scientific reports

Check for updates

OPEN

Deciphering the role of histone modifications in memory and exhausted CD8 T cells

Hua Huang[1,2,3,4,20], Amy E. Baxter[1,2,5,20], Zhen Zhang[3,4,6,20], Charly R. Good[3,4,20], Katherine A. Alexander[3,4,7,20], Zeyu Chen[1,2,8,9,10], Paula A. Agudelo Garcia[3,4,11], Parisa Samareh[3,4], Sierra M. Collins[3,4], Karl M. Glastad[3,4,12], Lu Wang[3,4,13,14], Gregory Donahue[3,4], Sasikanth Manne[1,2], Josephine R. Giles[1,2,19], Junwei Shi[3,15,16], Shelley L. Berger[3,4,17,18]✉ & E. John Wherry[1,2,19]✉

Exhausted CD8 T cells (T_{EX}) arising during chronic infections and cancer have reduced functional capacity and limited fate flexibility that prevents optimal disease control and response to immunotherapies. Compared to memory (T_{MEM}) cells, T_{EX} have a unique open chromatin landscape underlying a distinct gene expression program. How T_{EX} transcriptional and epigenetic landscapes are regulated through histone post-translational modifications (hPTMs) remains unclear. Here, we profiled key activating (H3K27ac and H3K4me3) and repressive (H3K27me3 and H3K9me3) histone modifications in naive CD8 T cells (T_N), T_{MEM} and T_{EX}. We identified H3K27ac-associated super-enhancers that distinguish T_N, T_{MEM} and T_{EX}, along with key transcription factor networks predicted to regulate these different transcriptional landscapes. Promoters of some key genes were poised in T_N, but activated in T_{MEM} or T_{EX} whereas other genes poised in T_N were repressed in T_{MEM} or T_{EX}, indicating that both repression and activation of poised genes may enforce these distinct cell states. Moreover, narrow peaks of repressive H3K9me3 were associated with increased gene expression in T_{EX}, suggesting an atypical role for this modification. These data indicate that beyond chromatin accessibility, hPTMs differentially regulate specific gene expression programs of T_{EX} compared to T_{MEM} through both activating and repressive pathways.

CD8 T cells are a proliferative and functional differentiation hierarchy. Following activation, quiescent naive CD8 T cells (T_N) undergo a complex and extensive rewiring of epigenetic, transcriptional regulation and gene expression programs. In the days after activation, T_N differentiate into two major divergent populations: short-lived cytotoxic effector CD8 T cells (T_{EFF}) and the precursors for long-lived, quiescent memory CD8 T cells

[1]Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. [2]Institute for Immunology and Immune Health, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. [3]Epigenetics Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. [4]Department of Cell and Developmental Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. [5]Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. [6]Institute of Health and Medicine, Hefei Comprehensive National Science Center, Hefei 230601, Anhui, China. [7]Cold Spring Harbor Laboratories, Cold Spring Harbor, NY 11724, USA. [8]Gene Regulation Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [9]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA. [10]Department of Cell Biology and Pathology, Harvard Medical School, Boston, MA 02115, USA. [11]Department of Biomedical Engineering, The Ohio State University, Columbus, OH 43210, USA. [12]Department of Biology, University of Rochester, Rochester, NY 14620, USA. [13]Sam and Ann Barshop Institute for Longevity and Aging Studies, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA. [14]Department of Biochemistry and Structural Biology, University of Texas Health Sciences Center at San Antonio, San Antonio, TX 78229, USA. [15]Department of Cancer Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [16]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, PA, USA. [17]Department of Genetics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. [18]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA. [19]Parker Institute for Cancer Immunotherapy, University of Pennsylvania, Philadelphia, PA, USA. [20]Hua Huang, Amy E. Baxter, Zhen Zhang, Charly R. Good, Katherine A. Alexander have contributed equally. ✉email: bergers@pennmedicine.upenn.edu; wherry@pennmedicine.upenn.edu

(T_{MEM})[1,2]. Whereas T_{EFF} mediate antiviral and anticancer effector functions, migrate throughout the body and help control initial disease, T_{MEM} precursors are more restrained. Following antigen clearance, most T_{EFF} die, but T_{MEM} precursors survive and differentiate into mature T_{MEM} that are quiescent and slowly self-renew. Furthermore, T_{MEM} can rapidly reactivate effector functions and proliferate following encounter with the same antigen, mounting robust recall responses. However, if antigen persists, such as during chronic viral infections and cancer, these early precursors instead differentiate into exhausted CD8 T cells (T_{EX}). In contrast to T_{MEM}, T_{EX} are maintained by persistent stimulation, resulting in chronic activation and altered effector capacity. Thus, T_{EX} mount weak recall responses to antigen restimulation and are associated with poor disease control compared to T_{MEM}[3]. Despite both cell populations originating from a common precursor cell (T_N), T_{MEM} and T_{EX} represent divergent differentiation paths, resulting in highly distinct cell types. This ability of a common progenitor population to give rise to differentiation hierarchies consisting of diverse cell types with distinct functions has parallels throughout developmental biology, for example in the gastrointestinal tract where an intestinal stem cell niche gives rise to all mucosal cell types[4]. Although much work has focused on the functional differences between T_{MEM} and T_{EX}, the precise epigenetic and transcriptional mechanisms associated with differentiation of these divergent cell states remains to be fully defined.

Epigenetic profiling of chromatin accessibility by Assay for Transposase-Accessible Chromatin with sequencing (ATAC-seq) has revealed that T_{MEM} and T_{EX} are as epigenetically distinct from each other as each cell type is from T_N[5–10]. Furthermore, T_{EX} have a unique chromatin accessibility profile with distinct open chromatin sites compared to other CD8 T cell subsets[5,6]. These observations support the idea that T_{MEM} and T_{EX} are separate lineages of mature CD8 T cells. The unique chromatin accessibility landscape of T_{EX} is established in part by the thymocyte selection associated high mobility group transcription factor (TF) TOX[11–16]. TOX is essential for the formation of T_{EX} but is dispensable for the generation of T_{MEM} despite transient expression of TOX following acute stimulation[13,17]. In addition, TOX may regulate further differentiation once the T_{EX} population is established[18], coordinating transitions between progenitor, intermediate and terminal T_{EX} subsets[19–25,18]. Thus, TOX plays a key role in establishing the unique T_{EX} open chromatin landscape profiled by ATAC-seq. However, the associations between this T_{EX} open chromatin landscape and regulation of gene expression through histone post-translational modifications (hPTMs), including how these hPTMs change as T_N differentiate into T_{MEM} or T_{EX}, remains incompletely understood.

CD8 T cell differentiation is regulated by the addition or removal of hPTMs by a variety of epigenetic enzymes. In all multicellular organisms, histone 3 lysine 27 acetylation (H3K27ac) is associated with active enhancers, whereas methylation at the same site (H3K27me3) is associated with decreased gene expression and the formation of facultative heterochromatin. In CD8 T cells, EZH2, a histone methyltransferase that establishes H3K27me3, and KDM6B, a lysine demethylase that removes methyl groups from H3K27, have both been reported to regulate cell fate specification between T_{EFF} and T_{MEM} populations, as well as formation of T_{MEM} capable of mounting robust recall responses[26–28]. Furthermore, the histone deacetylase HDAC3, which leads to chromatin compaction in part by removing acetyl groups from H3K27ac, restrains T_{EFF} development and dampens effector function[29]. In addition to H3K27me3, H3K9me3 is also associated with heterochromatin and regulates constitutive repression of repeated DNA elements and long-term repression of inactive regions. The methyltransferase Suv39h1 that establishes H3K9 tri-methylation silences T_{MEM}-associated genes to enable T_{EFF} differentiation, and may regulate T_{EX} effector function[30,31]. Furthermore, epigenetic enzymes such as protein arginine methyltransferase PRMT4 (CARM1)[32], chromatin remodelers SWI/SNF family members BAF and PBAF[33–37], and ASXL1[38], as well as enzymes catalyzing DNA methylation and demethylation (TET2[39,40] and DNMT3A[41–43] have all been implicated in regulating fate transitions between and/or within CD8 T cell subsets. The diverse functions of the epigenetic enzymes that potentially regulate CD8 T cell differentiation suggests that complex patterns of hPTMs may be a feature of the distinct T_N, T_{MEM} and T_{EX} epigenetic landscapes. Interrogating the differences in these hPTM patterns may provide insights into the diverse transcriptional regulation and gene expression programs of these CD8 T cell states.

Once established, T_{EX} are fate inflexible and do not differentiate into T_{EFF} or T_{MEM} cells. Furthermore, T_{EX} retain the epigenetic "scars" of exhaustion even after removal of antigen and "cure" of chronic infection[44]. In contrast, T_{MEM} are fate-flexible and are poised to rapidly respond when re-encountering antigen by differentiating into highly functional T_{EFF}. This T_{EX} fate inflexibility limits disease control. Targeted immunotherapies such as PD-1 pathway blockade "reinvigorate" T_{EX} and have revolutionized cancer therapy[45–47]. However, not all patients experience clinical benefit, in part because the burst of effector activity in reinvigorated T_{EX} is transient. This inability to provoke durable changes in T_{EX} function and differentiation is due, at least in part, to the failure of T_{EX}-targeted immunotherapies to rewire the T_{EX} chromatin landscape and epigenetically reprogram these cells into T_{EFF} or T_{MEM} cells[5,6,18]. Indeed, because PD-1 pathway blockade does not change the T_{EX} open chromatin landscape[5,10], reinvigorated T_{EX} revert to their original exhausted state over time[5]. Thus, although the open chromatin landscape of T_{EX} has been defined by ATAC-seq, a more comprehensive understanding of how changes in chromatin accessibility and hPTMs are associated with gene expression is required to develop immunotherapy strategies that provoke durable responses in T_{EX}. Furthermore, how epigenetic regulation through combinations of hPTMs might regulate the maintenance of T_{EX} epigenetic scars, but enable fate flexibility in T_{MEM}, remains unclear.

To address these questions, we interrogated hPTMs patterns in T_N, T_{MEM} and T_{EX} to investigate how active and repressive hPTMs were associated with the diverse gene expression and chromatin accessibility landscapes of these distinct CD8 T cell differentiation states. The epigenetic transition from a quiescent T_N state into antigen-experienced populations (i.e. from T_N to T_{MEM} or T_{EX}) was associated with broad genome-wide alterations, highlighting the extensive epigenetic remodeling associated with CD8 T cell activation. Moreover, these genome-wide changes were associated with distinct chromatin features for T_{MEM} and T_{EX} cells, highlighting a key role for hPTM patterns in the development of fate-flexible T_{MEM} versus fate-inflexible T_{EX}. Defining these hPTM

patterns may provide insight into the control of gene expression in T cell populations with diverse functions. These data provide a foundation for future epigenetic-based therapeutic approaches.

RESULTS

Histone modifications are associated with distinct gene expression landscapes of T_{MEM} and T_{EX}

To study how hPTMs could regulate T_{EX} development, we profiled activating and repressive histone modifications in T_{EX} compared to T_{MEM} and T_N. We used different strains of lymphocytic choriomeningitis virus (LCMV) to induce either an acutely resolving infection [Armstrong (Arm)] with development of T_{EFF} followed by T_{MEM}, or to establish a chronic infection [clone13 (Cl13)] that results in T_{EX} formation[3]. We adoptively transferred a physiological number of naive T cell receptor (TCR) transgenic LCMV $D^bGP_{33–41}$-specific CD8 T cells (P14 cells) into congenically distinct recipient mice and infected these recipient mice with LCMV Arm or Cl13 (Fig. 1a;[48,49]. At ~ day 30 post-infection, we isolated P14 cells from mice infected with either Arm (T_{MEM}, Fig. S1a-S1c) or Cl13 (T_{EX}, Fig. S1d-S1f). P14 cells from an uninfected mouse were isolated as naive controls (T_N). We then performed Cleavage Under Targets and Release Using Nuclease (CUT&RUN) for histone modifications H3K27ac, H3K4me3, H3K27me3, and H3K9me3 on T_N, T_{MEM}, and T_{EX} cells, alongside RNA-sequencing (Fig. 1a).

We first investigated whether hPTM patterns were distinct between CD8 T cells from acute and chronic infection compared to naive control cells. Principal component analysis (PCA) revealed that T_N, T_{MEM} and T_{EX} occupied separate regions of PCA space for each hPTM studied (Fig. 1b). Furthermore, for all hPTMs, T_N clustered separately from T_{EX} and T_{MEM} in PC1 whereas T_{MEM} and T_{EX} separated in PC2 (Fig. 1b). Thus, the highest numbers of differential hPTMs were identified between T_N and either T_{MEM} or T_{EX} (Fig. S1g), suggesting that the greatest magnitude of hPTM changes occurred as T_N were activated and differentiated into T_{EX} or T_{MEM} cell fates.

To examine how hPTMs were associated with changes in gene expression between the three CD8 T cell subtypes, we first mapped each hPTM region to the nearest gene. We selected the peak for each gene that was the most variable across conditions, mapping one peak per gene. We then correlated the fold change in RNA expression of these genes to the fold change in hPTMs at the gene-associated peak. Genes with higher H3K27ac were associated with increased gene expression ($R = 0.67$, Fig. 1c, top and Fig. S1h-S1i). In contrast, higher H3K27me3 was only weakly correlated with lower gene expression (R = -0.16, Fig. 1c, bottom and Fig. S1h-S1i). However, key CD8 T cell genes had concurrent changes in H3K27ac and H3K27me3. For example, the TF *Tcf7* was highly expressed in T_N, moderately expressed in T_{MEM} but lower in T_{EX} (Fig. S1j). These differences in RNA expression were associated with concordant changes in activating hPTMs: H3K27ac was highest in T_N, moderate in T_{MEM} and minimal in T_{EX}. In contrast, the repressive modification H3K27me3 was low in both T_N and T_{MEM}, but higher in T_{EX} (Fig. S1j). Thus, reduced *Tcf7* expression in T_{MEM} compared to T_N was associated with a decrease in activating hPTMs, whereas in T_{EX} this gene had a combination of both lower H3K27ac and higher H3K27me3. This combination of changes was associated with the lowest *Tcf7* RNA expression between CD8 T cell subsets.

The exhaustion-associated TF *Tox* is highly expressed in T_{EX} compared to T_{MEM} and the *Tox* locus has extensive open intronic chromatin in T_{EX}[13] (Fig. 1d). This region of the *Tox* gene was extensively marked with both activating H3K27ac and H3K4me3 in T_{EX}, concurrent with low H3K27me3 (Fig. 1d, Fig. S1k). In contrast, reduced expression of IL-2 receptor alpha (*Il2ra*) in T_{EX} compared to T_{MEM} was associated with higher H3K27me3 and lower H3K27ac and H3K4me3 (Fig. 1e, Fig. S1k). These analyses suggest that combinatorial changes of activating and repressive hPTMs accompany changes in gene expression between CD8 T cell fates.

To probe how the interplay between different hPTMs might regulate gene expression between CD8 T cell states, we next focused on hPTMs at promoters. H3K27ac at promoters is associated with active transcription, whereas H3K27me3 is present at inactive or poised promoters[50]. K-mean clustering identified 7 groups of genes differentially expressed between T_N, T_{EX} and T_{MEM} (Fig. 1f). One set of differentially expressed genes (DEGs) was highly expressed in T_{EX} compared to both T_N and T_{MEM} (Fig. 1f: C1; 1126 genes), whereas a second set of DEGs was highly expressed in T_{MEM} compared to both T_N and T_{EX} (Fig. 1f: C5; 528 genes). For both clusters of DEGs, H3K27ac levels were highest and broadest around the transcription start site (TSS) in the corresponding CD8 T cell type (Fig. 1g and h). For example, for genes upregulated in T_{EX} (C1), H3K27ac peaks at the TSS were highest and broadest in T_{EX} (Fig. 1g), whereas C5 DEGs (upregulated in T_{MEM}) had higher H3K27ac in T_{MEM} than T_{EX} and T_N (Fig. 1h). In contrast, the association between H3K27me3 and gene expression varied between cell state-associated genes. H3K27me3 was lower at the TSS in both T_{EX} and T_{MEM} compared to T_N for C1 genes, despite elevated gene expression for this cluster only in T_{EX} (Fig. 1i). However, H3K27me3 was lowest at the TSS of C5 genes (upregulated in T_{MEM}) in T_{MEM} compared to T_N and T_{EX} (Fig. 1j), suggesting that loss of this repressive mark may have a distinct role in enforcing gene expression in T_{MEM}. Furthermore, H3K27me3 was higher at the TSS of C5 genes in T_{EX} compared to both T_{MEM} and T_N (Fig. 1j) provoking the hypothesis that a subset of genes expressed in T_{MEM} are actively repressed in T_{EX}. Finally, consistent with the role of silencing hPTMs in gene repression, H3K27me3 deposition was distributed over ~ 10 kb around the TSS (Fig. 1j). These findings further support a role for changes in both activating and repressive hPTMs in regulating gene expression in T_{MEM} and T_{EX}, including at and outside the promoter, and also indicate that gain of activation marks, rather than loss of repressive modifications, may be a more common feature associated with increased gene expression.

We next investigated the genome-wide association of combinatorial changes in hPTMs with gene expression. DEGs between T_{EX} and T_{MEM} were binned into patterns based on higher or lower abundance of the hPTMs analyzed. For genes upregulated in T_{EX} compared to T_{MEM}, 5 of the 6 most frequent patterns were characterized by increased H3K27ac deposition in T_{EX} (Fig. 1k; G1, G2, G3, G5, and G6), which commonly co-occurred with higher H3K4me3 in T_{EX} (Fig. 1k; G1, G3, G5, G6). In contrast, associations between T_{EX} gene expression and

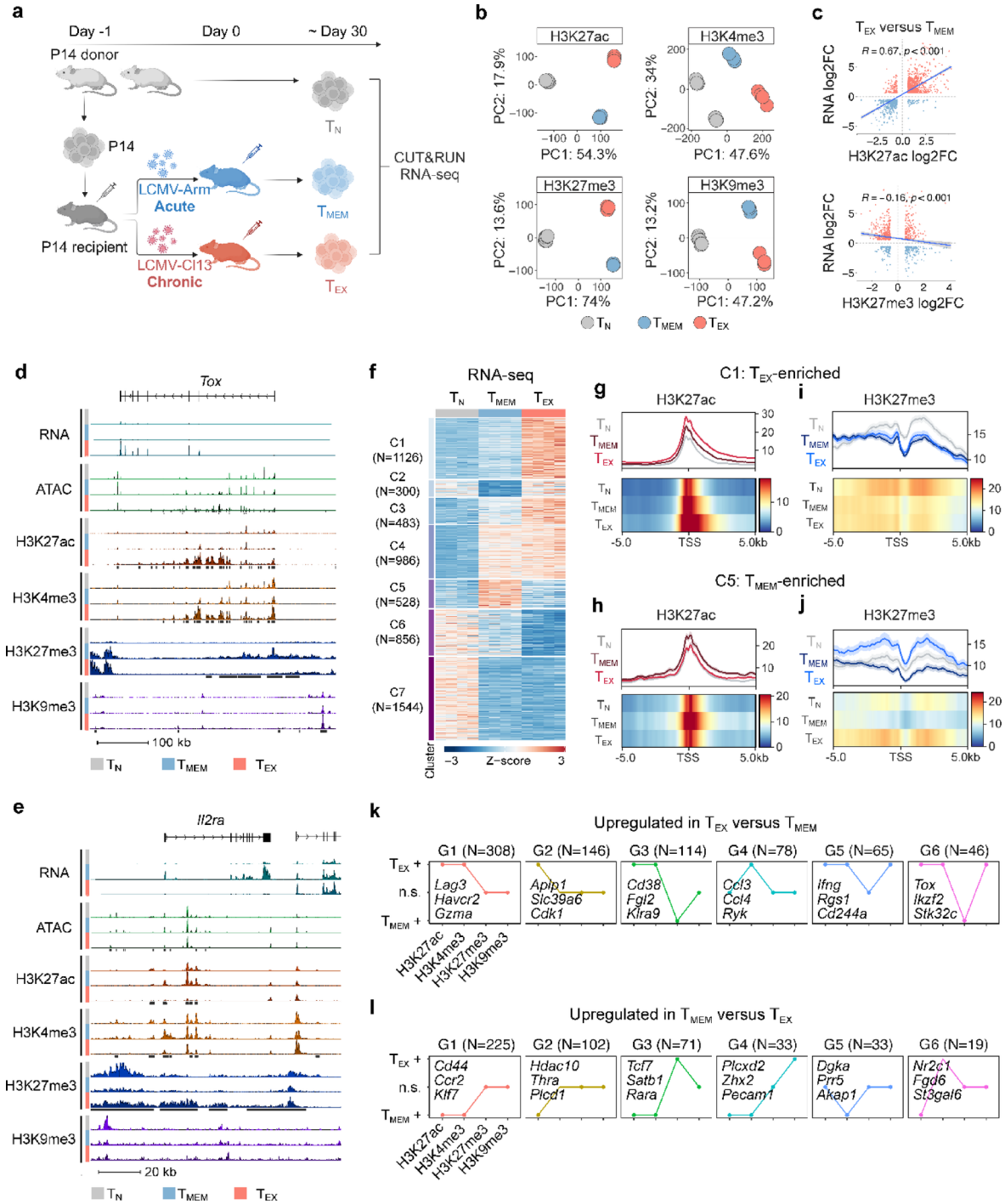


Fig. 1. Histone modifications act in concert to regulate gene expression in T_{EX} and T_{MEM}. (**a**) Experimental design. (**b**) PCA of H3K27ac, H3K4me3, H3K27me3 and H3K9me3 data for T_N, T_{MEM} and T_{EX}. For T_{MEM} and T_{EX}, $n = 4$ biological replicates; for T_N, $n = 8$ biological replicates. (**c**) Correlation of change in RNA expression between T_{MEM} and T_{EX} and change in H3K27ac (top), or H3K27me3 (bottom). R and associated P-value represent Pearson correlation. Each dot represents one gene with one associated peak. (**d-e**) Genome tracks showing RNA-seq, ATAC-seq and hPTM data. Differentially modified regions between T_{MEM} and T_{EX} for each hPTM are highlighted in black boxes. (**f**) Heatmap of DEGs showing K-mean clusters for all pairwise comparisons between T_N, T_{MEM}, and T_{EX}. (**g-h**) Meta plot (top) and heatmap plot (bottom) of H3K27ac at TSS for DEGs (**g**) cluster 1 (C1) and (**h**) cluster 5 (C5). (**i-j**) Meta plot (top) and heatmap plot (bottom) of H3K27me3 at TSS for DEGs (**i**) cluster 1 (C1) and (**j**) cluster 1 (C5). (**k-l**) Comparison of hPTM patterns between T_{MEM} and T_{EX} for (**k**) genes with increased expression in T_{EX} compared to T_{MEM} or (**l**) genes with increased expression in T_{MEM} compared to T_{EX}. Top six most frequent groups for each set of genes plotted.

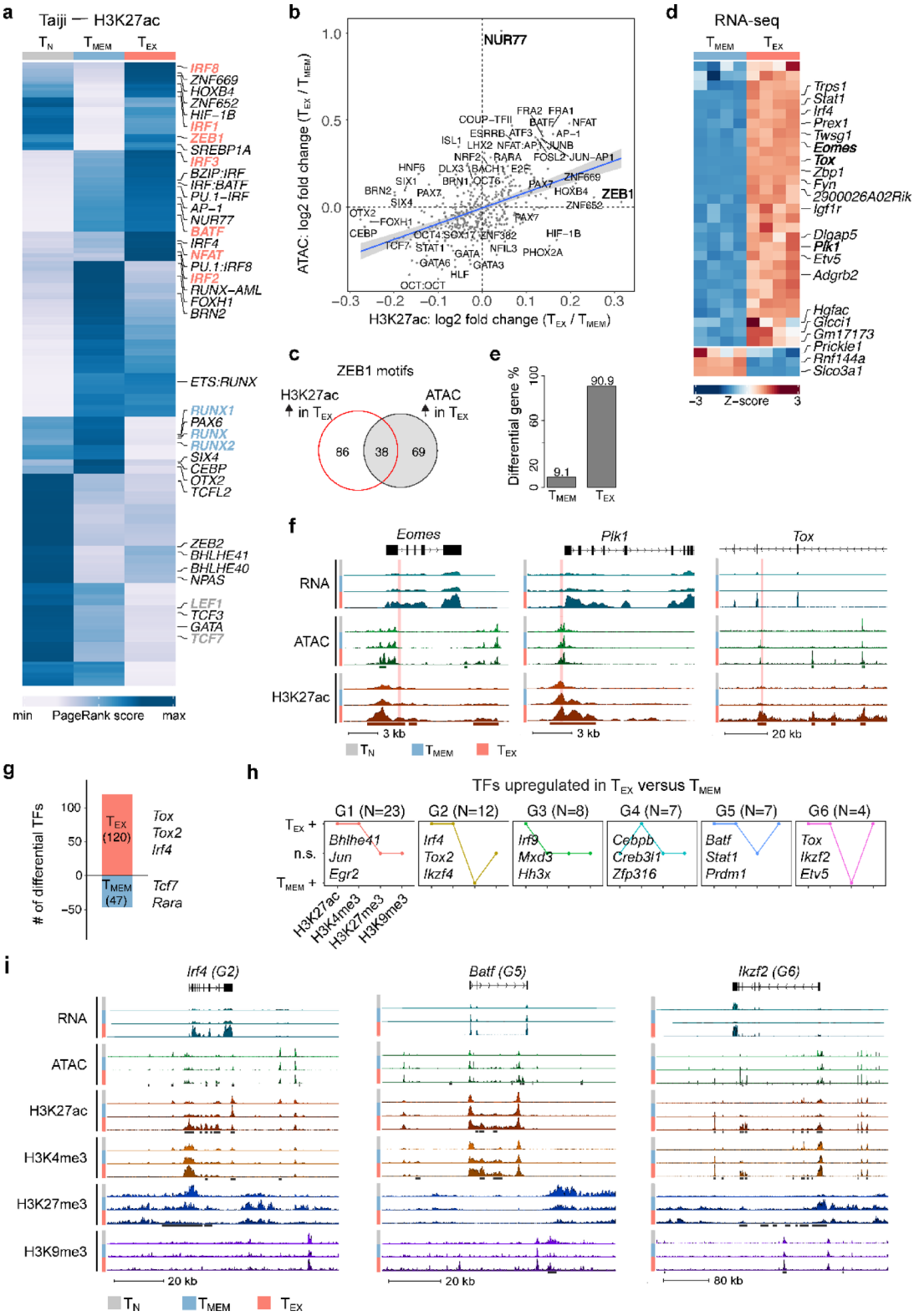
repressive marks were less consistent. H3K27me3 and H3K9me3 levels were unchanged in T_{EX} versus T_{MEM} in 3 of the 6 patterns (G1, G2 and G4), with a shift to lower H3K27me3 in T_{EX} only in patterns G3 and G6. Instead, H3K9me3 levels increased concurrently with H3K27ac and H3K4me3 in 2 patterns (G5 and G6), including for *Tox* (G6). Similar features were identified for genes upregulated in T_{MEM} (Fig. 1l). Indeed, the top 3 most frequent hPTM patterns were similar between genes upregulated in T_{MEM} or T_{EX} (Fig. 1k and l; G1, G2, G3). Together, these analyses suggest that gain of H3K27ac is the most prevalent and homogenous feature of differential gene expression between T_{EX} and T_{MEM}. In contrast, loss of the H3K27me3 is linked to increased expression of only a subset of genes whereas H3K9me3 may be associated with increased expression of some genes in T_{EX}.

H3K27ac identifies putative enhancers and predicts families of transcription factors acting at these enhancers in T_{EX}

Analysis of TF binding sites in differentially accessible chromatin regions has identified TFs driving distinct T_{MEM} and T_{EX} cell fates[10,51]. However, chromatin accessibility defined by ATAC-seq alone may not be sufficient to identify functional enhancers. Therefore, analysis of TF activity at total accessible chromatin sites may not reflect the TF regulatory networks active at enhancers controlling gene regulation. As H3K27ac amplitude is strongly correlated with enhancer activity[52,53], we hypothesized that identifying TF activity at H3K27ac sites could predict TFs active at putative enhancers. Thus, we performed Taiji PageRank analysis[54] on H3K27ac data to identify TF networks at putative enhancers that might drive T_{EX} differentiation. Taiji PageRank analysis combines peak intensity, TF motif binding site accessibility, and TF expression to predict TF activity. TFs associated with quiescence such as LEF1 and TCF7 ranked highly in T_N and higher in T_{MEM} than in T_{EX} (Fig. 2a)[55,56], whereas TFs reported to coordinate memory versus effector responses, including RUNX1 and RUNX2, ranked highest in T_{MEM}[57] (Fig. 2a). BATF, NFAT and IRF family members (IRF1, IRF3, IRF4 and IRF8) ranked highly in T_{EX} (Fig. 2a), supporting identified roles for these TFs in exhaustion[58–62]. To identify TFs with predominant roles at putative enhancers (i.e. H3K27ac sites) rather than all open chromatin regions, we compared Taiji PageRank scores generated using H3K27ac (Fig. 2a) to scores calculated using published ATAC-seq data (Fig. S2a)[5]. The majority of TFs that ranked highly in T_{EX} compared to T_{MEM} using H3K27ac also ranked highly in ATAC-seq-based analysis, including NFAT, AP-1 and BATF (Fig. 2b). Correspondingly, TFs that ranked more highly in T_{MEM} compared to T_{EX} in the H3K27ac analysis also ranked highly using ATAC-seq (e.g. TCF7) (Fig. 2b). However, some TFs were differentially ranked between the two analyses. Whereas NUR77 (NR4A1) was highly ranked in T_{EX} compared to T_{MEM} based on chromatin accessibility[63,64], this TF had comparable rankings in T_{EX} and T_{MEM} when assessed using enhancer-biased H3K27ac sites (Fig. 2b). In contrast, ZEB1 scored highly using H3K27ac in T_{EX} compared to T_{MEM}, but had similar PageRank scores in T_{MEM} and T_{EX} by ATAC-seq (Fig. 2b). The ZEB family of TFs, ZEB1 and ZEB2, may have reciprocal and potentially opposing roles in mature CD8 T cell differentiation. Of note, the ZEB2 binding motif remains undefined, preventing Taiji PageRank analysis. Whereas ZEB1 is required for T_{MEM} survival and function[10,65], ZEB2 promotes terminal differentiation and effector function[66,67]. In T_{EX}, ZEB1 regulates T_{EX} survival and persistence, whereas ZEB2 mediates T_{EX} cytotoxic function, suggesting these TFs modulate distinct, potentially opposing, gene regulatory networks[10,68]. A deeper understanding of TFs acting at putative enhancers may uncover relevant T_{EX} transcriptional networks, including transcriptional pathways regulated by the TF pair ZEB1/ZEB2.

To further investigate the roles of ZEB1 and NUR77 at H3K27ac sites compared to accessible chromatin regions, we used motif analysis to identify predicted ZEB1 and NUR77 binding sites within differentially H3K27ac-modified regions and differentially accessible chromatin regions. Although ~20% (38/193) of ZEB1 motifs were found in regions where both H3K27ac and chromatin accessibility were increased in T_{EX} compared to T_{MEM}, ~45% (86/193) of predicted ZEB1 binding sites had increased H3K27ac without a concurrent increase in chromatin accessibility (Fig. 2c). Of these 86 ZEB1 predicted binding sites with increased H3K27ac deposition in T_{EX}, but without increased chromatin accessibility (Fig. 2c), the vast majority, ~91%, were associated with genes increasing in expression in T_{EX} (Fig. 2d and e). Genes potentially regulated by ZEB1 binding at sites with changing H3K27ac deposition included the TFs *Eomes* and *Tox*, and the mitotic regulator *Plk1* (Fig. 2d and f). In contrast, NUR77 motifs were much more prevalent in regions of increased chromatin accessibility in T_{EX} than regions with increased H3K27ac, representing 71% of predicted binding sites (Fig. S2b). We identified an enrichment in these predicted binding sites toward genes with increased expression in T_{EX} (Fig. S2c-S2d), including the genes encoding TFs *Setbp1 and Ikzf2* (HELIOS), the pro-survival factor *Bcl2* and *Tnfsf4* (OX40L) (Fig. S2c and S2e). However, ~30% of these NUR77 motifs were located in genes highly expressed in T_{MEM}. These analyses support the differential rank in Taiji PageRank analysis for ZEB1 and NUR77 and suggest a preferential role for ZEB1 in T_{EX} through binding sites marked by H3K27ac.

Assessment of TF binding motifs at H3K27ac sites revealed TFs potentially acting at putative T_{EX} enhancers. However, both the accessibility of TF binding sites and expression of these TFs themselves must be tightly regulated to orchestrate broad changes in transcriptional networks during CD8 T cell differentiation. Therefore, we next asked how the TF genes were themselves regulated by hPTMs. First, we identified TFs from the AnimalTFDB database[69] with differential gene expression between T_{MEM} and T_{EX} cells. The majority of TFs identified in this pairwise comparison increased in expression in T_{EX}, with 120 TF genes upregulated in T_{EX} versus T_{MEM} compared to only 47 TFs with increased expression in T_{MEM} (Fig. 2g). We then assessed the genes encoding these TFs for associated changes in hPTMs (Fig. 2h). For TFs that increased in expression in T_{EX} compared to T_{MEM}, H3K27ac, H3K4me3 or both increased in all 6 of the hPTM patterns; however, only 2 patterns (G2 and G6) had decreased H3K27me3 and none had decreased H3K9me3 (Fig. 2h). For example, *Irf4*, *Batf* and *Ikzf2* all gained H3K27ac and H3K4me3 in T_{EX}, whereas H3K27me3 was unchanged for *Batf*, but lost for *Irf4* and *Ikzf2* (Fig. 2i). Similar patterns were observed for TFs with higher expression in T_{MEM}, where all of the top 6 patterns were associated with increased H3K27ac levels in T_{MEM} (Fig. S2f). Together, these analyses indicate that upregulation of TF expression is more frequently associated with an increase in



the activating hPTMs rather than removal of repressive hPTMs. This pattern of regulation (gain of activation-associated modifications in T_{EX}) was not unique to TFs, but was observed for all genes differentially expressed between T_{EX} and T_{MEM} (Fig. 1k). Thus, these analyses indicate that both expression of the TFs that coordinate CD8 T cell differentiation and the genes downstream of these TFs are regulated by similar patterns of hPTMs. Furthermore, these data suggest that regulating gene expression through the acquisition of activating hPTMs may be a common feature of CD8 T cell differentiation.

◀**Fig. 2**. Identification of predicted TF binding motifs under H3K27ac identities role for ZEB1 in T_{EX}. (**a**) Heatmap of normalized Taiji PageRank scores determined using RNA-seq and H3K27ac data. (**b**) Correlation plot comparing Taiji PageRank scores from H3K27ac to ATAC-seq data. Axes represent log2 fold change in Taiji PageRank scores between T_{EX} and T_{MEM}. (**c**) Venn diagram comparing number of ZEB1 motifs in regions with increased H3K27ac in T_{EX} to regions with increased chromatin accessibility (ATAC) in T_{EX}. (**d**) Heatmap showing DEGs between T_{MEM} and T_{EX} associated with regions with increased H3K27ac in T_{EX} that contain ZEB1 motifs. (**e**) Bar graph showing cell type expression of DEGs associated with ZEB1 motifs in regions with H3K27ac enriched in T_{EX} without concurrent increases in chromatin accessibility. (**f**) Genome tracks highlighting ZEB1 motifs in regions with increased H3K27ac levels in T_{EX} without changing chromatin accessibility. Differentially modified regions for H3K27ac and open chromatin are highlighted in boxes under tracks. Predicted ZEB1 binding sites are shown in red. (**g**) Number of differentially expressed TFs between T_{MEM} and T_{EX}, with representative TFs indicated. (**h**) Comparison of hPTMs between T_{EX} and T_{MEM} for TFs with increased expression in T_{EX}. Top six most frequent groups plotted. (**i**) Genome tracks showing RNA-seq, ATAC-seq and hPTM data for TFs with increased expression in T_{EX}. Differentially modified regions for each modification are highlighted in black bars.

Super enhancers associate with distinct transcriptional wiring of T_{EX}

Super enhancers (SEs) are large clusters of enhancers with the potential to bind numerous TFs and recruit co-factors[52,70], playing key roles in defining cell fate, controlling cell identity and/or driving disease[52]. We identified enhancers in each cell type based on the presence of H3K27ac at non-promoter regions (Fig. S3a), and ranked "stitched" enhancers using the ROSE algorithm[70,71]. Enhancers with high signal intensity were defined as super enhancers, whereas those with low signal intensity were defined as typical enhancers (TEs) (Fig. 3a and S3b). As expected, SEs were associated with higher expression of nearby genes compared to TEs (Fig. S3c). To investigate whether SEs were associated with the distinct T_{EX} and T_{MEM} cell fates, we ranked SEs within each CD8 T cell population (Fig. 3a). A number of top-ranked SEs were shared between all three CD8 T cell populations, such as *Ikzf1* (IKAROS), *Rapgef1*, *Fyn* and the transcriptional regulator *Id2*, whereas the TFs *Tbx21* (TBET), *Zeb2* and *Runx2* were shared between T_{MEM} and T_{EX} (Fig. 3a). In contrast, the SEs near genes involved in quiescence, such as *Bach2* and *Foxp1*, ranked highly in T_N (Fig. 3a, left). Although several SEs were highly ranked in both T_{MEM} and T_{EX}, key differences were identified. SEs more highly ranked in T_{MEM} than T_{EX} were located near genes associated with effector biology, such as *Rora*, *Klrb1b and Klrg1*, persistence-associated TFs *Bhlhe40*[68,72,73] and *Tcf7* (TCF1), and the cytokine receptor *Il2ra* were more highly ranked in T_{MEM} than T_{EX} (Fig. 3a, middle). In contrast, SEs close to the T_{EX}-associated TFs *Tox*, *Eomes* and *Batf*, as well as the inhibitory receptors *Pdcd1* and *Havcr2* ranked highly in T_{EX} (Fig. 3a, right). Furthermore, top-ranked SEs in T_{EX} were associated with unannotated RNAs or lncRNAs, including *2310001H17Rik* (Fig. 3a, right). Thus, H3K27ac-associated SEs likely play a role in regulating expression of key lineage-defining TFs, effector molecules and other genes driving the distinct T_N, T_{MEM} and T_{EX} cell differentiation paths.

Chromatin accessibility can infer SEs[52,70] and in T_{EX} such analysis has been used to investigate the regulation of *Tox* expression[13]. Therefore, we investigated whether defining SEs by H3K27ac rather than chromatin accessibility could provide additional insights into SE regulation of the T_{EX} cell fate. We directly compared SEs defined by chromatin accessibility[5] (Fig. S3d) to SEs identified via H3K27ac (Fig. 3a and b). Many SEs were highly ranked using both approaches, including SEs associated with *Tox* and *Pdcd1* (Fig. 3b). However, the majority of SEs identified by chromatin accessibility were not identified using H3K27ac (Fig. S3e). For example, *Tigit*, encoding an inhibitory receptor[74], *Nr4a2*, encoding a TF reported to promote T cell exhaustion[13,15,64], and TNF family member *Tnfsf10*, all ranked highly for SE activity defined by chromatin accessibility but were not identified as H3K27ac-defined SEs (Fig. 3b and S3f). In contrast, *Havcr2*, encoding the inhibitory receptor TIM3, and killer cell lectin-like receptor *Klra8* ranked highly when SEs were identified by H3K27ac, but not by chromatin accessibility (Fig. 3b and c). Furthermore, Gene Ontology (GO) analysis revealed functional divisions within SEs. Whereas GO terms for gene-associated SEs identified by chromatin accessibility were more likely to have roles in cell survival and differentiation, genes associated with SEs identified based on H3K27ac were enriched for GO terms involved in cell adhesion, division and inflammatory responses (Fig. S3g). Therefore, H3K27ac identified additional potential SE-regulated genes both with known roles in T cell exhaustion and genes that have not previously been deeply interrogated in T_{EX}.

As antigen-experienced cells, T_{EX} and T_{MEM} share a core epigenetic and transcriptional network that distinguishes these cells from T_N. However, T_{EX} and T_{MEM} also have distinct chromatin accessibility and transcriptional circuits that are cell-type specific and define these two differentiation trajectories[5,6,8,10]. More than half of all SEs we identified (52%) were shared between T_N, T_{MEM} and T_{EX}, suggesting a common role in CD8 T cell biology (Fig. 3d). Furthermore, 161 SEs were shared between T_{MEM} and T_{EX}, reflecting common pathways in non-naive T cells. Moreover, only 6% of SEs were unique to T_{EX} (58/985), and even fewer were unique to T_{MEM} (3%; 33/985; Fig. 3d). Therefore, we next examined whether T cell fate-specific SEs were associated with cell type specific expression of the SE-associated gene. Indeed, T_{MEM}-specific SEs were associated with high gene expression only in T_{MEM} and included *Il2ra* and *Cd44* (Fig. 3e). In contrast, genes with SEs specific to T_{EX} were highly expressed only in T_{EX}, including the SE-associated genes *Tox* and the inhibitory receptors *Entpd1*, *Pdcd1* and *Havcr2* (Fig. 3e and f). This analysis also revealed T_{EX}-enriched SE-associated genes that have not been extensively studied in T_{EX}, including *Setbp1*[13], *Trps1* and *Ubash3b* (Fig. 3e and f). Together, these data add further support to the hypothesis that SEs play key roles driving both the shared and distinct chromatin regulatory and transcriptional circuitry of T_{EX} versus T_{MEM} cells and identify previously understudied H3K27ac-enriched SEs associated with genes in T_{EX} that warrant further investigation.

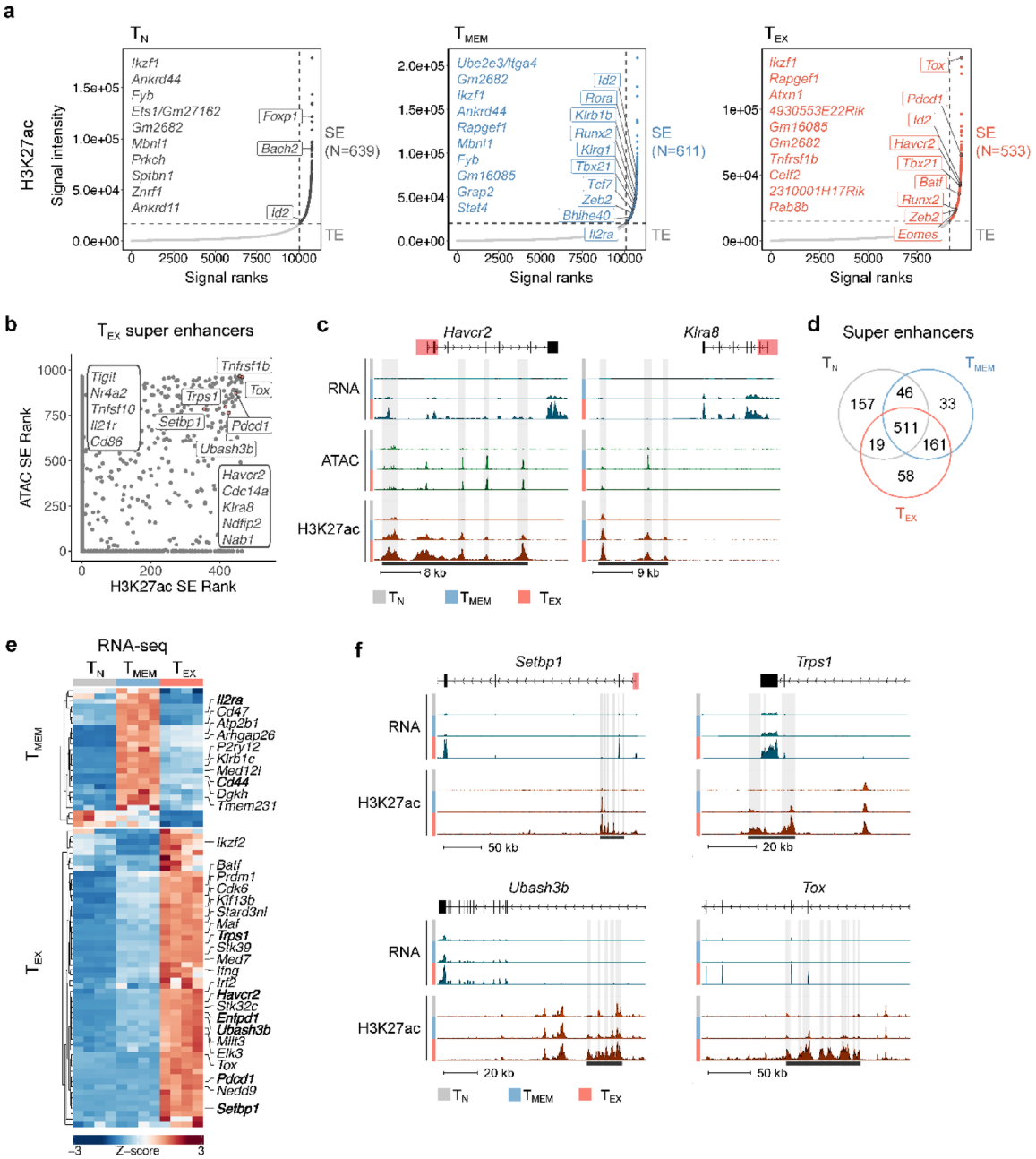


Fig. 3. Super enhancers drive transcriptional phenotype of T_{EX}. (**a**) Distribution of H3K27ac signal across stitched enhancer regions in T_N, T_{MEM} and T_{EX}. Top 10 ranked putative SEs plus selected SEs are highlighted. N indicates the number of putative SEs identified for each cell type. Stitched enhancers above horizontal dashed lines are associated with putative SEs; enhancers below horizontal dashed lines are typical enhancers (TE). (**b**) Comparison of SE ranks identified using H3K27ac signal (x-axis) and ATAC signal (y-axis). Selected SEs are labeled. (**c**) Genome tracks showing RNA-seq, ATAC-seq and H3K27ac data. Putative SE regions identified using H3K27ac data but not ATAC-seq data are highlighted in black boxes below tracks. Individual enhancers within SE regions are highlighted in grey. Promoter regions (± 2,500 bp of TSS) are indicated in red. (**d**) Venn diagram showing cell-type specificity of SEs identified by H3K27ac. (**e**) Heatmap showing gene expression of genes within 50 kb of SE identified only in T_{MEM} and T_{EX}, respectively. Selected DEGs are highlighted. (**f**) Genome tracks showing RNA-seq and H3K27ac data. Putative SE regions identified using H3K27ac data are highlighted in black box. Individual enhancers within SE regions as highlighted in grey. Promoter regions (± 2,500 bp of TSS) are indicated in red.

Chromatin state analysis reveals state-specific transitions from T_N to T_{MEM} or T_{EX}

To study how genome-wide chromatin states change during CD8 T cell differentiation, we applied ChromHMM, an algorithm that uses multiple hPTMs to segment the genome into distinct states[75]. We used ATAC-seq, H3K27ac, H3K4me3, H3K27me3, and H3K9me3 data to identify the four major promoter states; active (I),

poised (II), repressive (III) and repetitive/heterochromatin (IV) promoters. Active promoters (I) were defined by open chromatin and by the active marks H3K27ac and H3K4me3 (Fig. S4a-S4b). As expected, the vast majority of (~95%) of genes with active promoters in one CD8 T cell population were highly expressed in that cell type (Fig. S4c). Pathway analysis of genes with active promoters in T_N revealed that these promoters were associated with core cellular function pathways, including DNA repair, chromatin segregation and translation (Fig. S4d), suggesting that, in T_N, genes involved in basic cellular functions are regulated by promoters with active marks. Repressed (III) and repetitive/heterochromatin (IV) promoters were defined by deposition of the repressive marks H3K27me3 and/or H3K9me3 (Fig. S4a-S4b). Accordingly, the vast majority of genes with repressed or repetitive/heterochromatin promoters were not expressed in the cell type with those repressed or repetitive/heterochromatin features (Fig. S4c). In T_N, genes with repressed promoters were enriched for more specialized pathways with limited roles in CD8 T cells, such as muscle contraction, sensory perception of pain and response to pheromone (Fig. S4d).

The poised promoter state was first described in embryonic stem cells (ESCs) as nucleosomes bearing both H3K4me3 and H3K27me3[76–78]. The genes associated with these dual modified promoters were not expressed in ESCs, but instead were turned on as cells acquired identity and lineage commitment in the developing embryo[79,80]. This poised state has also been described in "multipotent" T_N[81]. To investigate how genes with poised promoters are associated with T_{MEM} and T_{EX} cell fates, we first identified genes with poised (state II) promoters in T_N (Fig. S4a-S4b). As expected, in T_N the majority (~57.7%) of genes with promoters bearing both H3K4me3 and H3K27me3 were not highly expressed (Fig. S4c), supporting the assignments of these promoters as poised. Within the 4,945 poised T_N promoters, we identified promoters that shifted to an activated state only in T_{EX} or only in T_{MEM} (Fig. 4a, indicated by star and triangle respectively). We then focused on promoters that were associated with increased gene expression in each cell type (Fig. 4b). For example, the Src family kinase *Yes1* and the IL-2 receptor alpha (*Il2ra*) showed this pattern of activation only in T_{MEM} (Fig. 4b and c). Furthermore, promoters for genes posited to have roles in CD8 T cell persistence and tissue residency, including *Prss12*[82], *Nt5e* (CD73[83]), and *Ier3* (IEX-1[84]) transitioned from poised-to-active state only in T_{MEM} (Fig. 4b). GO analysis

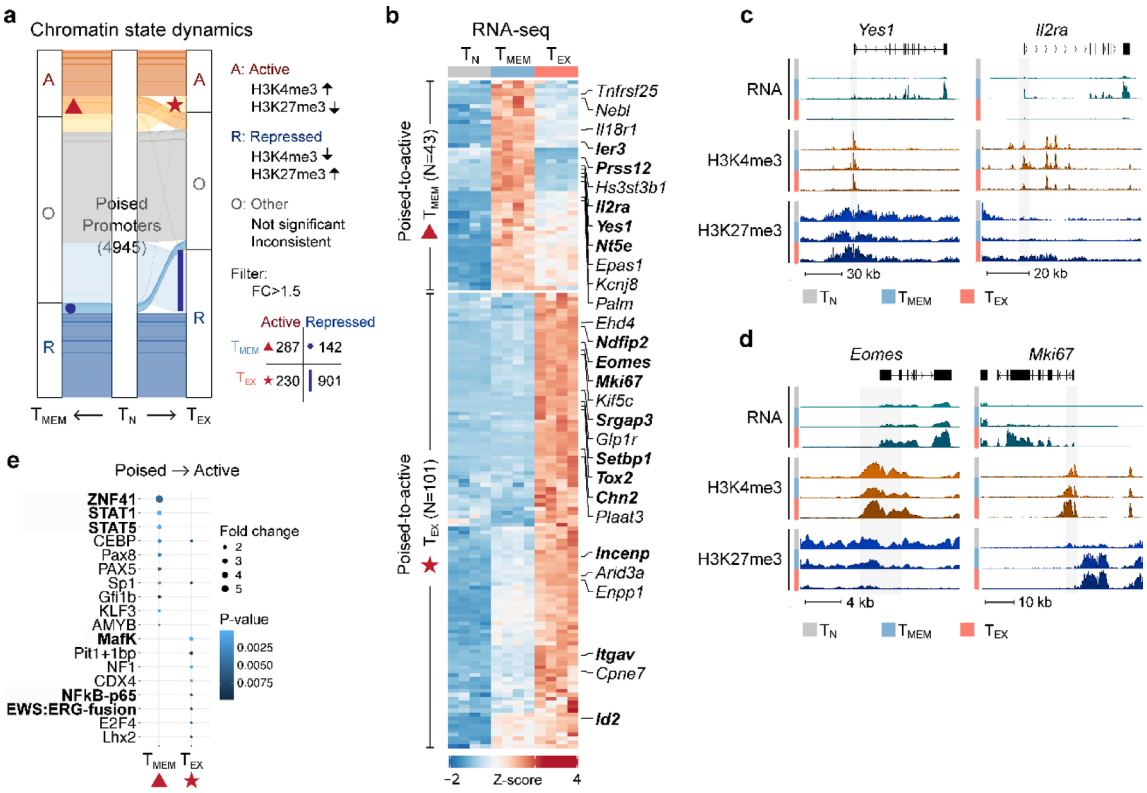


Fig. 4. Key T_{MEM} and T_{EX} genes are poised in T_N and activated upon differentiation. (**a**) Alluvial plot showing how hPTMs at promoters poised in T_N change in T_{MEM} and T_{EX}. Active promoters were defined as either significantly gaining H3K4me3 or losing H3K27me3 or both; repressed promoters were defined as either significantly losing H3K4me3 or gaining H3K27me3 or both. Statistical cutoff of FC > 1.5. Number represents total number of T_N poised protomers; inset table shows number of poised-to-active and poised-to-repressed promoters in T_{MEM} and T_{EX}. (**b**) RNA expression heatmap for DEGs with poised promoters which were activated in either T_{MEM} (triangle) or T_{EX} (star). Selected DEGs are highlighted in bold. (**c**-**d**) Genome tracks showing poised promoters in T_N that switched to active (**c**) only in T_{MEM} or (**d**) only in T_{EX}. Promoters were excluded from analysis and are highlighted in grey bars. (**e**) Bubble plot showing changes in predicted TF binding site accessibility for poised-to-active T_{MEM} or T_{EX} genes, with select TFs highlighted.

revealed functions in metabolism and signaling, including cytokine signaling (Fig. S4e). Together these analyses suggest that genes poised in T_N that become activated in T_{MEM} are involved in key aspects of T_{MEM} biology.

In contrast, genes with promoters that shifted from T_N poised to active in T_{EX} included the TFs *Eomes*, which drives terminal differentiation of T_{EX}[19,61] and *Tox2*, which has a reported role in T_{EX}[15], as well as other T_{EX}-associated genes including *ItgaV* (CD51)[61], *Srgap3*[85], and *Ndfip2*[85] (Fig. 4b and d). Several cell cycle-associated genes also showed this pattern of regulation, including *Mki67* (KI67), *Incenp* and *Chn2* (Fig. 4b and d) consistent with the more extensive cell division history of T_{EX} including ongoing cell cycle[10,19]. Of note, the promoter of *Id2* was poised in T_N, and shifted to an activated state only in T_{EX}, despite RNA expression in both T_{MEM} and T_{EX} (Fig. 4b[86–88]. Further analysis revealed that, although H3K27me3 was lost at the *Id2* promoter in both T_{MEM} and T_{EX}, T_{EX} retained activating H3K4me3 and this modification decreased in T_{MEM} compared to T_{EX} (Fig. S4f). These analyses highlight the complexity of gene expression regulation by a suite of hPTMs at and around the promoter. Finally, several genes that have not previously been well studied in T_{EX} were also identified, including *Setbp1* (Fig. 4b), which was also associated with a SE in T_{EX} (Fig. 3e and f). GO analysis identified pathways linked to development, differentiation, and proliferation (Fig. S4e). Together, these data suggest that genes with poised promoters in T_N are key genes involved in the divergent differentiation trajectories of T_{MEM} and T_{EX}, including lineage-driving TFs and genes associated with T_{EX} function (e.g. continued proliferation and survival during high antigen stress).

We next identified TFs associated with expression of genes with poised-to-active promoters that might therefore drive the distinct T_{EX} and T_{MEM} cell fates. We performed TF motif analysis on these poised-to-active promoters (Fig. 4a and b). Motifs for ZNF41, a zinc finger family TF, were enriched in T_{MEM} poised-to-active promoters, as were both STAT1 and STAT5 motifs (Fig. 4e). Indeed, STAT1 is required for CD8 T cell clonal expansion and memory formation[89], and STAT5 has a key role in early effector and memory-precursor CD8 T cell differentiation[90]. In contrast, predicted MafK binding sites were enriched in poised-to-active T_{EX} promoters (Fig. 4e). MafK forms heterodimers with BACH2 to help direct the repressive activity of BACH2 to specific genes[91]. BACH2 is a key transcriptional coordinator involved in T cell quiescence in T_N, T_{MEM} including stem cell memory cells, but also in stem cell-like progenitor T_{EX}[92,93] suggesting a potential T_{EX}-associated BACH2-MafK regulatory module. Furthermore, FLI1, an ETS family TF, dampens effector CD8 T cell transcriptional networks[94]. The EWS: ERG fusion motif, which is also predicted to be bound by FLI1, was enriched in poised-to-active promoters in T_{EX} (Fig. 4e), supporting a potential role for FLI1 or other ETS family members in restraining effector biology in T_{EX}. Finally, NFkB-p65 motifs were enriched in T_{EX} poised-to-active promoters (Fig. 4e). NFkB signaling has broad roles in T cells, regulating initial TCR-mediated T cell activation, proliferation and effector function as well as T_{MEM} survival[95,96]. Moreover, NFkB transcriptional circuitry is augmented after treatment with immunotherapies targeting inhibitory receptors, such as PD-1 blockade, or costimulation, such as CD137 (41BB) agonism[5,97]. The enrichment for NFkB-p65 motifs at T_{EX} poised-to-active promoters suggests that reengaged NFkB circuitry in immunotherapy-reinvigorated T_{EX} may be driving expression of previously poised genes, with potential implications for improving therapies. Together, these data suggest that a subset of the distinct T_{EX} and T_{MEM} transcriptional programs consist of genes with poised promoters in T_N that then acquire either active or repressed chromatin states during differentiation.

We next investigated how chromatin modifications could reinforce distinct T_{MEM} and T_{EX} cell states through repression of alternative fates. We identified promoters that were poised in T_N and then became repressed in T_{MEM} or T_{EX} either through gain of H3K27me3, loss of H3K4me3 or both potentially to silence genes of alternative cell fates (Fig. 4a). In total, 901 promoters switched from a poised to repressed state in T_{EX} compared to only 142 promoters for T_{MEM} (Fig. 4a, blue line for T_{EX} compared to blue circle for T_{MEM}). Genes with promoters that switched from poised-to-repressed as T_N differentiated into T_{MEM} on average lost the activating mark H3K4me3 and also maintained the repressive mark H3K27me3 (Fig. S4g). For example, several genes associated with T cell differentiation including the TFs *Tox2*, *Eomes* and *Ikzf2* (HELIOS), the exhaustion-associated ectonuclease *Entpd1* (CD39), costimulatory molecule *Tnfsf4* (OX40L) and glycoprotein *Itm2a* all had higher H3K27me3 and lower H3K4me3 at the promoters for these genes in T_{MEM} compared to T_N and T_{EX} (Fig. S4h). This poised-to-repressed promoter state was associated with lower RNA expression of these genes in T_{MEM} than in T_{EX}. (Fig. S4i), indicating that genes highly expressed in T_{EX} and poised in T_N are actively repressed in T_{MEM}. In contrast, genes that became repressed during T_{EX} differentiation gained repressive H3K27me3 at the promoter, however there was variable loss of H3K4me3 at these promoters (Fig. S4j). For example, in T_{EX}, whereas *Ier3* (IEX1), *Tnfrsf25* (DR3), *Hdac10* and *Mapk12* gained H3K27me3 at the promoter, only *Mapk12* lost H3K4me3 (Fig. S4h, S4j-S4k). Furthermore, a subset of genes with promoters that switched from poised-to-repressed in T_{EX} switched from poised-to-active in T_{MEM} (e.g. *Ier3*, *Il2ra*, *Tnfrsf25*) and vice versa (e.g. *Tox2*) (Fig. 4b and S4h). Together, these data suggest that active repression of genes poised in T_N may help enforce the distinct T_{MEM} and T_{EX} states and potentially limit conversion between these two populations, with gain of H3K27me3 predominantly driving this repression in T_{EX}.

Atypical H3K9me3 narrow peaks are enriched for CTCF motifs and occur at distinct repeat classes

H3K9me3 is typically associated with constitutive heterochromatin and is classically involved in silencing gene expression, playing a critical role in cell differentiation[98,99]. This hPTM classically exhibits broad peaks across the genome[98], however analysis of H3K9me3 peak width in CD8 T cells revealed a wide range in peak size, from < 5 kb to > 100 kb, with most peaks less than 10 kb (Fig. S5a)[100,101]. Therefore, we examined how H3K9me3 localization and deposition changed during CD8 T cell differentiation. Regions with higher H3K9me3 in T_{MEM} compared to T_{EX} (T_{MEM}-enriched peaks; $n = 1565$) were broad, covering ~ 40.1 kb bases on average (Fig. 5a, left; Fig. 5b). In contrast, regions with higher H3K9me3 in T_{EX} compared to T_{MEM} (T_{EX}-enriched peaks; $n = 1279$) were narrower, averaging only ~ 15.7 kb bases (Fig. 5a, right; Fig. 5c). Directly comparing the distribution of

peak sizes between T_{EX}-enriched and T_{MEM}-enriched H3K9me3 peaks showed that T_{EX}-enriched peaks were substantially narrower than T_{MEM}-enriched peaks (Fig. 5d). Moreover, only 10.2% ($n=130$) of T_{EX}-enriched peaks were broad (>=15 kb), compared to 42.9% ($n=672$) of T_{MEM}-enriched peaks (Fig. 5e). Analysis of the number of base pairs in the genome covered by H3K9me3 revealed that although the majority of the genome is covered by broad peaks, over 25% of the T_{EX}-enriched base pairs are in narrow peaks (Fig. S5b). The majority of narrow T_{EX}-enriched H3K9me3 peaks were found in intergenic and intronic regions, with only a small proportion (~3%) located in promoter regions (Fig. S5c). These results suggest that H3K9me3 deposition patterns at T_{MEM} or T_{EX}-enriched peaks have distinct characteristics, with H3K9me3 enriched in narrow peaks in T_{EX} and broad peaks in T_{MEM}.

A major function for H3K9me3 is to repress repetitive genomic elements, maintaining genome integrity[102]. Therefore, we examined whether the atypical narrow H3K9me3 peaks enriched in T_{EX} also were associated with repetitive elements like broad H3K9me3 peaks. We first assessed repeat coverage of narrow, broad and H3K9me3 peaks that were not significantly different (n.s.) between T_{EX} and T_{MEM}. Analysis of T_{EX}-enriched H3K9me3 peaks shows that 53% of narrow peaks covered repeats, compared to around 34% of broad peaks (Fig. S5d). This analysis suggests that broad and narrow H3K9me3 peaks may function to repress repetitive genomic elements, but that narrow H3K9me3 peaks could be additionally specialized to this role.

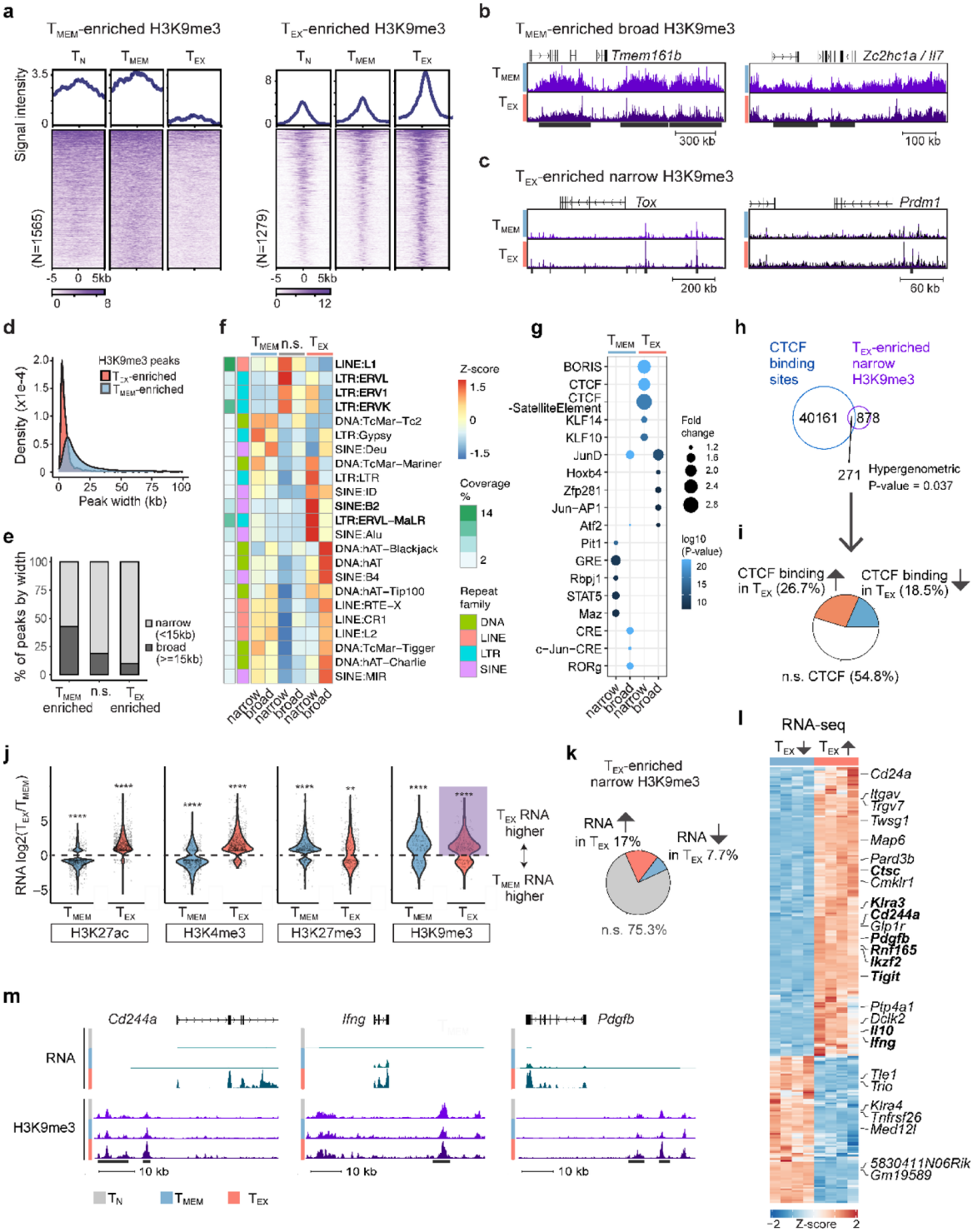
We next investigated whether specific families of repetitive elements were enriched underneath the T_{EX}-enriched narrow H3K9me3 peaks. T_{EX}-enriched narrow H3K9me3 peaks had reduced coverage of LINE elements (27.7%) (Fig. S5e, top) compared to n.s. narrow peaks (60.8%) and T_{MEM}-enriched narrow peaks (52%), as did T_{EX}-enriched broad peaks (Fig. S5e, bottom). However, T_{EX}-enriched narrow peaks showed an increase in LTR element coverage compared to other H3K9me3 peak sets (Fig. S5e). To further investigate this finding, we assessed repeat element subclasses. Narrow H3K9me3 peaks that were not enriched in T_{MEM} or T_{EX} were strongly associated with LINE: L1 and LTR: ERV subclasses compared to n.s. broad H3K9me3 peaks (Fig. 5f). T_{MEM}-enriched broad and narrow peaks had similar patterns of repeat element coverage, wheras T_{EX}-enriched broad and narrow peaks were associated with distinct repeat element subclasses (Fig. 5f). For example, T_{EX}-enriched narrow H3K9me3 peaks showed a unique enrichment in the LTR subclass ERVL-MaLR and the retrotransposon SINE B2 (Fig. 5f), suggesting a distinct regulatory role for these elements in T_{EX} cells.

Repeat elements have been co-opted to serve critical roles in chromatin organization, enhancer function, and gene regulation[103,104]. SINE B2 is one example of repeat elements serving functional roles. These elements are rodent-specific retrotransposons that contain binding sites for CTCF[105], a zinc finger protein that acts as a transcriptional activator, repressor, and has a role in genome organization[106]. H3K9me3 at SINE B2 repetitive elements can regulate CTCF binding at these sites[107]. Therefore, we performed an unbiased motif analysis of T_{MEM}- and T_{EX}-enriched H3K9me3 peaks. Indeed, the top most enriched motifs under T_{EX} narrow peaks were CTCF and the CTCF related factor (CTCFL or BORIS) (Fig. 5g). To validate whether CTCF bound directly at T_{EX}-enriched H3K9me3 narrow peaks, we performed CUT&RUN for CTCF in $T_{N,}$ $T_{MEM,}$ and T_{EX}. Nearly a quarter (24%, $n=271$) of T_{EX}-enriched H3K9me3 narrow peaks were bound by CTCF (Fig. 5h, $p=0.037$). Of these, 45% had differential CTCF binding: 26.7% showed increased CTCF binding in T_{EX} compared to T_{MEM}, while 18.5% had reduced CTCF binding in T_{EX} (Fig. 5i). For example, CTCF binding increased in H3K9me3 peaks associated with *Ctsc* (CAPTHESIN C), a peptidase that coordinates activation of serine proteases including granzymes, whereas CTCF binding decreased at a H3K9me3 peak close to *Klra3*, an NK cell receptor expressed in a cytotoxic subset of T_{EX} (Fig. S5f). Together, these findings suggest that narrow H3K9me3 peaks may have distinct associations with repetitive elements in different CD8 T cell subtypes, and that altered H3K9me3 deposition in T_{EX} may impact CTCF binding at specific sites in the genome, with potential implications for cell type-specific genome organization.

To investigate whether these T_{EX}-enriched narrow H3K9me3 peaks were associated with chromatin accessibility, we examined the overlap between these peaks and T_{EX}-accessible chromatin regions. Approximately 25% (287/1149) of T_{EX}-enriched narrow H3K9me3 peaks were located in regions of open chromatin in T_{EX} (Fig. S5g), for example near *Tox* (Fig. S5i). Genes with narrow H3K9me3 peaks associated with open chromatin, were significantly enriched in cytokine-mediated signaling and leukocyte cell-cell adhesion pathways (Fig. S5h). These observations suggests that the association of a subset of narrow H3K9me3 peaks with open chromatin may contribute to the role for this hPTM to regulate expression of key cell-type associated genes.

T_{EX}-enriched H3K9me3 peaks correlate with gene activation

H3K9me3 is typically associated with repression of gene expression[108]. However, a subset of genes with increased expression in T_{EX} compared to T_{MEM} had higher H3K9me3 deposition in T_{EX} (Fig. 1k, G5 and G6). Thus, we next investigated the association between H3K9me3 and gene expression in T_{EX}. Regions with differentially enriched hPTMs between T_{MEM} or T_{EX} were identified, then filtered for nearby genes that were differentially expressed between T_{MEM} and T_{EX} (Fig. 5j). As expected, genes near regions with increased H3K27ac or H3K4me3 in one cell type had overall higher expression in that cell type (Fig. 5j), consistent with gene-activating functions of H3K27ac and H3K4me3. In contrast, genes near regions of increased H3K27me3 or H3K9me3 in T_{MEM} had lower expression in T_{MEM} than T_{EX} (Fig. 5j), consistent with the repressive roles of H3K27me3 and H3K9me3. However, although increased H3K27me3 was negatively associated with gene expression in T_{EX}, H3K9me3 was positively associated with gene expression, with nearby genes more highly expressed in T_{EX} versus T_{MEM} (Fig. 5j, purple box). Given our observations that increased H3K9me3 deposition in T_{EX} was predominantly localized to narrow peaks (Figs. 5a and e and 89.8% of peaks), we tested if the link between H3K9me3 and increased gene expression was associated with narrow H3K9me3 peaks, or a general feature of H3K9me3 in T_{EX}. Both narrow and broad H3K9me3 peaks showed the same pattern of increased deposition of H3K9me3 near genes with higher expression in T_{EX} (Fig. S5j).



To assess the breadth of impact of these H3K9me3 peaks on regulating the transcriptional programs of T_{EX}, we next asked what proportion of T_{EX}-enriched peaks were associated with changes in gene expression (Fig. 5k). Given that the vast majority of H3K9me3 peaks in T_{EX} were narrow, we focused the subsequent analyses on these atypically narrow peaks. The majority of genes close to T_{EX}-enriched H3K9me3 narrow peaks did not change expression (n.s., Fig. 5k), suggesting that not all H3K9me3 has gene regulatory functions. However, 17% of H3K9me3 T_{EX}-enriched narrow peaks were associated with gene upregulation in T_{EX} cells, whereas 7.7% were associated with decreased expression (Fig. 5k), such that narrow H3K9me3 T_{EX}-enriched peaks were twice as likely to be associated with gene upregulation compared to downregulation. Of note, T_{MEM}-enriched narrow H3K9me3 peaks did not show this positive association with gene expression (Fig. S5k). Furthermore, 12.3% of genes upregulated in T_{EX} compared to T_{MEM} had an increase in H3K9me3 deposition in a narrow peak within 50 kb (Fig. S5l). This fraction increased to 34% of genes when assessing H3K9me3 deposition within 250 kb of the gene (Fig. S5l), suggesting that a subset of T_{EX} upregulated genes were potentially regulated at least in part by atypical "activating" H3K9me3.

◀**Fig. 5**. T_{EX}-enriched atypical H3K9me3 peaks cover CTCF sites and are associated with gene expression. (**a**) Signal intensity heatmap of T_{MEM}-enriched and T_{EX}-enriched H3K9me3 regions. (**b**) Genome track showing broad H3K9me3 regions enriched in T_{MEM}. (**c**) Genome track showing narrow H3K9me3 regions enriched in T_{EX}. (**d**) Peak size distribution of T_{MEM}- and T_{EX}-enriched H3K9me3 regions. (**e**) Percentage of H3K9me3 peaks that are broad versus narrow. Peaks > = 15 kb are defined as broad, peaks < 15 kb are defined as narrow. N.s. = not significantly different between T_{MEM} and T_{EX}. (**f**) Heatmap of Z-scored repeat element class coverage of T_{MEM}-enriched, non-significant (n.s.) and T_{EX}-enriched H3K9me3 narrow and broad peaks. (**g**) Predicted TF binding motifs in T_{MEM}- and T_{EX}-enriched narrow and broad peaks compared to n.s. H3K9me3 peaks. Top 5 motifs shown per comparison. (**h**) Venn diagram showing overlap between locations of CTCF binding sites and T_{EX}-enriched narrow H3K9me3 peaks. P value represents a hypergeometric test. (**i**) Pie chart showing change in CTCF binding at sites within T_{EX}-enriched narrow H3K9me3 peaks. (**j**) Violin plot of log2 fold change in RNA expression between T_{MEM} and T_{EX} near differentially modified regions for each hPTM (H3K27ac, H3K4me3, H3K27me3 and H3K9me3). (**k**) Pie chart showing changes in RNA expression between T_{MEM} and T_{EX} of genes near T_{EX}-enriched narrow H3K9me3 peaks. (**l**) Heatmap of RNA expression for genes from Fig. 5k. (**m**) Genome tracks showing DEGs near to narrow T_{EX}-enriched H3K9me3 regions.

Finally, we performed a GO analysis to examine differences in the biological processes associated with H3K9me3 deposition in T_{EX}. For example, genes nearby T_{EX}-enriched narrow peaks with decreased expression in T_{EX}, indicative of a repressive function for H3K9me3, were enriched for negative regulation of lymphocyte activation (Fig. S5m). In contrast, genes with increased expression close to regions with increased H3K9me3 deposition in T_{EX}, i.e. "activating" H3K9me3, were associated with cell cycle, cell death, response to cytokine stimulus and leukocyte activation (Fig. S5m). Thus, genes potentially regulated by both canonical repressive and non-canonical "activating" H3K9me3 play broad roles in T_{EX} biology. Indeed, upregulated genes in T_{EX} located near T_{EX}-enriched narrow H3K9me3 peaks included the key T cell exhaustion TFs *Tox*, *Prdm1* (BLIMP-1) and *Ikzf2*, inhibitory receptors *Cd244a* (2B4) and *Tigit*, and functional molecules *Ifng*, *Il10* and *Pdgfb* (Fig. 5c and l-m). Together, these results suggest that the H3K9me3 modification may have distinct characteristics in T_{EX}, with increased deposition localized to non-conventional narrow peaks, a subset of which are located near key genes that increase in expression in T_{EX}, including *Tox*.

Discussion

Here we profiled the hPTMs and chromatin epigenetic landscape of CD8 T cells as they differentiate from T_N into two functionally different cell fates, T_{MEM} and T_{EX}. T_N, T_{MEM} and T_{EX} cells had distinct epigenetic profiles across both activating (H3K27ac and H3K4me3) and repressive (H3K27me3 and H3K9me3) hPTMs, with the majority of hPTM changes occurring as T_N were activated and differentiated into T_{MEM} or T_{EX}. The unique transcriptional networks of T_{MEM} and T_{EX} were co-regulated by combinations of hPTMs, with gain of activating hPTMs playing a dominant role compared to loss of repressive modifications. Differentiation from T_N into T_{MEM} or T_{EX} resulted in both activation and repression of discrete subsets of genes that were poised in T_N, suggesting that hPTMs play a role in both upregulating T_{MEM} versus T_{EX} transcriptional networks, and repressing transcription of genes associated with the opposing cell fate. Whereas increased deposition of H3K9me3 in T_{MEM} was associated with decreased gene expression in T_{MEM}, a subset of genes with increased expression in T_{EX}, including the TF *Tox*, had increased H3K9me3 in nearby non-canonical narrow peaks. Thus, our analyses reveal the complexity of hPTMs in guiding alternative CD8 T cell fates, with potentially atypical roles in T_{EX}.

T_{EX} have a unique open chromatin accessibility and transcriptional landscape compared to T_{MEM}. However, precisely how this cell-fate specific chromatin accessibility may mediate cell-fate associated gene expression remains poorly understood. Analysis of individual hPTMs revealed that in both T_{MEM} and T_{EX} gain of activating modifications in one cell type was strongly associated with increased gene expression in this cell type. In contrast, loss of repressive modifications was only loosely correlated with increased gene expression in this cell type. Supporting this observation, we found that activating modifications were frequently gained in the most common combinatorial patterns of hPTMs associated with cell-type specific increases in gene expression, but that often these changes in H3K27ac and H3K4me3 did not co-occur with loss of H3K27me3 and/or H3K9me3. Thus, these analyses suggest that active modifications are key components for gene activation in CD8 T cells, whereas repressive hPTMs may serve a fine-tuning role, selectively regulating a subset of potentially cell-fate related genes.

To further investigate how combinations of activating and repressive marks regulate T_{MEM} and T_{EX} gene expression, we focused on poised chromatin states (with both H3K4me3 and H3K27me3) and the dynamics of these states as CD8 T cells differentiate. Genes in a poised state displayed high cellular plasticity, enabling them to quickly respond to antigen stimulation and facilitate rapid cell differentiation[109–111]. Thus, poised chromatin states are linked to genes driving cellular identity[110]. Since T_{MEM} and T_{EX} share a common progenitor (T_N), we investigated how genes in a poised state in T_N might shape the distinct transcriptional networks of these populations. Discrete subsets of genes poised in T_N shifted to an active state in either T_{MEM} or T_{EX}. In T_{MEM}, this subset included genes such as *Ier3* and *Il2ra*, and in T_{EX} included multiple TFs such as *Eomes*, *Setbp1*, *Tox2* and *Id2* in addition to genes such as *Ki67*. Furthermore, genes that shifted from poised-to-active promoter states only in T_{MEM} or only in T_{EX} were regulated by distinct families of TFs, with STAT family binding motifs enriched in T_{MEM} poised-to-active genes compared to MafK and NFkB-p65 in T_{EX}. Thus, the distinct functions of T_{MEM} and T_{EX} were regulated by discrete sets of TFs that coordinated upregulation of key genes primed for activation in quiescent T_N. Whether this poised state is imprinted during thymic development when many T cell genes are "tested" or reflects even earlier developmental poising will be interesting to examine in the future. Furthermore,

T_{MEM} and T_{EX} are distinct endpoints in complex differentiation trajectories originating from a common T_N precursor. It will be interesting to investigate the dynamics of how promoters poised in T_N are activated/repressed throughout the full trajectory of T_{MEM} and T_{EX} differentiation, for example as early T_{MEM} precursors differentiate into T_{MEM}. T_{EX} are epigenetically inflexible and do not convert to T_{EFF} or T_{MEM} cell states. Analysis of genes that were poised in T_N but shifted to a repressed state in T_{MEM} or T_{EX} highlighted the role of active repression in forming and maintaining these distinct CD8 T cell populations. For example, expression of the inhibitory receptor *Entpd1* was repressed in T_{MEM}, whereas *Il2ra* was repressed in T_{EX}. Together, these data suggest that, although gain of activating hPTMs plays a dominant role in the upregulation of discrete TF regulatory networks in T_{MEM} and T_{EX}, repressive hPTMs also coordinate the repression of opposing CD8 T cell fates. These findings highlight the multidimensional roles of hPTMs in T_{EX}, and emphasize the importance of both activating and repressive modifications. Understanding the combinations of these hPTMs provides a comprehensive view of the regulatory landscape, and reveals the complex patterns regulating gene expression in T_{EX}.

The top three "patterns" of hPTM changes associated with gene upregulation were the same between T_{MEM} and T_{EX}, indicating that the broad associations of these hPTMs and gene expression are comparable between CD8 T cell states. However, in T_{EX}, a subset of upregulated genes was associated with gain of both H3K27ac and H3K4me3, but also a gain of H3K9me3, typically a repressive modification. These genes included key T_{EX} TFs, such as *Tox* and *Ikzf2*, effector genes including *Ifng*, and the gene including the inhibitory receptor *Cd244a* (2B4). This finding was in contrast to the widely understood role of H3K9me3 and its association with gene repression, especially in embryonic stem cells[108,99], suggesting that this modification may have a distinct function in T_{EX}. Indeed, we found that classically broad H3K9me3 was, as expected, associated with decreased gene expression in T_{MEM}, indicating that H3K9me3 performs this typical role in T_{MEM}. However, in T_{EX}, H3K9me3 was predominantly gained in atypically narrow peaks and, furthermore, these narrow H3K9me3 peaks were associated with gene activation. Narrow H3K9me3 peaks were enriched for CTCF motifs and the repetitive elements SINE: B2 and LTR: ERVL-MaLR, suggesting that CTCF binding could be regulated by H3K9me3 deposition in T_{EX}. Specifically, increased H3K9me3 at SINE B2 sites in T_{EX} may influence CTCF binding patterns and therefore genome organization during CD8 T cell exhaustion, potentially contributing to increased gene expression at these locations. These data provoke the hypothesis that gene activation in T_{EX} is, at least in part, mediated by atypical H3K9me3 deposition that influences CTCF binding to alter higher order chromatin structure, which may in turn regulate gene expression. This result highlights an unusual role of H3K9me3 in gene regulation in T_{EX}. Understanding whether and how this role impacts chromatin organization, and what the relationship is between these hPTM patterns and CTCF binding in T_{EX} biology will be of interest for future studies.

Multiple recent studies have used ATAC-seq to examine the unique chromatin accessibility landscape of T_{EX} and to investigate how this landscape impacts T_{EX} biology[5,6,8,10,7,9,112,113]. However, chromatin accessibility is only one feature of a dynamic epigenetic landscape and analysis of hPTMs may provide additional insights into CD8 T cell subset differentiation and T_{EX} biology. Using H3K27ac deposition to identify cell-state associated SEs uncovered additional SEs compared to examination of chromatin accessibility alone. For example, we discovered a potential role for SEs in regulating expression of *Havcr2* (TIM3) and *Klra8* using H3K27ac, whereas ATAC-seq did not identify these SEs. In addition, analysis of predicted TF activity at H3K27ac sites provided further insight into TF function at enhancers and SEs. For example, the TF NUR77 was highly ranked in T_{EX} compared to T_{MEM} only when activity was predicted using open chromatin data, but not when H3K27ac-decorated regions were analyzed. This observation suggests that NUR77 may be predominantly acting in T_{EX} at genomic locations not associated with enhancer or SE activities. In contrast, the TF ZEB1 was predicted to have high importance in T_{EX} compared to T_{MEM} at H3K27ac-associated enhancers and SEs, but not across open chromatin regions in general. Thus, understanding the additional layer of regulatory networks modulated by hPTMs on top of chromatin accessibility provides further insight into how the distinct T_{MEM} and T_{EX} transcriptional programs are established. Further work is required to investigate how TFs such as ZEB1 function at SEs.

In this study, we interrogated the epigenetic landscape of two distinct CD8 T cell fates, T_{MEM} and T_{EX}, and their common precursor T_N. Despite these two populations representing endpoints of an antigen-driven differentiation hierarchy, T_{MEM} and T_{EX} themselves are heterogeneous and contain further proliferative and functional hierarchies, including subsets that function as stem cell-like reservoirs for more terminally differentiated, effector-like populations[19–25,18,10]. It is likely that observations made for bulk T_{EX} and T_{MEM} populations reflect an average hPTM landscape and average gene expression across distinct T_{MEM} and T_{EX} subsets. For example, it is likely that additional hPTM associations with transcriptional circuits will be identified in subsets of T_{EX} such as the key progenitor T_{EX} and the downstream T_{EX} intermediate and terminal populations. This subset-specific variation in hPTMs could contribute to functional differences between T_{EX} subsets, potentially shaping the diversity of T cell responses. Similar heterogeneity exists in T_{MEM} as well. Thus, future research should explore how hPTMs mediate control of gene expression within these additional T_{MEM} and T_{EX} subsets, including at earlier timepoints in the differentiation trajectories of these subsets. Our analysis of hPTMs uncovered a potential interplay between hPTM deposition and higher-order chromatin structure in regulating gene expression in T_{EX}. How the deposition of these hPTMs, including atypical H3K9me3 in T_{EX}, is orchestrated remains unknown, including the role of T_{MEM} and T_{EX} TF networks in recruiting epigenetic enzymes to sites of both histone modification and chromatin remodeling. It will be interesting in the future to investigate how TF networks, hPTMs, and three-dimensional genome structure coordinate establishment and maintenance of the distinct transcriptional landscape of T_{MEM} and T_{EX} and their subsets. Thus, understanding how the T_{MEM} and T_{EX} differentiation hierarchies are epigenetically regulated will provide key insight into the epigenetic scar of exhaustion, fate-flexibility, and could be used to inform effective clinical therapies.

Materials and methods

Mice

Animals were housed in a specific pathogen-free facility at the University of Pennsylvania at ~20 °C with 55% humidity and a dark-light cycle of 12 h–12 h. Animals were provided with ad libitum access to food and water throughout the duration of the experiment. All experiments and breeding were approved by the Institutional Animal Care and Use Committee guidelines for the University of Pennsylvania. All procedures were performed in accordance with Institutional Animal Care and Use Committee Protocol 803071. Transgenic mice expressing a TCR specific for the LCMV peptide D_bGP_{33-41} (P14 donor mice) were bred in-house at the University of Pennsylvania on a C57BL/6 background purchased from Charles River. Donor mice were used at ~8 weeks of age. Recipient C57BL/6 mice were purchased from Charles River and used at 6–8 weeks of age. Recipient mice were sex matched with donor mice. Euthanasia was performed using CO_2 inhalation in a CO_2 unit as recommended by the Panel on Euthanasia of the American Veterinary Medical Association and the University of Pennsylvania.

Infections

LCMV Armstrong (Arm) and LCMV clone 13 (Cl13) were grown in house and titered as previously described[5]. Recipient mice were either infected intraperitoneally (i.p) with 2×10^5 PFU LCMV Armstrong to model an acute infection or intravenously (i.v.) with 4×10^6 PFU LCMV Cl13 to establish a chronic infection.

Adoptive cell transfer

PBMCs were isolated from the peripheral blood of naive P14 donor mice using gradient centrifugation (Histopaque-1083). 1,000 naive P14 cells were adoptively transferred i.v. into sex-matched recipient mice. P14 cells were isolated from donor mice of a distinct congenic background than recipient mice to enable donor P14 cells to be distinguished from recipient CD8 T cells. Recipient mice were infected with LCMV Arm or LCMV Cl13 one day following adoptive cell transfer.

Cell sorting

Spleens were collected at d30 of LCMV Arm and d32 of LCMV Cl13 infection. Donor P14 mice or littermates were used for the naive condition. Single cell suspensions were prepared by mechanical disruption of spleens through a 70 µm cell strainer. Red blood cells were lysed in ACK buffer (3 min, RT) and CD8 T cells isolated using EasySep CD8 T cell negative selection kit (Stem Cell, Cat# 19853) following manufacturer's instructions. CD8 T cells were washed in FACS buffer (2% FCS in PBS) and surface stained with an antibody cocktail in FACS buffer for 30 min at 4 °C. Donor P14 CD8 T cells were sorted on a BD FACS Aria II using congenic markers for identification. Samples were sorted to >95% purity.

CUT&RUN

CUT&RUN was performed as previously described with slight modifications[94,114]. 10,000 sorted cells were washed twice (600 g x 5 min) with 1 ml of cold wash buffer (20 mM HEPES-NaOH, pH 7.5, 150 mM NaCl, 0.5 mM Spermidine (Sigma 85558-1G) supplemented with protease inhibitor cocktail (Sigma 4693132001) in 1.5 ml tubes. Next, cells were resuspended in 1 ml of cold wash buffer, 20 µl of BioMagPlus Concanavalin A beads (Bangs laboratories BP531) were added and samples were mixed by rotation (4 °C, 20 min). Samples were briefly spun at 100 g, placed on DynaMag™-2 Magnet (Thermo 12321D), and liquid was removed. Primary antibodies were diluted 1:100 in 250 µl of cold antibody buffer (20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 2 mM EDTA, 0.1% digitonin (Millipore 300410-1GM) supplemented with protease inhibitor cocktails) and incubated with samples (4 °C, overnight, with rotation). The following day, samples were washed once with 1 ml cold wash buffer. Protein A-MNase (pA-MN) was diluted 1:200 in 250 µl of cold digitonin buffer (20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 0.1% digitonin supplemented with protease inhibitor cocktails) and added to samples (4 °C, 1 h, with end-to-end rotation). Samples were washed twice with 1 ml of cold digitonin buffer, resuspended in 150 µl of cold digitonin buffer and placed on a pre-cooled metal block on ice for 5 min. pA-MN digestion was initiated by adding 3 µl of 0.1 M CaCl2 to samples, mixed by gently flicking tubes 20 times and samples placed back on metal block for 30 min. Digestion was stopped by adding 150 µl of 2 x stop buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.02% Digitonin, 50 µg/ml RNase A (Thermo EN0531), 50 µg/ml Glycogen (Thermo R0561), and 4 pg/ml yeast heterologous spike-in DNA). Samples were incubated at 37 °C for 10 min and then spun (16,000 g, 5 min, 4 °C). Supernatant containing cleaved chromatin was transferred to a new tube, 3 µl of 10% SDS and 2.5 µl of 20 mg/ml proteinase K (Denville Scientific CB3210-5) were added and samples were incubated at 70 °C for 10 min, followed by phenol: chloroform: isoamyl alcohol (Thermo 15593049) and chloroform (Sigma 288306) extraction. Supernatant containing DNA (~300 µl) was transferred to new tubes pre-loaded with 20 µg of glycogen and then mixed with 750 µl of cold 100% ethanol for precipitation at -20 °C overnight. Tubes were centrifuged at 20,000 g for 30 min at 4 °C. DNA pellets were washed once with 1 ml of cold 100% ethanol, air-dried, and stored at -20 °C.

DNA libraries were prepared as previously described with slight modifications[94,115,116]. DNA pellets were dissolved in nuclease free H_2O and library preparation performed using NEBNext Ultra II DNA Library Prep Kit (NEB E7645L). Adaptor was diluted to 1:25 for adaptor ligation. For samples labeled with CTCF, DNA was barcoded and amplified for 12 PCR cycles. For histone modification samples, adaptor-ligated DNA was first selected with 25 µl and second with 45 µl of AMPure XP beads, followed by PCR amplification (H3K4me3: 11 cycles, H3K9me3: 9 cycles, H3K27ac: 14 cycles, and H3K27me3: 10–12 cycles). All libraries were cleaned up using AMPure XP beads (Beckman Coulter A63881).

CUT&RUN antibodies

Antigen	Host	Subtype	Vendor	Cat#	Lot#
Rb IgG	Guinea Pig	IgG	Antibody online	ABIN101961	39794
H3K4me3	Rb	pAb	Abcam	8580	GR3275843-1
H3K9me3	Rb	pAb	Active motif	39161	21518003
H3K27ac	Rb	IgG	Active motif	39133	28518012
H3K27me3	Rb	mAb	CST	9733 S	14
CTCF	Rb	antiserum	Millipore	07-729	3273150

RNA isolation

RNA-seq was performed as previously described with minor modifications[117]. 90,000 sorted cells were pelleted (600 g, 5 min, 4 °C) and pellets washed once with cold PBS. Pellets were resuspended in 0.5 ml of TRIzol (Thermo 15596018) and stored at -80 °C. Total RNA (~30–120 ng) was extracted using RNA Clean & Concentrator-5 (ZYMO R1013) following manufacturer's instruction and immediately followed by RNA library preparation. mRNA was isolated using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB E7490L). Libraries were prepared using NEBNext Ultra II Directional RNA Library Prep Kit (NEB E7760L) and following manufacturer's instructions.

Sequencing

Library quality was assessed using the Agilent 2100 Bioanalyzer (Agilent G2939BA) and libraries were quantified using a Qubit 2.0 fluorometer (Thermo Q32866) and by qPCR using NEBNext Library Quant Kit for Illumina (NEB E7630L) according to manufacturer's instructions. Libraries were pooled at equal molarity and sequenced with NextSeq 500/550 High Output Kit (75 cycles) v2.5 kit (Illumina 20024906) on NextSeq 550 sequencing system (Illumina SY-415-1002). 20–30 million reads for each library were sequenced using paired-end sequencing (42:6:0:42).

RNA-seq data processing and analysis

Paired-end reads were aligned and processed using STAR[118] v2.7.1a with mm10 Gencode reference genome and default parameters. Paired-end read counts of genes were quantified by featureCounts using Gencode primary assembly annotation reference genome version vM24. Genes with raw reads greater than 5 were used for downstream analysis. Normalized read counts and differential analyses were generated using DESeq2[119]. Differentially expressed genes (DEGs) were identified with filters absolute fold change > 1.5 and adjusted P-value < 0.05. All pairwise DEGs across different cell types were combined and grouped into 7 clusters using the K-mean algorithm. Heatmap plots were generated using ComplexHeatmap[120,121] packages. Other summary bar plots, violin plots, volcano plots were generated using R.

Pathway analysis of differentially expressed genes was performed using Metascape[122]. DESeq2[119] normalized factors were used to normalize bam files. Normalized bigwig files were generated using bamCoverage[123] with parameters -ignore chM --minMappingQuality 5 -ignoreDuplicates -skipNAs and were visualized in UCSC genome browser and R package Gviz[124].

ATAC-seq data analysis and processing

ATAC-seq data were aligned and processed using Bowtie2[125] v2.3.5 using mouse mm10 reference genome. Picard[126] tools v1.96 was used to remove presumed PCR duplicates using the MarkDuplicates command. Bam files containing uniquely mapped reads were created using Samtools[127] v1.1. Blacklist regions defined by ENCODE[128], random chromosomes and mitochondria were removed, and filtered bam files were used for downstream analysis.

Union peaks were downloaded from GEO. Read per million (RPM/CPM) normalized bigwig files were created using deepTools bamCoverage[123,129]. Replicates were pooled together using wiggleTools and UCSC toolkit bedGraphToBigwig[130], and tracks were imported and viewed using UCSC genome browser[131].

CUT&RUN data processing

FastQC[133] v0.11.2 and MultiQC[134] were used to check data quality. Reads were aligned to the mouse mm10 Gencode reference genome using Bowtie2[126] v2.3.5, following parameters suggested by Skene et al.[114] --local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700 -k1 -N1. Picard[126] tools v1.96 was used to remove presumed PCR duplicates using MarkDuplicates command. Bam files containing uniquely mapped reads were created using Samtools[127,135] v1.1. Fragments between 40 and 700 bp were kept. Blacklist regions defined by ENCODE, random chromosomes and mitochondria were removed, and filtered bam files were used for downstream analysis.

CUT&RUN signals were called using MACS[132] v2.1 using the broadPeak setting with adjusted P-value cutoff 0.01. In broad histone modifications, customized parameters were adjusted for different hPTMs based on length of signals and sample background variation. Consensus peaks shown in at least two biological replicates were used, and merged conditional peaks with IgG peaks removal were finally used as a union peak list for downstream quantification. Venn diagrams were generated using the ChIPpeakAnno[136] package findOverlapsOfPeaks() and makeVennDiagram() function.

Read counts were quantified across all samples based on union peak using featureCounts[137], and validated using bedtools coverage. All sample read counts were normalized using DESeq2, and principal component analysis (PCA) plots of all replicates were generated using R function prcomp. Statistical significantly differential hPTMs analyses were performed using DESeq2. The histone modification regions with adjusted P-value < 0.05 and fold change > 1.5 were defined as significantly differential modification regions, and were quantified in bar plots using ggplots. The volcano plots were generated using R ggplot2. Approximate posterior estimation for GLM shrinkage method "apeglm"[138] in DESeq2 was applied to H3K4me3 to alleviate its batch effects and final fold change calculation.

Binding motif enrichment of selected differential histone modification regions were identified using findMotifsGenome.pl from HOMER[139] v4 using each corresponding union peaks as background and size as given with mask options.

For consistent visualization, DESeq2 normalization factors were used to adjust bam files to create normalized bigwig files using bamCoverage. Bigwig files of replicates were pooled together using WiggleTools[140] mean setting. Tracks were loaded to UCSC genome browser and Gviz[124] R package for visualization. Heatmaps and metaplots were generated using deepTools plotHeatmap[123,129].

CUT&RUN data analysis

Annotations -- chromatin modification region annotation

Genes proximal to peaks (hPTMs) were annotated against mm10 genome using annotatePeaks.pl from HOMER[139] v4, ChIPseeker[141] with 10,000 base pairs (bp) flank regions, as well as GREAT[142]. Gene position information was extracted from the Gencode mm10 database, excluding pseudo genes or ambiguous undefined genes. Regions within 2,500 bp of the TSS were defined as promoters. Annotation pie chart and bar chart of gene locations were generated based on filtered gene annotation.

For one-gene-one-peak mapping, the peak with maximum variations across different conditions representing the gene were selected. For one-gene-two-peak mapping, two peaks including the gene promoter and the non-promoter peak with maximum variation were used. For one-peak-multi-gene mapping, all genes annotated to target peaks were used. For hPTM patterns analysis and gene expression correlation, one-gene-one-peak mapping was used.

Taiji transcription factors analysis

Taiji[54] analyses were performed using H3K27ac CUT&RUN bam data and RNA-seq read count data to predict key TFs. For comparison, the same analysis was performed using RNA-seq data and ATAC-seq from published papers[5,6]. The identified Taiji page rank scores of TFs were Z-score normalized across three conditions T_N, T_{MEM} and T_{EX} to identify key TFs corresponding to each condition genomewide. Heatmaps were generated based on the page rank Z-scores, and plotted using ComplexHeatmap[120,121].

Specific TF binding sites were identified using the MEME suite FIMO tool[143]. Venn diagrams of motif binding sites in H3K27ac and ATAC-seq were generated using ChIPpeakAnno[136].

Super enhancer analysis

Promoter regions were defined as regions located within 2,500 bp of TSS of each gene using the Gencode mm10 reference genome. Enhancers were defined as non-promoter regions with H3K27ac bound or ATAC-seq open accessibility. The enhancers between T_N, T_{MEM} and T_{EX} were compared using ChIPpeakAnno[136].

Super enhancers were identified using the stitching and rank ordering algorithm, ROSE[52,70], using enhancers defined by H3K27ac and ATAC-seq, respectively. In brief, nearby enhancers were stitched together, ranked and plotted by signal enrichment levels. The enhancers with signals above tangent point (slope = 1) were defined as super enhancers and the rest as typical enhancers. The SEs with only one peak stitched were further filtered out. SE were annotated to potential regulated genes using Homer[139] and GREAT[142], and further verified in the UCSC genome track. The SEs were ranked by signal intensity and plotted using R.

The comparison of SEs across T_N, T_{MEM} and T_{EX} were performed as follows. First, an initial comparison and venn diagram were generated using ChIPpeakAnno[136]. Next, the SE signal intensities for overlapping and conditional-specific SEs were quantified for each cell type by summing the normalized reads of individual enhancers under the examined region. SEs were then filtered and refined: SEs with a fold change greater than 1.5 between any two conditions were classified as conditional-specific SEs, whereas other SEs were designated as shared respectively. An updated venn diagram was generated to reflect the refined SEs.

The enrichment signal intensity of SE and TE per condition was compared, and metaplots were generated using deepTools plotProfile[123,129]. Nearest genes were used to access gene expression differences between SE and TE, which were displayed in boxplot.

Comparison of SEs identified using H3K27ac and ATAC-Seq were generated using ChIPpeakAnno[136], with hypergeometric P-values calculated for each pairwise comparison.

Bivalency analysis

Chromatin states were identified using chromHMM[75] by separating promoters with non-promoters, respectively. Promoter regions were defined as regions located within 2,500 bp of TSS of each gene and the rest of the regions were defined as non-promoter regions. ChromHMM was performed using consensus peaks of histone modifications H3K27ac, H3K27me3, H3K9me3 and ATAC with predefined 10 states, using concatenated mode with binarizing the peaks. The chromatin states coverage was quantified by base pairs. Initial states with non-low signals were further categorized into 4 major states, including (I) active, (II) poised, (III) repressed and (IV) repetitive states, based on the presence of histone modifications, and validated through heatmap plot of all marks in naive state.

Alluvial plots were generated to examine poised promoter dynamics from T_N to T_{MEM} and T_{EX} states. The significant changes were defined as histone changing with absolute fold change > 1.5. The poised-to-activated promoters were defined as either gaining H3K4me3 or losing H3K27me3 or both, and poised-to-repressed promoters were defined as either losing H3K4me3 or gaining H3K27me3 or both. The identified dynamic regions were compared to significant DEGs.

The enriched motifs were identified using Homer[139] v4 with parameters -size given and -mask, and all coding gene promoters were used as background for identifying all poised promoter motifs, while random genomes were used as background for identifying unique motifs enriched in the dynamic poised promoters. Pathway analyses were generated using Metascape on major chromatin states and dynamic promoters with consistent gene expression. Heatmaps were generated using R package ComplexHeatmap[121]. Examples of gene tracks were generated from UCSC genome tracks and Gviz[124] R package. Summary bar plots and dot plots were generated using R ggplot2.

H3K9me3 analysis

Genome-wide H3K9me3 bound regions were split into broad and narrow peaks based on 15 kb cutoff. Peaks were annotated to nearest genes and genes within 250 kb.

Mouse (mm10) repeat database was downloaded from UCSC RepeatMasker tool[144,145]. Five runs of random background controls were generated using non-H3K9me3 bound genome regions with peak amount and width matching to T_{EX}-enriched H3K9me3. The 5 random background regions were merged and used as one background control, average values were used for repeat coverage calculation.

Three levels of repeats, including repeat family, class and name, were used to calculate repeat coverage over peaks. Repeat covered regions were identified using bedtools[146] intersect of query regions with the repeat database. Repeat coverage was calculated as base pairs covered by any type of repeats.

Comparisons of genome-wide H3K9me3 bound regions with CTCF binding sites and ATAC-seq open chromatin regions were performed using R package ChIPpeakAnno[136] and tracks were generated using Gviz[124].

Data availability

RNA-seq and CUT&RUN data generated in this study is deposited in the National Center for Biotechnology Information Gene Expression Omnibus under accession numbers GSE285248 (RNA-seq) and GSE285245 (CUT&RUN). ATAC-seq data used in this study is from GSE86797.

Received: 20 December 2024; Accepted: 23 April 2025
Published online: 19 May 2025

References

1. Turner, S. J., Bennett, T. J. & Gruta, N. L. L. CD8 + T-Cell memory: The why, the when, and the how. *Cold Spring Harb Perspect. Biol.* **13**, a038661 (2021).
2. Kaech, S. M. & Cui, W. Transcriptional control of effector and memory CD8 + T cell differentiation. *Nat. Rev. Immunol.* **12**, 749–761 (2012).
3. McLane, L. M., Abdel-Hakeem, M. S. & Wherry, E. J. CD8 T cell exhaustion during chronic viral infection and Cancer. *Annu. Rev. Immunol.* **37**, 457–495 (2019).
4. Gehart, H. & Clevers, H. Tales from the Crypt: New insights into intestinal stem cells. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 19–34 (2019).
5. Pauken, K. E. et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 Blockade. *Science* **354**, 1160–1165 (2016).
6. Sen, D. R. et al. The epigenetic landscape of T cell exhaustion. *Science* **354**, 1165–1169 (2016).
7. Scott-Browne, J. P. et al. Dynamic changes in chromatin accessibility occur in CD8 + T cells responding to viral infection. *Immunity* **45**, 1327–1340 (2016).
8. Daniel, B. et al. Divergent clonal differentiation trajectories of T cell exhaustion. *Nat. Immunol.* **23**, 1614–1627 (2022).
9. Philip, M. et al. Chromatin States define tumour-specific T cell dysfunction and reprogramming. *Nature* **545**, 452–456 (2017).
10. Giles, J. R. et al. Shared and distinct biological circuits in effector, memory and exhausted CD8 + T cells revealed by Temporal single-cell transcriptomics and epigenetics. *Nat. Immunol.* **23**, 1600–1613 (2022).
11. Alfei, F. et al. TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
12. Scott, A. C. et al. TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
13. Khan, O. et al. TOX transcriptionally and epigenetically programs CD8 + T cell exhaustion. *Nature* **571**, 211–218 (2019).
14. Yao, C. et al. Single-cell RNA-seq reveals TOX as a key regulator of CD8 + T cell persistence in chronic infection. *Nat. Immunol.* **20**, 890–901 (2019).
15. Seo, H. et al. TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8 + T cell exhaustion. *Proc. Natl. Acad. Sci.* **116**, 12410–12415 (2019).
16. Wang, X. et al. TOX promotes the exhaustion of antitumor CD8 + T cells by preventing PD1 degradation in hepatocellular carcinoma. *J. Hepatol.* **71**, 731–741 (2019).
17. Sekine, T. et al. TOX is expressed by exhausted and polyfunctional human effector memory CD8 + T cells. *Sci. Immunol.* **5**, eaba7918 (2020).
18. Beltra, J. C. et al. Developmental relationships of four exhausted CD8 + T cell subsets reveals underlying transcriptional and epigenetic landscape control mechanisms. *Immunity* **52**, 825–841e8 (2020).
19. Paley, M. A. et al. Progenitor and terminal subsets of CD8 + T cells cooperate to contain chronic viral infection. *Science* **338**, 1220–1225 (2012).
20. Utzschneider, D. T. 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).
21. Im, S. J. et al. Defining CD8 + T cells that provide the proliferative burst after PD-1 therapy. *Nature* **537**, 417–421 (2016).
22. He, R. et al. Follicular CXCR5-expressing CD8 + T cells curtail chronic viral infection. *Nature* **537**, 412–416 (2016).
23. Leong, Y. A. et al. CXCR5 + follicular cytotoxic T cells control viral infection in B cell follicles. *Nat. Immunol.* **17**, 1187–1196 (2016).

24. Wu, T. et al. The TCF1-Bcl6 axis counteracts type I interferon to repress exhaustion and maintain T cell stemness. *Sci. Immunol.* **1**, eaai8593 (2016).
25. Hudson, W. H. 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**, 1043–1058e4 (2019).
26. He, S. et al. Ezh2 phosphorylation state determines its capacity to maintain CD8 + T memory precursors for antitumor immunity. *Nat. Commun.* **8**, 2125 (2017).
27. Gray, S. M., Amezquita, R. A., Guan, T., Kleinstein, S. H. & Kaech, S. M. Polycomb repressive complex 2-Mediated chromatin repression guides effector CD8 + T cell terminal differentiation and loss of multipotency. *Immunity* **46**, 596–608 (2017).
28. Li, J. et al. KDM6B-dependent chromatin remodeling underpins effective virus-specific CD8 + T cell differentiation. *Cell. Rep.* **34**, (2021).
29. Tay, R. E. et al. Hdac3 is an epigenetic inhibitor of the cytotoxicity program in CD8 T cells. *J. Exp. Med.* **217**, e20191453 (2020).
30. Niborski, L. L. et al. CD8 + T cell responsiveness to anti-PD-1 is epigenetically regulated by Suv39h1 in melanomas. *Nat. Commun.* **13**, 3739 (2022).
31. Pace, L. et al. The epigenetic control of stemness in CD8 + T cell fate commitment. *Science* **359**, 177–186 (2018).
32. Kumar, S. et al. CARM1 Inhibition enables immunotherapy of resistant tumors by dual action on tumor cells and T cells. *Cancer Discov.* **11**, 2050–2071 (2021).
33. Baxter, A. E. et al. The SWI/SNF chromatin remodeling complexes BAF and PBAF differentially regulate epigenetic transitions in exhausted CD8 + T cells. *Immunity* **56**, 1320–1340e10 (2023).
34. McDonald, B. et al. Canonical BAF complex activity shapes the enhancer landscape that licenses CD8 + T cell effector and memory fates. *Immunity* **56**, 1303–1319e5 (2023).
35. Battistello, E. et al. Stepwise activities of mSWI/SNF family chromatin remodeling complexes direct T cell activation and exhaustion. *Mol. Cell.* **83**, 1216–1236e12 (2023).
36. Belk, J. A. et al. Genome-wide CRISPR screens of T cell exhaustion identify chromatin remodeling factors that limit T cell persistence. *Cancer Cell.* **40**, 768–786e7 (2022).
37. Kharel, A. et al. Loss of PBAF promotes expansion and effector differentiation of CD8 + T cells during chronic viral infection and cancer. *Cell. Rep.* **42**, 112649 (2023).
38. Kang, T. G. et al. Epigenetic regulators of clonal hematopoiesis control CD8 T cell stemness during immunotherapy. *Science* **386**, eadl4492 (2024).
39. Carty, S. A. et al. The loss of TET2 promotes CD8 + T cell memory differentiation. *J. Immunol.* **200**, 82–91 (2018).
40. Dimitri, A. J. et al. TET2 regulates early and late transitions in exhausted CD8 + T cell differentiation and limits CAR T cell function. *Sci. Adv.* **10**, eadp9371 (2024).
41. Ladle, B. H. et al. De Novo DNA methylation by DNA methyltransferase 3a controls early effector CD8 + T-cell fate decisions following activation. *Proc. Natl. Acad. Sci. U S A.* **113**, 10631–10636 (2016).
42. Youngblood, B. et al. Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature* **552**, 404–409 (2017).
43. Ghoneim, H. E. et al. De Novo epigenetic programs inhibit PD-1 Blockade-Mediated T cell rejuvenation. *Cell* **170**, 142–157e19 (2017).
44. Abdel-Hakeem, M. S. et al. Epigenetic scarring of exhausted T cells hinders memory differentiation upon eliminating chronic antigenic stimulation. *Nat. Immunol.* **22**, 1008–1019 (2021).
45. Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* **348**, 56–61 (2015).
46. Curran, M. A., Montalvo, W., Yagita, H. & Allison, J. P. PD-1 and CTLA-4 combination Blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. *Proc. Natl. Acad. Sci.* **107**, 4275–4280 (2010).
47. Butterfield, L. H. & Najjar, Y. G. Immunotherapy combination approaches: mechanisms, biomarkers and clinical observations. *Nat. Rev. Immunol.* **24**, 399–416 (2024).
48. Pircher, H. et al. Molecular analysis of the antigen receptor of virus-specific cytotoxic T cells and identification of a new Vα family. *Eur. J. Immunol.* **17**, 1843–1846 (1987).
49. Pircher, H. et al. Characterization of virus-specific cytotoxic T cell clones from allogeneic bone marrow chimeras. *Eur. J. Immunol.* **17**, 159–166 (1987).
50. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why?? *Mol. Cell.* **49**, 825–837 (2013).
51. Chen, Z. et al. TCF-1-Centered transcriptional network drives an effector versus exhausted CD8 T Cell-Fate decision. *Immunity* **51**, 840–855e5 (2019).
52. Hnisz, D. et al. Super-Enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
53. Creyghton, M. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U S A.* **107**, 21931–21936 (2010).
54. Zhang, K., Wang, M., Zhao, Y., Wang, W. & Taiji System-level identification of key transcription factors reveals transcriptional waves in mouse embryonic development. *Sci. Adv.* **5**, eaav3262 (2019).
55. Willinger, T. et al. Human Naive CD8 T cells Down-Regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in Vivo1. *J. Immunol.* **176**, 1439–1446 (2006).
56. Zhao, X., Shan, Q. & Xue, H. H. TCF1 in T cell immunity: a broadened frontier. *Nat. Rev. Immunol.* **22**, 147–157 (2022).
57. Pipkin, M. E. Runx proteins and transcriptional mechanisms that govern memory CD8 T cell development. *Immunol. Rev.* **300**, 100–124 (2021).
58. Quigley, M. et al. Transcriptional analysis of HIV-specific CD8 + T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nat. Med.* **16**, 1147–1151 (2010).
59. Chen, Y. et al. BATF regulates progenitor to cytolytic effector CD8 + T cell transition during chronic viral infection. *Nat. Immunol.* **22**, 996–1007 (2021).
60. Martinez, G. J. et al. The transcription factor NFAT promotes exhaustion of activated CD8 + T cells. *Immunity* **42**, 265–278 (2015).
61. Wherry, E. J. et al. Molecular signature of CD8 + T cell exhaustion during chronic viral infection. *Immunity* **27**, 670–684 (2007).
62. Man, K. et al. Transcription factor IRF4 promotes CD8 + T cell exhaustion and limits the development of Memory-like T cells during chronic infection. *Immunity* **47**, 1129–1141e5 (2017).
63. Liu, X. et al. Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. *Nature* **567**, 525–529 (2019).
64. Chen, J. et al. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530–534 (2019).
65. Guan, T. et al. ZEB1, ZEB2, and the miR-200 family form a counterregulatory network to regulate CD8 + T cell fates. *J. Exp. Med.* **215**, 1153–1168 (2018).
66. Omilusik, K. D. et al. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8 + effector and memory T cell populations during infection. *J. Exp. Med.* **212**, 2027–2039 (2015).
67. Dominguez, C. X. et al. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J. Exp. Med.* **212**, 2041–2056 (2015).
68. Wu, J. E. et al. In vitro modeling of CD8 $^{+}$ T cell exhaustion enables CRISPR screening to reveal a role for BHLHE40. *Sci. Immunol.* **8**, eade3369 (2023).
69. Shen, W. K. et al. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res.* **51**, D39–D45 (2023).

70. Whyte, W. A. et al. Master transcription factors and mediator Establish Super-Enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
71. Lovén, J. et al. Selective Inhibition of tumor oncogenes by disruption of Super-Enhancers. *Cell* **153**, 320–334 (2013).
72. Li, C. et al. The transcription factor Bhlhe40 programs mitochondrial regulation of resident CD8 + T cell fitness and functionality. *Immunity* **51**, 491–507e7 (2019).
73. Salmon, A. J. et al. BHLHE40 regulates the T-Cell effector function required for tumor microenvironment remodeling and immune checkpoint therapy efficacy. *Cancer Immunol. Res.* **10**, 597–611 (2022).
74. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499 (2015).
75. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with chromhmm. *Nat. Protoc.* **12**, 2478–2492 (2017).
76. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
77. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
78. Macrae, T. A., Fothergill-Robinson, J. & Ramalho-Santos, M. Regulation, functions and transmission of bivalent chromatin during mammalian development. *Nat. Rev. Mol. Cell. Biol.* **24**, 6–26 (2023).
79. Lesch, B. J. & Page, D. C. Poised chromatin in the mammalian germ line. *Development* **141**, 3619–3626 (2014).
80. Lesch, B. J., Dokshin, G. A., Young, R. A., McCarrey, J. R. & Page, D. C. A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. *Proc. Natl. Acad. Sci.* 110, 16061–16066 (2013).
81. Russ, B. E. et al. Distinct epigenetic signatures delineate transcriptional programs during Virus-Specific CD8 + T cell differentiation. *Immunity* **41**, 853–865 (2014).
82. Renkema, K. R. et al. KLRG1 + memory CD8 T cells combine properties of Short-Lived effectors and Long-Lived memory. *J. Immunol.* **205**, 1059–1069 (2020).
83. Fang, F. et al. The cell-surface 5′-nucleotidase CD73 defines a functional T memory cell subset that declines with age. *Cell. Rep.* **37**, 109981 (2021).
84. Zhang, Y. et al. Impaired apoptosis, extended duration of immune responses, and a lupus-like autoimmune disease in IEX-1-transgenic mice. *Proc. Natl. Acad. Sci.* **99**, 878–883 (2002).
85. Good, C. R. et al. An NK-like CAR T cell transition in CAR T cell dysfunction. *Cell* **184**, 6081–6100e26 (2021).
86. Cannarile, M. A. et al. Transcriptional regulator Id2 mediates CD8 + T cell immunity. *Nat. Immunol.* **7**, 1317–1325 (2006).
87. Yang, C. Y. et al. The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8 + T cell subsets. *Nat. Immunol.* **12**, 1221–1229 (2011).
88. Li, Y. et al. Id2 epigenetically controls CD8 + T-cell exhaustion by disrupting the assembly of the Tcf3-LSD1 complex. *Cell. Mol. Immunol.* **21**, 292–308 (2024).
89. Quigley, M., Huang, X. & Yang, Y. STAT1 signaling in CD8 T cells is required for their clonal expansion and memory formation following viral infection in vivo. *J. Immunol.* **180**, 2158–2164 (2008).
90. Tripathi, P. et al. STAT5 is critical to maintain effector CD8 + T cell responses. *J. Immunol.* **185**, 2116–2124 (2010).
91. Oyake, T. et al. Bach proteins belong to a novel family of BTB-Basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. *Mol. Cell. Biol.* **16**, 6083–6095 (1996).
92. Utzschneider, D. T. et al. Early precursor T cells Establish and propagate T cell exhaustion in chronic infection. *Nat. Immunol.* **21**, 1256–1266 (2020).
93. Yao, C. et al. BACH2 enforces the transcriptional and epigenetic programs of stem-like CD8 + T cells. *Nat. Immunol.* **22**, 370–380 (2021).
94. Chen, Z. et al. In vivo CD8 + T cell CRISPR screening reveals control by Fli1 in infection and cancer. *Cell* **184**, 1262–1280e22 (2021).
95. Gerondakis, S., Fulford, T. S., Messina, N. L. & Grumont, R. J. NF-κB control of T cell development. *Nat. Immunol.* **15**, 15–25 (2014).
96. Gerondakis, S. & Siebenlist, U. Roles of the NF-κB pathway in lymphocyte development and function. *Cold Spring Harb Perspect. Biol.* **2**, a000182 (2010).
97. Pichler, A. C. et al. TCR-independent CD137 (4-1BB) signaling promotes CD8+-exhausted T cell proliferation and terminal differentiation. *Immunity* **56**, 1631–1648e10 (2023).
98. Nicetto, D. & Zaret, K. S. Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr. Opin. Genet. Dev.* **55**, 1–10 (2019).
99. Padeken, J., Methot, S. P. & Gasser, S. M. Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat. Rev. Mol. Cell. Biol.* **23**, 623–640 (2022).
100. Katznelson, A. et al. Heterochromatin protein ERH represses alternative cell fates during early mammalian differentiation. 06.06.597604 Preprint at (2024). https://doi.org/10.1101/2024.06.06.597604 (2024).
101. Barral, A. et al. SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Mol. Cell.* **82**, 816–832e12 (2022).
102. Tam, P. L. F., Cheung, M. F., Chan, L. Y. & Leung, D. Cell-type differential targeting of SETDB1 prevents aberrant CTCF binding, chromatin looping, and cis-regulatory interactions. *Nat. Commun.* **15**, 15 (2024).
103. Liao, X. et al. Repetitive DNA sequence detection and its role in the human genome. *Commun. Biol.* **6**, 1–21 (2023).
104. Diehl, A. G., Ouyang, N. & Boyle, A. P. Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. *Nat. Commun.* **11**, 1796 (2020).
105. Bourque, G. et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* **18**, 1752–1762 (2008).
106. Kim, S., Yu, N. K. & Kaang, B. K. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp. Mol. Med.* **47**, e166 (2015).
107. Gualdrini, F. et al. H3K9 trimethylation in active chromatin restricts the usage of functional CTCF sites in SINE B2 repeats. *Genes Dev.* **36**, 414–432 (2022).
108. Becker, J. S., Nicetto, D. & Zaret, K. H3K9me3-Dependent heterochromatin: barrier to cell fate changes. *Trends Genet.* **32**, 29–41 (2016).
109. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553 (2007).
110. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
111. Khoa, L. T. P. et al. Quiescence enables unrestricted cell fate in Naive embryonic stem cells. *Nat. Commun.* **15**, 1721 (2024).
112. P Mognol, G. et al. Exhaustion-associated regulatory regions in CD8 + tumor-infiltrating T cells. *Proc. Natl. Acad. Sci. U S A.* **114**, E2776–E2785 (2017).
113. Riegel, D. et al. Integrated single-cell profiling dissects cell-state-specific enhancer landscapes of human tumor-infiltrating CD8 + T cells. *Mol. Cell.* **83**, 622–636e10 (2023).
114. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).
115. Liu, X. S. et al. Rescue of fragile X syndrome neurons by DNA methylation editing of the *FMR1* gene. *Cell* **172**, 979–992e6 (2018).

116. Cao, Z. et al. ZMYND8-regulated IRF8 transcription axis is an acute myeloid leukemia dependency. *Mol. Cell.* **81**, 3604–3622e10 (2021).
117. Zhang, Z. et al. Efficient engineering of human and mouse primary cells using peptide-assisted genome editing. *Nat. Biotechnol.* **42**, 305–315 (2024).
118. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
119. Love, M. I., Huber, W. & Anders, S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
120. Gu, Z. Complex heatmap visualization. *iMeta* **1**, e43 (2022).
121. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
122. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
123. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187 (2014).
124. Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and bioconductor. in Statistical Genomics: Methods and Protocols (eds Mathé, E. & Davis, S.) 335–351 (Springer, New York, NY, doi:https://doi.org/10.1007/978-1-4939-3578-9_16. (2016).
125. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nat. Methods.* **9**, 357–359 (2012).
126. Picard Tools - By Broad Institute. https://broadinstitute.github.io/picard/
127. Danecek, P. et al. Twelve years of samtools and BCFtools. *GigaScience* **10**, giab008 (2021).
128. ENCODE. https://www.encodeproject.org/
129. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
130. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and bigbed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).
131. UCSC Genome Browser Home. https://genome.ucsc.edu/index.html
132. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
133. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: (2010). http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
134. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
135. Li, H. et al. The sequence alignment/map format and samtools. *Bioinformatics* **25**, 2078–2079 (2009).
136. Zhu, L. J. et al. ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinform.* **11**, 237 (2010).
137. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
138. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).
139. Heinz, S. et al. Simple combinations of Lineage-Determining transcription factors prime *cis*-Regulatory elements required for macrophage and B cell identities. *Mol. Cell.* **38**, 576–589 (2010).
140. Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P. & Flicek, P. WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* **30**, 1008–1009 (2014).
141. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: An R/Bioconductor package for chip peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
142. McLean, C. Y. et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
143. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
144. Flynn, J. et al. (ed, M.) RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci.* **117** 9451–9457 (2020).
145. Jurka, J. Repbase update: A database and an electronic journal of repetitive elements. *Trends Genet.* **16**, 418–420 (2000).
146. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

Acknowledgements

Protein A-MNase (batch 6) and yeast heterologous spike-in DNA were kindly provided by Dr. Steve Henikoff. The authors also thank Terri D. Bryson from Henikoff Laboratory for sharing the pA-MNase purification protocol.

Author contributions

E.J.W. and S.L.B. conceived the project. A.E.B., Z.C., Z.Z, P.A.AG., E.J.W. and S.L.B. designed the experiments. A.E.B., Z.C., Z.Z., P.A.AG., P.S. and S.C. performed experiments. H.H. analyzed data and prepared figures with help from A.E.B., Z.Z., C.R.G and K.A.A. K.M.G., L.W., G.D., S.M., J.R.G. and J.S. consulted on data analysis. H.H., A.E.B., Z.Z., C.R.G, K.A.A., S.L.B. and E.J.W. wrote the manuscript. All authors reviewed the manuscript.

Declarations

Competing interests

E.J.W. is a member of the Parker Institute for Cancer Immunotherapy. E.J.W. is an advisor for Arsenal Biosciences, Coherus, Danger Bio, IpiNovyx, New Limit, Marengo, Pluto Immunotherapeutics, Prox Bio, Related Sciences, Santa Ana Bio, and Synthekine. E.J.W. is a founder of Arsenal Biosciences, Danger Bio, Prox Bio and holds stock in Coherus. J.R.G. is a consultant for Arsenal Biosciences, Cellanome, GVM1, and Seismic Therapeutics. The remaining authors declare no competing interests.

Ethical approval

This study is reported in accordance with ARRIVE guidelines.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/1

0.1038/s41598-025-99804-0.

Correspondence and requests for materials should be addressed to S.L.B. or E.J.W.