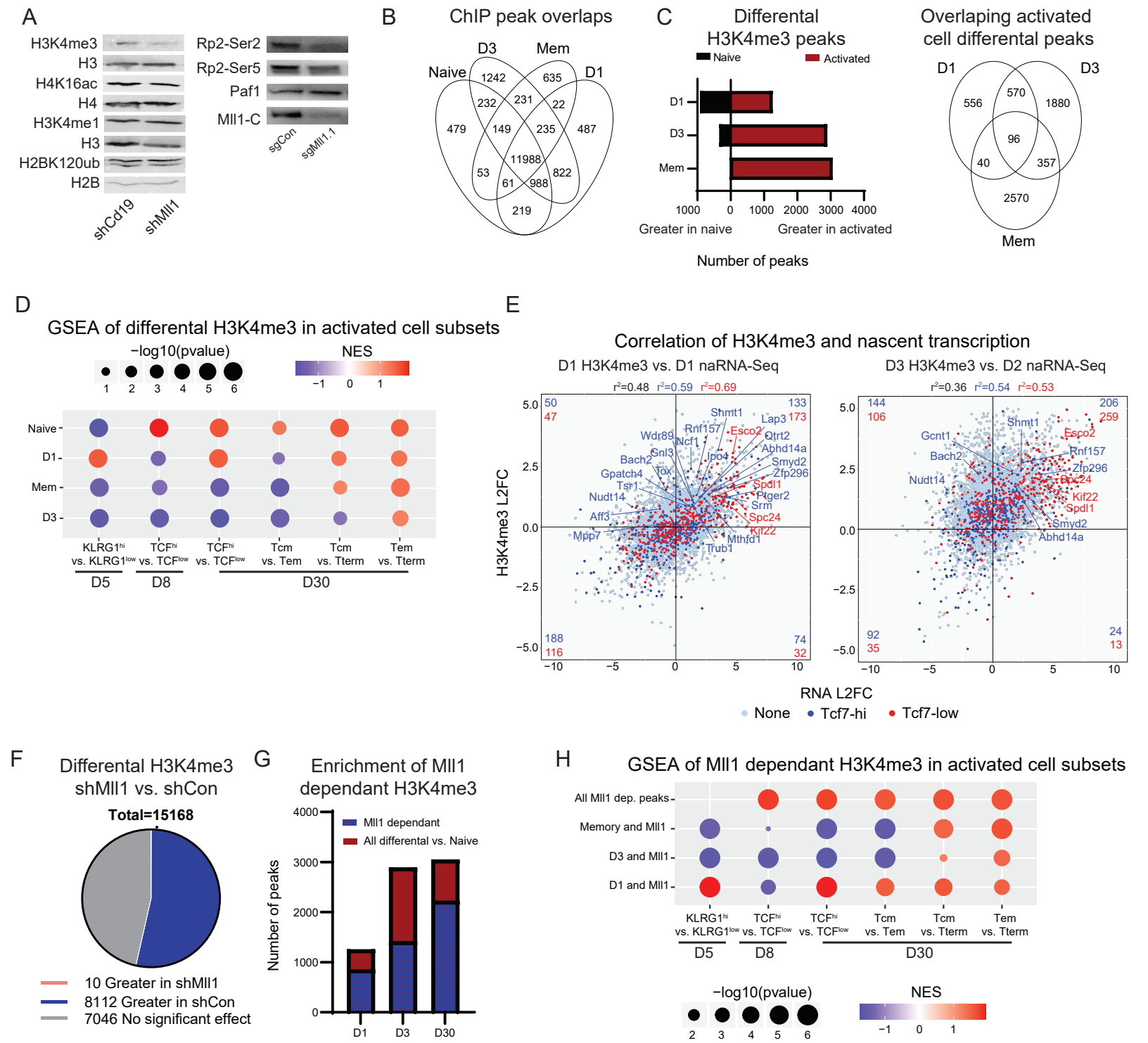
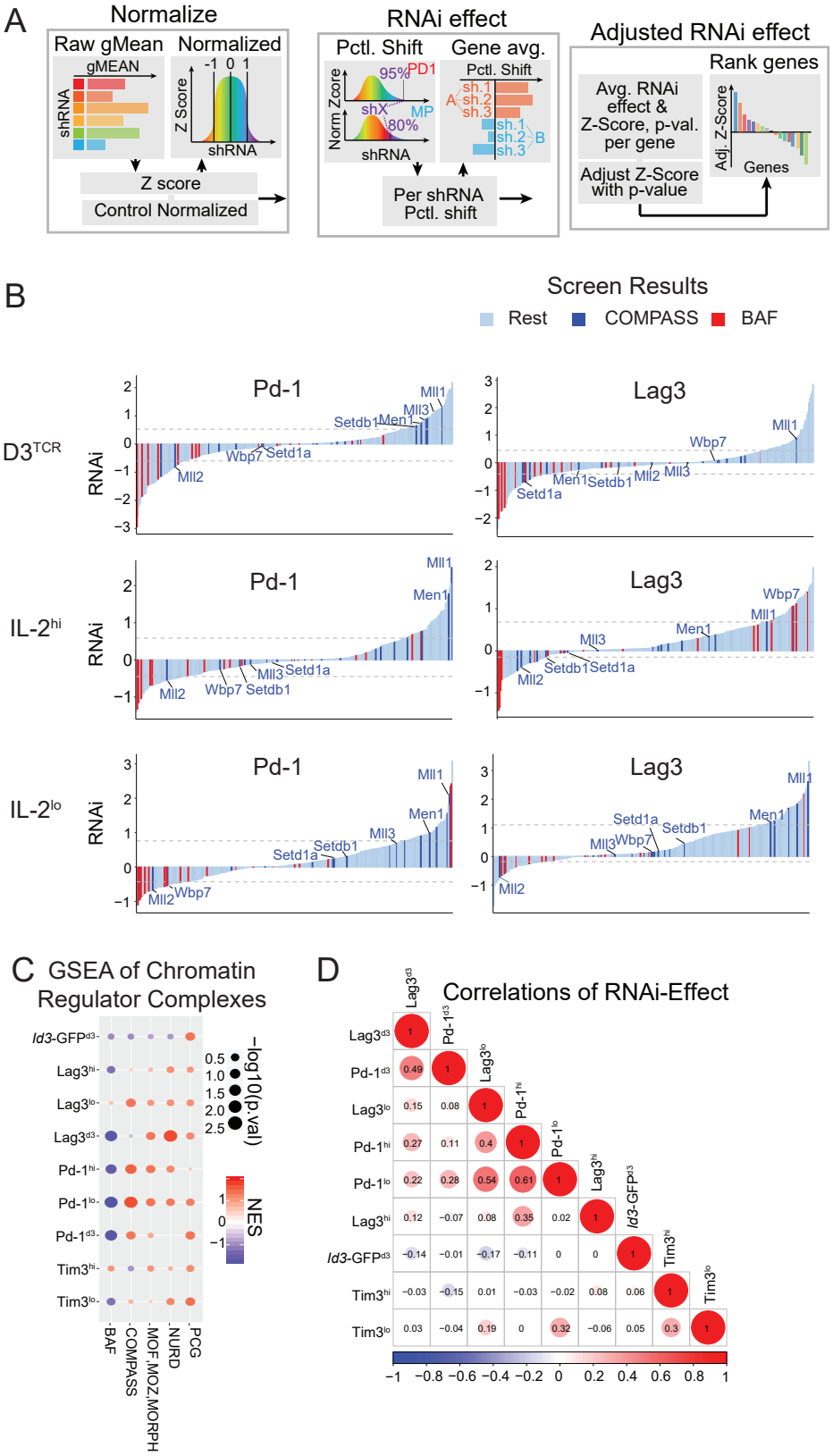
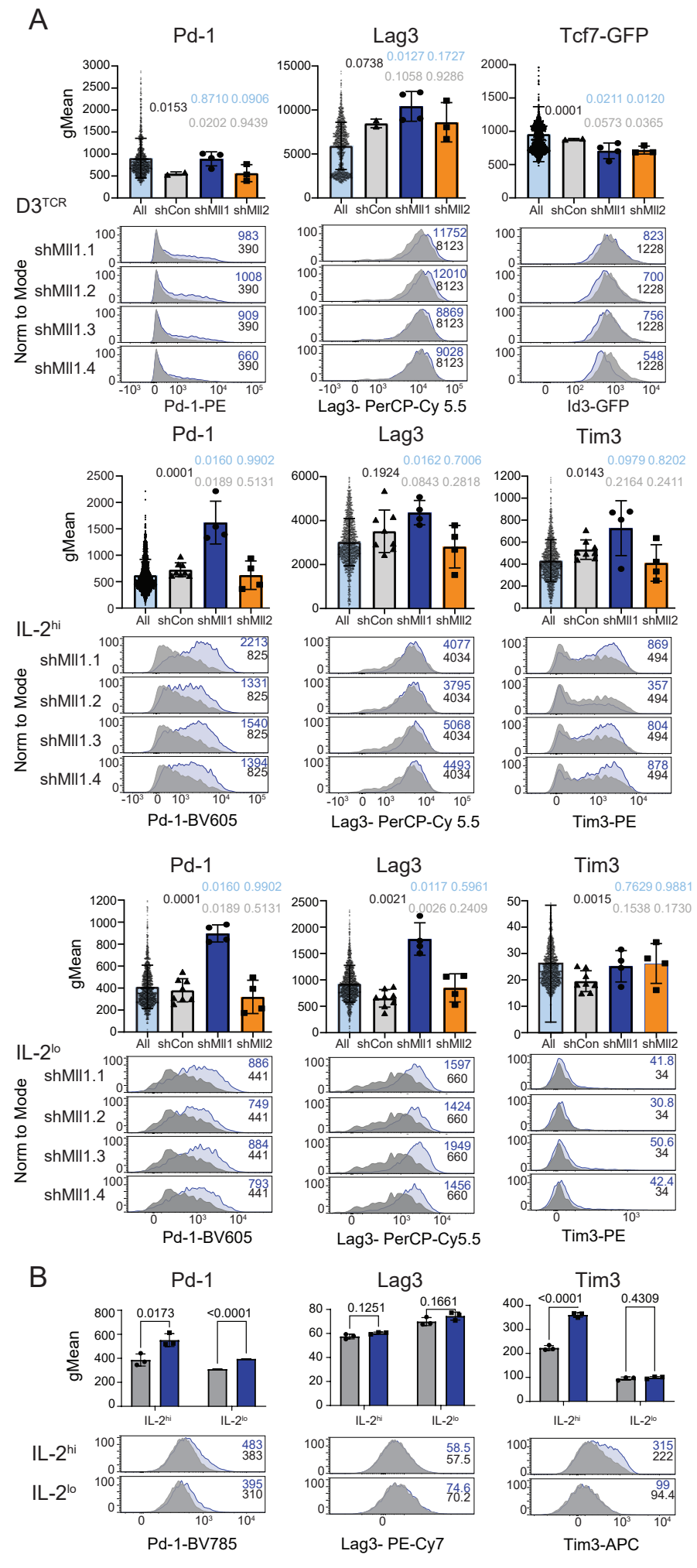

A
NES
−2 0 2
−log10(pvalue)
1 2 3 4 5 6
Naive
TCR
TCF7+
Tcf7
Tox
Klf2
Tbx21
Il2ra
Blimp1
EE
MP
DP
TE
Klrg1lo
Klrg1hi
shMll1 vs. shCon
Klrg1lo shMll1 vs. shCon
64
40
14
33
Tcf7l2
Jund
Bcl6
Id3
Il6ra
Junb
Ccl5
(−log10(pvalue))
Log2 F.C
Klrg1hi shMll1 vs. shCon
38
6
18
42
Cx3cr1
Ccl5
Zeb2
Jun
Bhlhe40
Bcl6
Tcf7lo
Tcf7hi
B
Leiden clusters
12(T4)
9(T3)
1(T5)
7(T1)
10(T2)
6(T6)
4(T8)
8(T11)
5(T7)
3(T9)
11(T10)
0(T12)
2(T13)
Latent time
Cell type
D3
D8
shCon
shMll1
C
NES
−2 0 2
−log10(pvalue)
1 2 3 4 5 6
Phenotype GSEA
Naive
TCR
TCF7+
EE
MP
DP
TE
T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12 T13
Time resolved GSEA
Best 1
Best 2
Best 3
Best 4
Best 5
Best 6
Best 7
Best 8
Best 9
Best 10
D
Percent of Cluster whole
Percent of Cell type
E
shCon
shMll1
F
Log2 F.C.
−1 0 1
Cluster vs. Rest
shMll1 vs. shCon
Transcription factors/Chromatin regulators
Zeb2
Runx1
Tbx21
Id2
Prdm1
Arid2
Nr4a1
Fli1
Bcl2
Tcf7
Batf
Kmt2a
Nr4a2
Tox
Cell surface markers
Cx3cr1
Klrg1
Ifngr1
Cd44
Il7r
Pdcd1
Sell
Ccr7
Cd27
Cd69
Slamf6
Ctla4
Effector molecules
Gzma
Prf1
Ifng
Tnf
COMPASS Complex
Cxxc1
Wdr82
Kmt2c
Men1
Kmt2d
Kmt2b
Setdb1
Kmt2a
T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12 T13
All T3 T6 T5 T9 T8 T7 T11
G
Log2 F.C.
−1 0 1
Cdca7
Rad51
Trip13
Uhrf1
Cenph
Wee1
Ccne1
Prdx4
Hells
Chaf1a
Cdc6
Brca2
Ccne2
Mcm5
1500009L16Rik
Shmt1
Hirip3
Slirp
Eif2b3
Taf1d
Itgb1
Cbx1
Etfb
Mrps5
Ptrh2
Lcp1
Snrpd3
Ppm1g
Ssrp1
Psmd1
Dnajc2
Serbp1
Gspt1
Cdc16
Cse1l
Ythdf2
Sarnp
Rbbp7
Pdia6
Wdhd1
Gins2
Rfc2
Fkbp5
Hsp90b1
Usp14
Slbp
Tagln2
Cdc123
Gmnn
Cenpk
Orc1
Hmgb2
Ctla4
Sdhb
Rps18
Mrpl27
Fcf1
Mtx1
Ptrhd1
Lck
Sf3b6
Ddx24
Larp7
Cxcr6
Zmiz1
Arl6ip5
Dnajc15
Cltb
Abcb9
Tmem71
Rasgrp2
6H vs. 0H
24H vs. 0H
48H vs. 0H
shMll1 vs. shCon

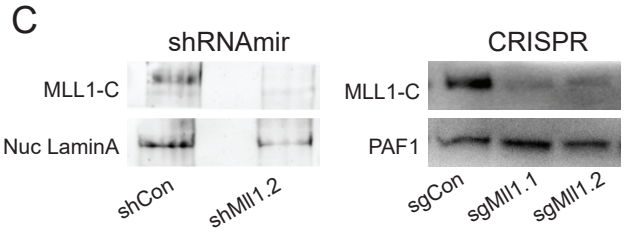
Figure 3

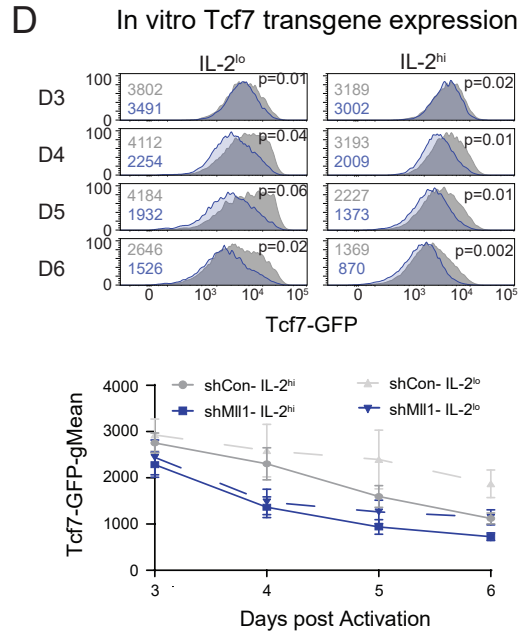

A
H3K4me3
H3
H4K16ac
H4
H3K4me1
H3
H2BK120ub
H2B
shCd19
shMll1
Rp2-Ser2
Rp2-Ser5
Paf1
Mll1-C
sgCon
sgMll1.1
B
ChIP peak overlaps
D3
Mem
Naive
D1
1242
635
232
231
22
479
149
235
487
11988
53
822
61
988
219
C
Differential H3K4me3 peaks
Naive
Activated
D1
D3
Mem
1000
0
1000
2000
3000
4000
Greater in naive
Greater in activated
Number of peaks
Overlaping activated cell differental peaks
D1
D3
556
570
1880
96
40
357
2570
Mem
D
GSEA of differental H3K4me3 in activated cell subsets
−log10(pvalue)
1
2
3
4
5
6
NES
−1
0
1
Naive
D1
Mem
D3
KLRG1hi vs. KLRG1low
TCFhi vs. TCFlow
TCFhi vs. TCFlow
Tcm vs. Tem
Tcm vs. Tterm
Tem vs. Tterm
D5
D8
D30
E
Correlation of H3K4me3 and nascent transcription
D1 H3K4me3 vs. D1 naRNA-Seq
r²=0.48 r²=0.59 r²=0.69
D3 H3K4me3 vs. D2 naRNA-Seq
r²=0.36 r²=0.54 r²=0.53
H3K4me3 L2FC
RNA L2FC
5.0
2.5
0.0
−2.5
−5.0
−10
−5
0
5
10
50
47
133
173
188
116
74
32
144
106
206
259
92
35
24
13
Shmt1
Lap3
Rnf157
Esco2
Wdr89
Ncf1
Qtrt2
Abhd14a
Bach2
Gnl3
Ipo4
Smyd2
Gpatch4
Tox
Zfp296
Tsr1
Spdl1
Nudt14
Piger2
Aff3
Srm
Spc24
Mpp7
Kif22
Mthfd1
Trub1
Gcnt1
Bach2
Rnf157
Nudt14
Zfp296
Spc24
Kif22
Spdl1
Smyd2
Abhd14a
None
Tcf7-hi
Tcf7-low
F
Differental H3K4me3 shMll1 vs. shCon
Total=15168
10 Greater in shMll1
8112 Greater in shCon
7046 No significant effect
G
Enrichment of Mll1 dependant H3K4me3
Mll1 dependant
All differental vs. Naive
Number of peaks
4000
3000
2000
1000
0
D1
D3
D30
H
GSEA of Mll1 dependant H3K4me3 in activated cell subsets
All Mll1 dep. peaks
Memory and Mll1
D3 and Mll1
D1 and Mll1
KLRG1hi vs. KLRG1low
TCFhi vs. TCFlow
TCFhi vs. TCFlow
Tcm vs. Tem
Tcm vs. Tterm
Tem vs. Tterm
D5
D8
D30
−log10(pvalue)
2
3
4
5
6
NES
−1
0
1

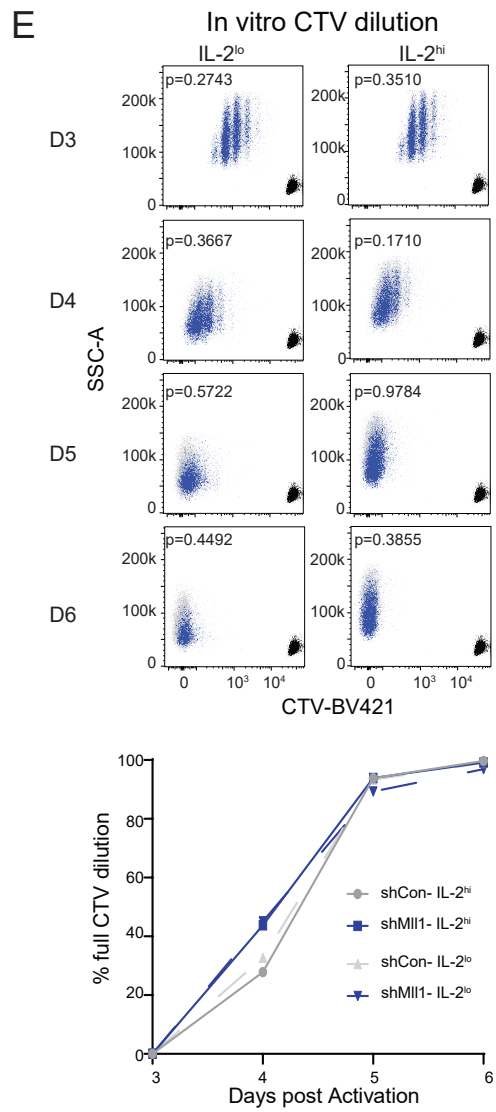

A
Normalize
Raw gMean
Normalized
gMEAN
shRNA
Z Score
shRNA
Z score
Control Normalized
RNAi effect
Pctl. Shift
Gene avg.
Norm Zcore
95%
PD1
shX
MP
80%
shRNA
Pctl. Shift
Per shRNA Pctl. shift
Adjusted RNAi effect
Avg. RNAi effect & Z-Score, p-val. per gene
Adjust Z-Score with p-value
Rank genes
Adj. Z-Score
Genes
B
Screen Results
Rest
COMPASS
BAF
D3TCR
IL-2hi
IL-2lo
RNAi
Pd-1
Lag3
Id3-GFP
Tim3
C
GSEA of Chromatin Regulator Complexes
Id3-GFPd3-
Lag3hi-
Lag3lo-
Lag3d3-
Pd-1hi-
Pd-1lo-
Pd-1d3-
Tim3hi-
Tim3lo-
BAF
COMPASS
MOF,MOZ,MORPH
NURD
PCG
-log10(p.val)
NES
D
Correlations of RNAi-Effect

Sup. Figure 1

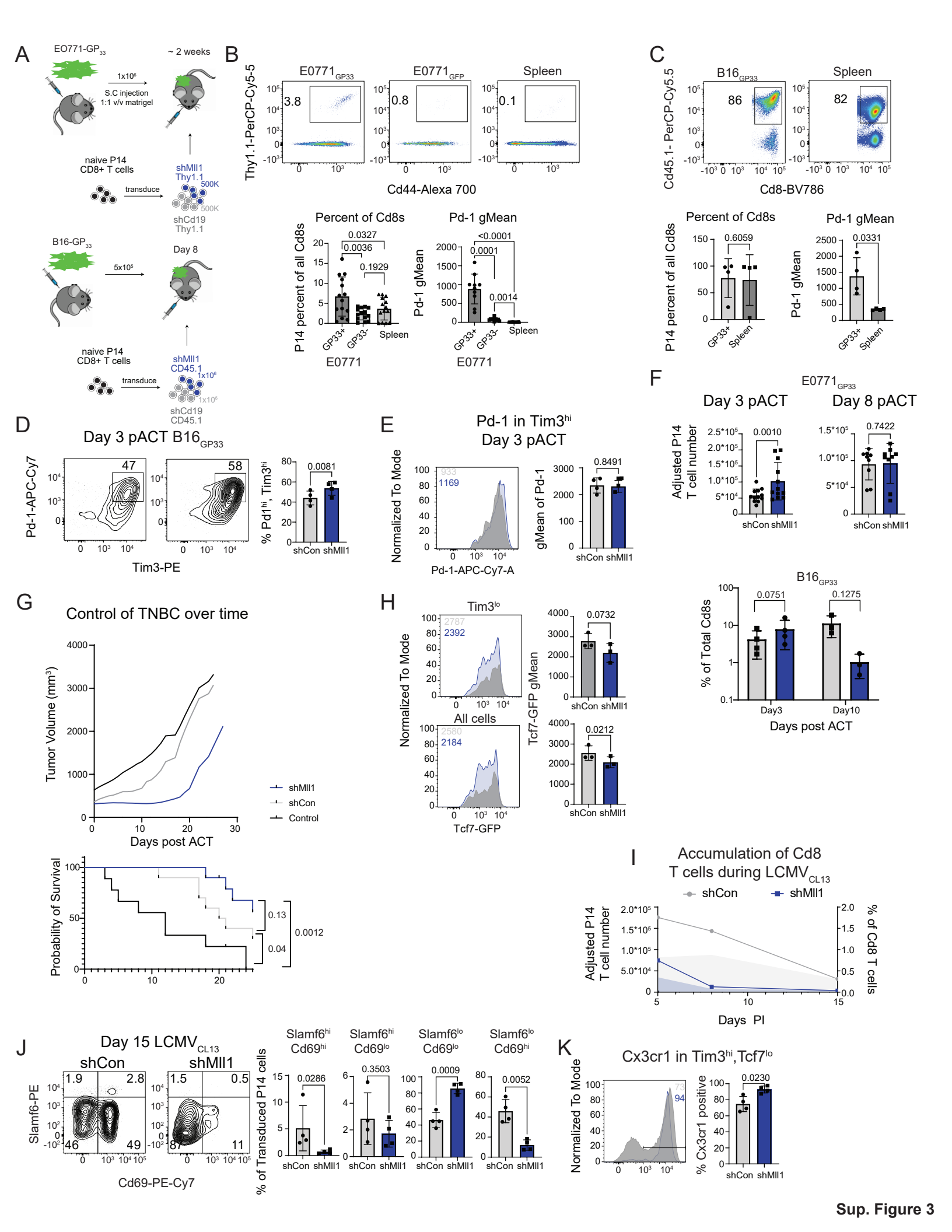

A
Pd-1
Lag3
Tcf7-GFP
gMean
D3TCR
All shCon shMll1 shMll2
shMll1.1
shMll1.2
shMll1.3
shMll1.4
Norm to Mode
Pd-1-PE
Lag3- PerCP-Cy 5.5
Id3-GFP
Tim3
IL-2hi
Pd-1-BV605
Tim3-PE
IL-2lo
Pd-1-BV605
Lag3- PerCP-Cy5.5
B
IL-2hi
IL-2lo
Pd-1-BV785
Lag3- PE-Cy7
Tim3-APC


C
shRNAmir
CRISPR
MLL1-C
Nuc LaminA
PAF1
shCon
shMll1.2
sgCon
sgMll1.1
sgMll1.2

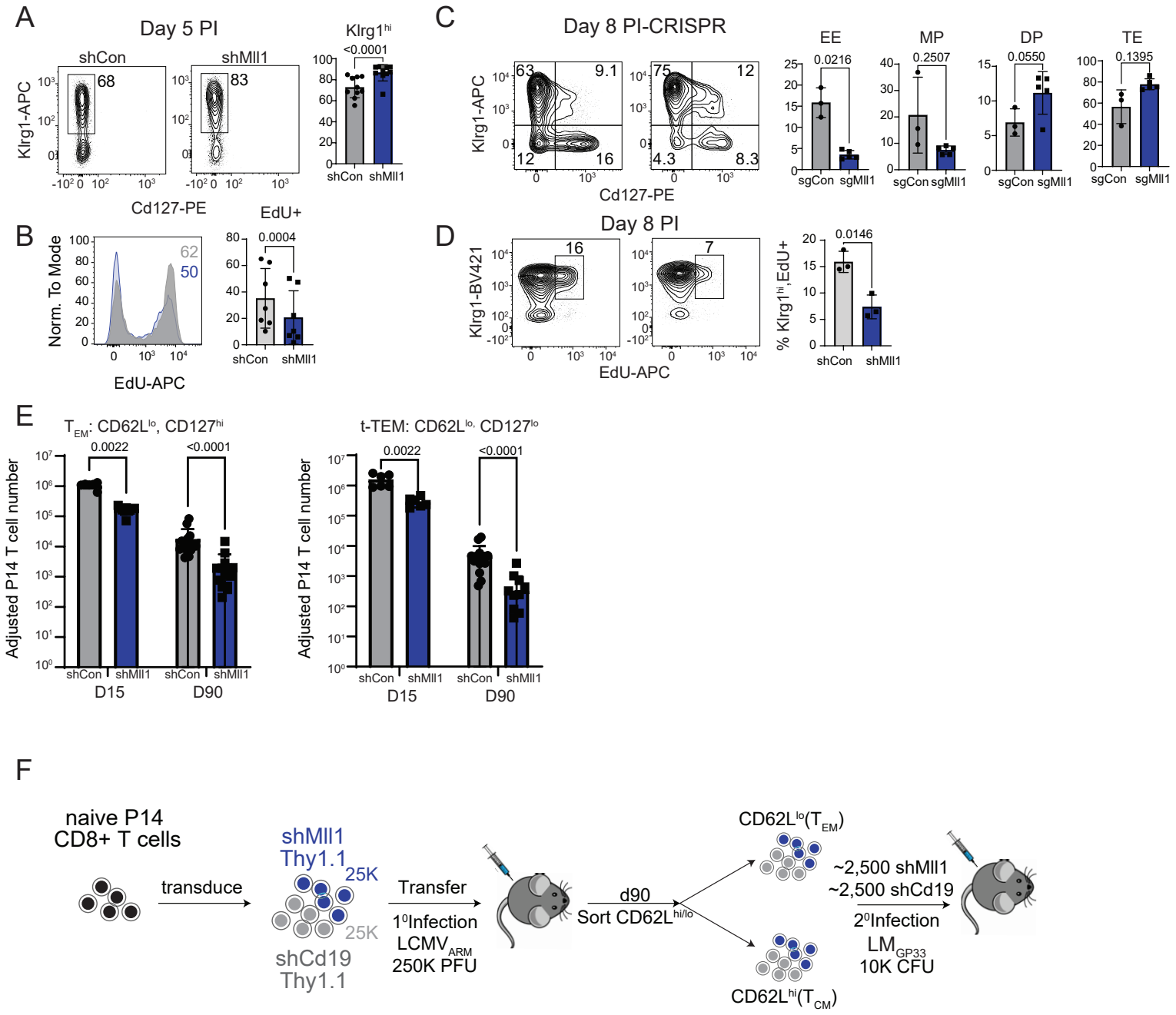

D
In vitro Tcf7 transgene expression
IL-2lo
IL-2hi
D3
D4
D5
D6
Tcf7-GFP
Tcf7-GFP-gMean
shCon- IL-2hi
shMll1- IL-2hi
shCon- IL-2lo
shMll1- IL-2lo
Days post Activation


E
In vitro CTV dilution
IL-2lo
IL-2hi
D3
D4
D5
D6
SSC-A
CTV-BV421
% full CTV dilution
shCon- IL-2hi
shMll1- IL-2hi
shCon- IL-2lo
shMll1- IL-2lo
Days post Activation

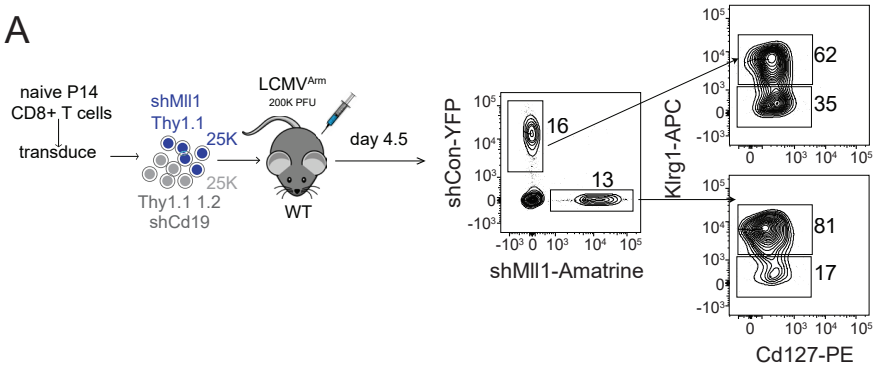
Sup. Figure 2

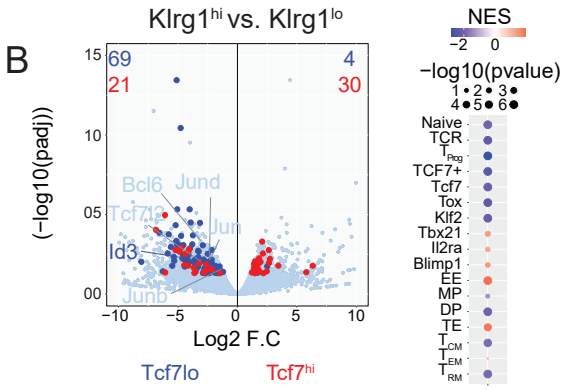

A
EO771-GP33
~ 2 weeks
1x10^6
S.C injection
1:1 v/v matrigel
naive P14
CD8+ T cells
transduce
shMll1
Thy1.1
500K
500K
shCd19
Thy1.1
B16-GP33
5x10^5
Day 8
naive P14
CD8+ T cells
transduce
shMll1
CD45.1
1x10^6
1x10^6
shCd19
CD45.1
B
E0771GP33
E0771GFP
Spleen
Thy1.1-PerCP-Cy5-5
Cd44-Alexa 700
3.8
0.8
0.1
Percent of Cd8s
Pd-1 gMean
P14 percent of all Cd8s
Pd-1 gMean
0.0327
0.0036
0.1929
<0.0001
0.0001
0.0014
GP33+
GP33-
Spleen
E0771
C
B16GP33
Spleen
Cd45.1- PerCP-Cy5.5
Cd8-BV786
86
82
Percent of Cd8s
Pd-1 gMean
P14 percent of all Cd8s
Pd-1 gMean
0.6059
0.0331
GP33+
Spleen
D
Day 3 pACT B16GP33
Pd-1-APC-Cy7
Tim3-PE
47
58
% Pd1hi, Tim3hi
0.0081
shCon
shMll1
E
Pd-1 in Tim3hi
Day 3 pACT
Normalized To Mode
933
1169
Pd-1-APC-Cy7-A
gMean of Pd-1
0.8491
shCon
shMll1
F
E0771GP33
Day 3 pACT
Day 8 pACT
Adjusted P14
T cell number
0.0010
0.7422
shCon
shMll1
B16GP33
% of Total Cd8s
0.0751
0.1275
Day3
Day10
Days post ACT
G
Control of TNBC over time
Tumor Volume (mm3)
Days post ACT
shMll1
shCon
Control
Probability of Survival
0.13
0.04
0.0012
H
Tim3lo
2787
2392
All cells
2580
2184
Normalized To Mode
Tcf7-GFP
Tcf7-GFP gMean
0.0732
0.0212
shCon
shMll1
I
Accumulation of Cd8
T cells during LCMVCL13
shCon
shMll1
Adjusted P14
T cell number
% of Cd8 T cells
Days PI
J
Day 15 LCMVCL13
shCon
shMll1
Slamf6-PE
Cd69-PE-Cy7
1.9
2.8
46
49
1.5
0.5
87
11
% of Transduced P14 cells
Slamf6hi
Cd69hi
Slamf6hi
Cd69lo
Slamf6lo
Cd69lo
Slamf6lo
Cd69hi
0.0286
0.3503
0.0009
0.0052
shCon
shMll1
K
Cx3cr1 in Tim3hi,Tcf7lo
Normalized To Mode
73
94
% Cx3cr1 positive
0.0230
shCon
shMll1

Sup. Figure 3


A
Day 5 PI
shCon
shMll1
Klrg1-APC
Cd127-PE
68
83
Klrg1hi
<0.0001
shCon
shMll1
B
Norm. To Mode
62
50
EdU-APC
EdU+
0.0004
C
Day 8 PI-CRISPR
Klrg1-APC
Cd127-PE
63
9.1
12
16
75
12
4.3
8.3
EE
MP
DP
TE
0.0216
0.2507
0.0550
0.1395
sgCon
sgMll1
D
Day 8 PI
Klrg1-BV421
EdU-APC
16
7
% Klrg1hi,EdU+
0.0146
E
TEM: CD62Llo, CD127hi
t-TEM: CD62Llo, CD127lo
Adjusted P14 T cell number
0.0022
<0.0001
D15
D90
F
naive P14
CD8+ T cells
transduce
shMll1
Thy1.1
25K
25K
shCd19
Thy1.1
Transfer
1°Infection
LCMVARM
250K PFU
d90
Sort CD62Lhi/lo
CD62Llo(TEM)
CD62Lhi(TCM)
~2,500 shMll1
~2,500 shCd19
2°Infection
LMGP33
10K CFU

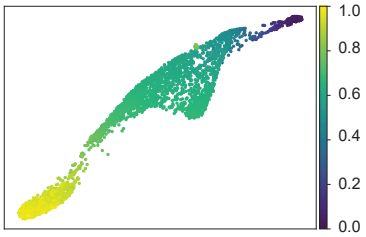
Sup. Figure 4


A
naive P14
CD8+ T cells
transduce
shMll1
Thy1.1
25K
25K
Thy1.1 1.2
shCd19
LCMV^Arm
200K PFU
WT
day 4.5
shCon-YFP
shMll1-Amatrine
16
13
Klrg1-APC
Cd127-PE
62
35
81
17


B
Klrg1^hi vs. Klrg1^lo
69
21
4
30
(−log10(padj))
Log2 F.C
Bcl6
Jund
Tcf7
Jun
Id3
Junb
Tcf7lo
Tcf7^hi
NES
−2
0
−log10(pvalue)
Naive
TCR
TCF7+
Tcf7
Tox
Klf2
Tbx21
Il2ra
Blimp1
EE
MP
DP
TE

Sup. Figure 5

A



B

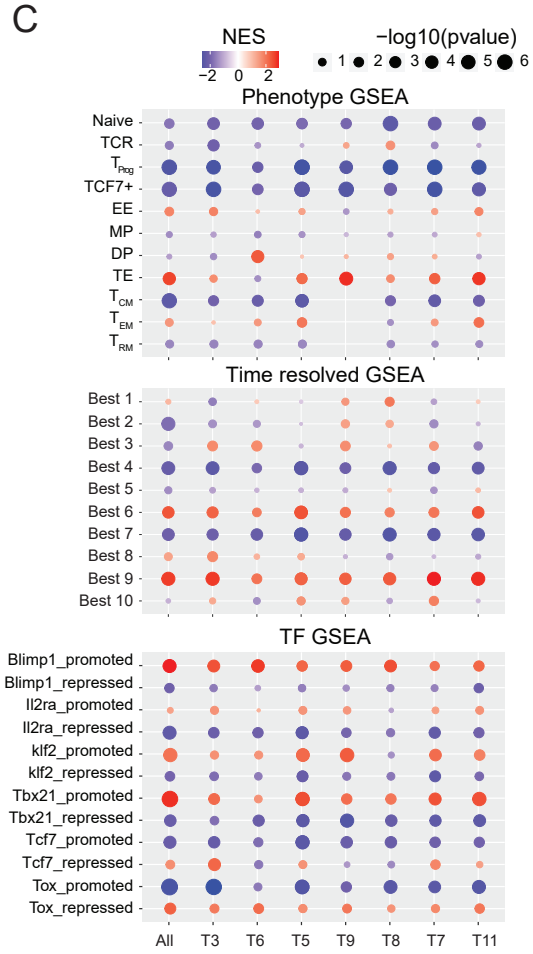
shCon Transition probabilities

Transition from	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
T1	0.000	0.324	0.744	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.000
T2	0.000	0.000	0.000	0.000	0.727	0.376	0.000	0.000	0.000	0.000	0.000	0.000	0.000
T3	0.000	0.122	0.000	0.000	0.657	0.670	0.000	0.000	0.000	0.000	0.000	0.000	0.000
T4	0.000	0.136	0.000	0.000	0.387	0.191	0.000	0.000	0.358	0.000	0.000	0.068	0.000
T5	0.000	0.000	0.000	0.000	0.000	0.000	0.284	0.543	0.223	0.000	0.000	0.000	0.000
T6	0.000	0.000	0.000	0.000	0.084	0.000	0.000	0.000	0.334	0.000	0.000	0.000	0.000
T7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.370	0.000	0.000
T8	0.000	0.000	0.000	0.000	0.000	0.000	0.156	0.000	0.000	0.000	0.349	0.000	0.000
T9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.338	0.000	0.000	0.000	0.000	0.000
T10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.733	0.000
T11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.803	0.000
T12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.225
T13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

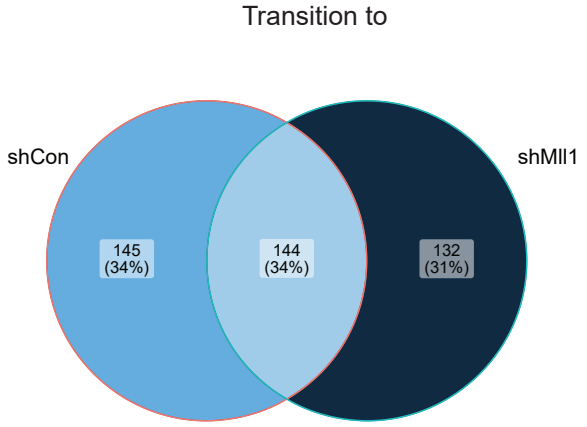
shMll1 Transition probabilities

Transition from	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
T1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.000
T2	0.000	0.000	0.016	0.000	0.053	0.105	0.012	0.024	0.147	0.000	0.044	0.000	0.000
T3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000	0.000
T4	0.000	0.030	0.000	0.000	0.153	0.164	0.030	0.029	0.280	0.000	0.062	0.066	0.000
T5	0.000	0.000	0.000	0.000	0.000	0.389	0.000	0.000	0.362	0.000	0.000	0.000	0.000
T6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.169	0.000	0.000	0.000	0.000
T7	0.000	0.000	0.000	0.000	0.070	0.118	0.000	0.023	0.430	0.000	0.455	0.000	0.000
T8	0.000	0.000	0.000	0.000	0.026	0.000	0.000	0.000	0.047	0.000	0.288	0.000	0.000
T9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000
T10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.078	0.785	0.000
T11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.680	0.000
T12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.272
T13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

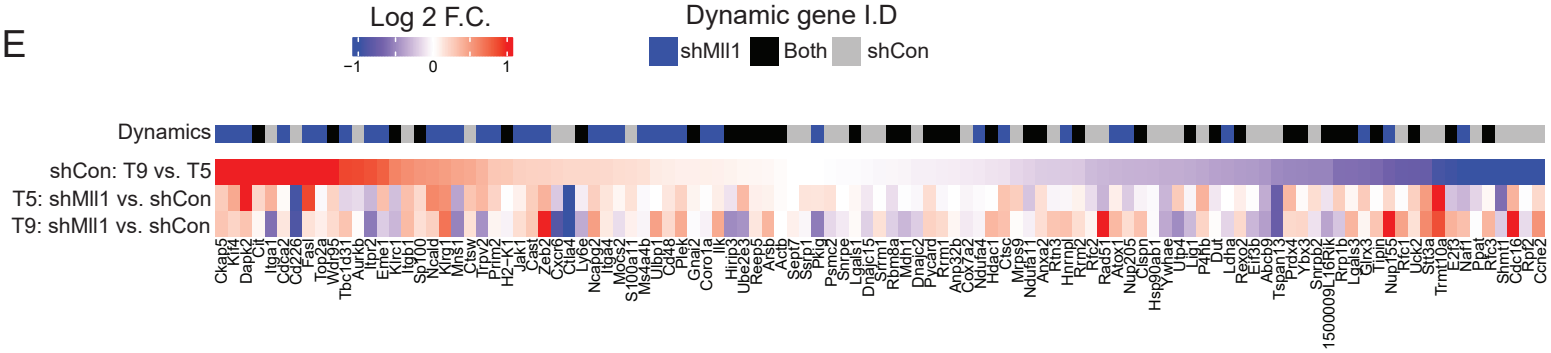
C



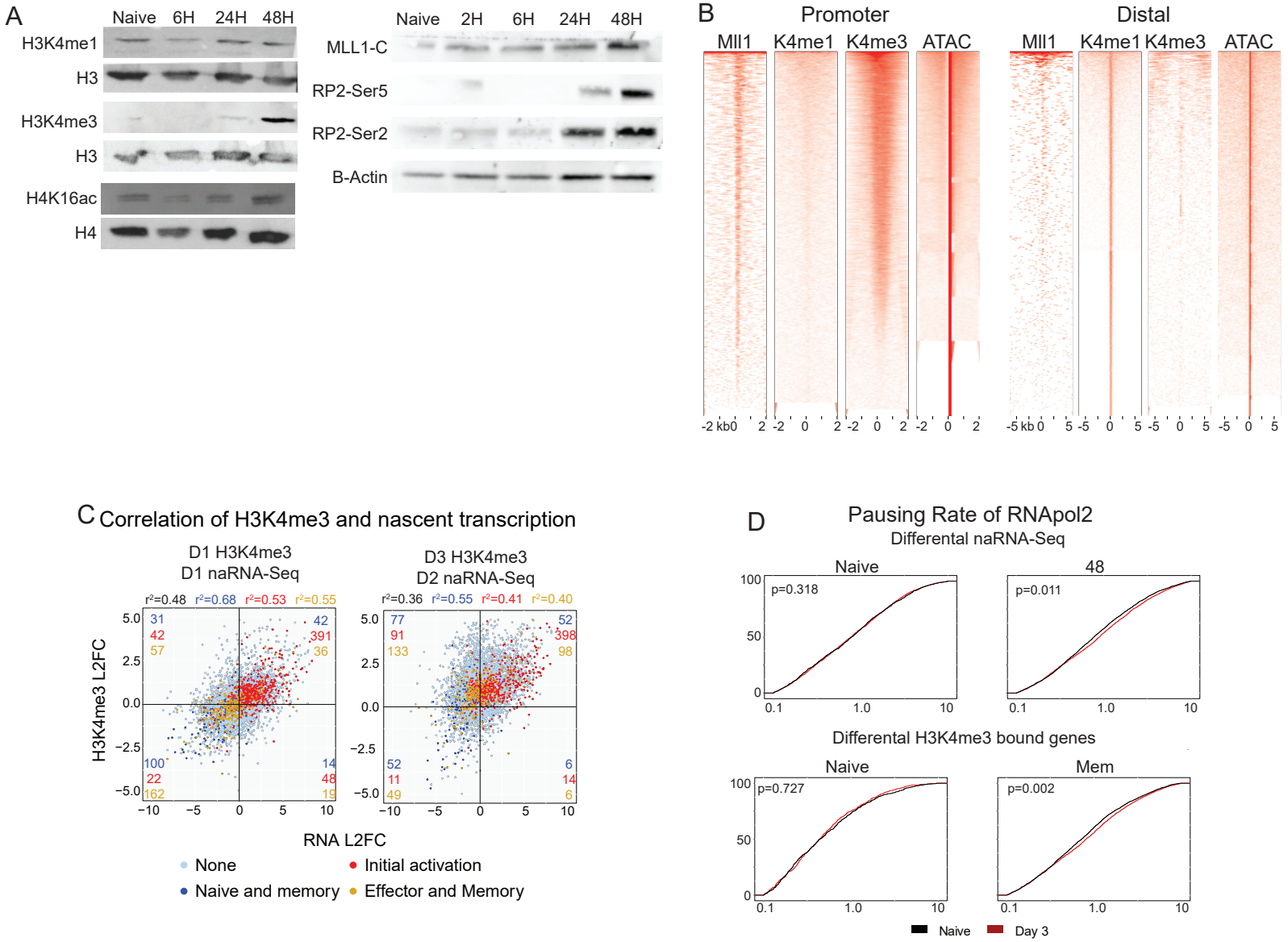
D



E



Sup. Figure 6


A
Naive 6H 24H 48H
H3K4me1
H3
H3K4me3
H3
H4K16ac
H4
Naive 2H 6H 24H 48H
MLL1-C
RP2-Ser5
RP2-Ser2
B-Actin
B
Promoter
Distal
Mll1 K4me1 K4me3 ATAC
Mll1 K4me1 K4me3 ATAC
-2 kb 0 2
-5 kb 0 5
C Correlation of H3K4me3 and nascent transcription
D1 H3K4me3
D1 naRNA-Seq
r²=0.48 r²=0.68 r²=0.53 r²=0.55
D3 H3K4me3
D2 naRNA-Seq
r²=0.36 r²=0.55 r²=0.41 r²=0.40
H3K4me3 L2FC
RNA L2FC
None
Initial activation
Naive and memory
Effector and Memory
D
Pausing Rate of RNApol2
Differental naRNA-Seq
Naive
48
p=0.318
p=0.011
Differental H3K4me3 bound genes
Naive
Mem
p=0.727
p=0.002
Naive
Day 3

Sup. Figure 7