

## Distinct transcriptomic and epigenomic modalities underpin human memory T cell subsets and their activation potential

James R. Rose<sup>1,3</sup>, Bagdeser Akdogan-Ozdilek<sup>1,3</sup>, Andrew R. Rahmberg<sup>2</sup>, Michael D. Powell<sup>1</sup>, Sakeenah L. Hicks<sup>1</sup>, Christopher D. Scharer<sup>1</sup> & Jeremy M. Boss<sup>1</sup>✉

Human memory T cells (MTC) are poised to rapidly respond to antigen re-exposure. Here, we derived the transcriptional and epigenetic programs of resting and ex vivo activated, circulating CD4<sup>+</sup> and CD8<sup>+</sup> MTC subsets. A progressive gradient of gene expression from naïve to T<sub>CM</sub> to T<sub>EM</sub> is observed, which is accompanied by corresponding changes in chromatin accessibility. Transcriptional changes suggest adaptations of metabolism that are reflected in altered metabolic capacity. Other differences involve regulatory modalities comprised of discrete accessible chromatin patterns, transcription factor binding motif enrichment, and evidence of epigenetic priming. Basic-helix-loop-helix factor motifs for AHR and HIF1A distinguish subsets and predict transcription networks to sense environmental changes. Following stimulation, primed accessible chromatin correlate with an augmentation of MTC gene expression as well as effector transcription factor gene expression. These results identify coordinated epigenetic remodeling, metabolic, and transcriptional changes that enable MTC subsets to ultimately respond to antigen re-encounters more efficiently.

<sup>1</sup>Department of Microbiology and Immunology, and the Emory Vaccine Center, Emory University School of Medicine, Atlanta, GA 30322, USA. <sup>2</sup>Barrier Immunity Section, Laboratory of Viral Diseases, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD, USA. <sup>3</sup>These authors contributed equally: James R. Rose, Bagdeser Akdogan-Ozdilek. ✉email: [jmboss@emory.edu](mailto:jmboss@emory.edu)

Memory T cells (MTC) arise after a naïve T cell responds to an initial interaction with antigen and play an essential role in mounting a robust secondary response upon reinfection. These cells, composed of both CD4 and CD8 expressing T cells, exhibit enhanced effector functions and a heightened response to subsequent activations with antigen<sup>1</sup>. Human MTC are extremely long-lived, with one study finding antigen-specific T cells still detectable in individuals 75 years after vaccination, suggesting a capacity for self-renewal not seen in shorter-lived effector cells<sup>2</sup>. Naïve and memory CD4<sup>+</sup> and CD8<sup>+</sup> T-cell subsets are commonly subdivided based on cell surface expression of CCR7, CD45RA, and CD62L into naïve (Nav, CCR7<sup>+</sup>CD45RA<sup>+</sup>), central memory (T<sub>CM</sub>, CCR7<sup>+</sup>CD45RA<sup>-</sup>), effector memory (T<sub>EM</sub>, CCR7<sup>-</sup>CD45RA<sup>-</sup>), and terminally differentiated effector memory (T<sub>EMRA</sub>, CCR7<sup>-</sup>CD45RA<sup>+</sup>)<sup>3,4</sup>. CCR7 and CD62L are highly expressed in T<sub>CM</sub> and naïve T cells and enable cellular homing to secondary lymphoid organs. T<sub>EM</sub> exhibit lower expression of both CCR7 and CD62L and are thought to instead localize to inflamed tissues where they exhibit a higher degree of effector function<sup>5</sup>. Additionally, naïve T cells can be separated from T<sub>EM</sub> and T<sub>CM</sub> cells by higher expression of the CD45RA isoform<sup>6</sup>, which is downregulated in T<sub>CM</sub> and T<sub>EM</sub> cells. However, a subset of effector-memory-like human CD8<sup>+</sup> T cells in the blood express high levels of CD45RA (CCR7<sup>-</sup>CD45RA<sup>+</sup>) while simultaneously expressing effector function genes such as *GZMB* and *PRF1* (encoding granzyme B and perforin, respectively)<sup>7</sup>. These cells (T<sub>EMRA</sub>) are believed to be highly differentiated effector cells or potentially even senescent versions of effector MTC, which arise after chronic infections with virus<sup>8–10</sup>. Despite their importance in immunological memory, the full spectrum of transcriptional and epigenetic characteristics of these main memory subsets, the transcription factor programs associated with each, and how they play a role in memory responses remains to be fully understood.

Previous studies have profiled various epigenetic and transcriptional aspects of CD8<sup>+</sup> MTC<sup>11–13</sup>; however, the full phenotypic and epigenetic characteristics of the subsets found within the entire circulating human MTC compartment remains largely under-defined. At the same time, changes to the transcriptomic and epigenetic landscape of T cells after initial activation of naïve cells have been shown to be important in both establishing the differentiated memory subsets, and in rapid recall response upon re-stimulation<sup>14–16</sup>. Understanding how these memory-subset-specific epigenetic changes affect their formation and function upon re-encounter with antigen will ultimately inform more effective therapeutic design.

To better understand the epigenomic parameters of human MTC subsets and how such parameters dictate gene expression, we investigated transcriptional and epigenetic differences between CD4<sup>+</sup> and CD8<sup>+</sup> memory cell subtypes (T<sub>CM</sub>, T<sub>EM</sub>, and T<sub>EMRA</sub>) from human blood, as well as in response to *ex vivo* stimulation. Analysis of changes in mRNA transcripts and chromatin accessibility revealed that MTC share a substantial set of genes expressed at similar levels irrespective of lineage and cell subset. Moreover, we observed evidence of a progressive increase in the amount of differentiation from naïve T cells, to T<sub>CM</sub>, and finally to T<sub>EM</sub> and T<sub>EMRA</sub> memory populations. One set of genes that were upregulated, included those that reflected changes in metabolic capacity of the MTC subsets. Biochemical analyses confirmed differences in the metabolic capacity of various MTC subsets. Combining gene expression and chromatin accessibility analyses before and after stimulation identified a series of patterned regulatory modalities that may define memory-subset differentiation, as well as increased reactivity in response to secondary activations. Such dynamic regions of chromatin were enriched for motifs known to bind a small group of transcription

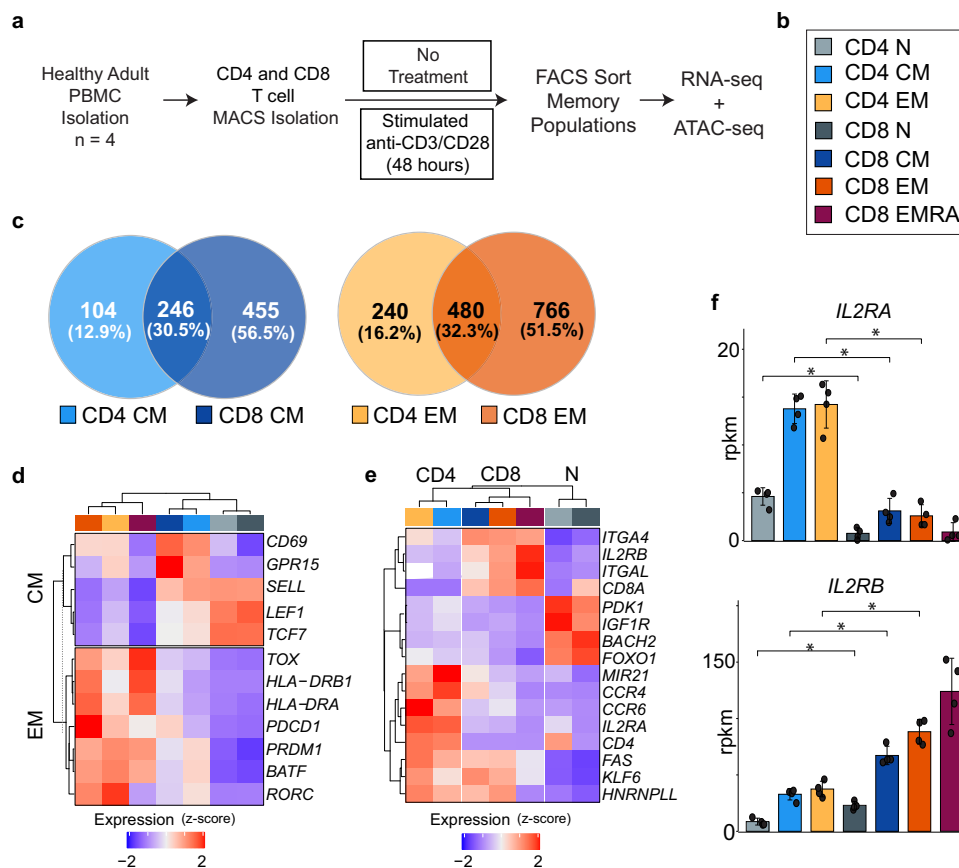
factors (TFs) from the basic leucine zipper (bZIP), high mobility group (HMG), T-box, and basic helix-loop-helix (bHLH) families. Altogether, these data lay out the full spectrum of transcriptional and epigenetic differences of the primary memory-subset categories found in human blood, while also identifying the unique transcription factor networks associated with memory-subset differentiation and highlighting loci that may be important in establishing a memory cell's response to activation upon secondary antigen exposure.

## Results

**The shared transcriptional programs of human CD4<sup>+</sup> and CD8<sup>+</sup> MTC.** To define the transcriptional profiles of the major subsets of MTC, naïve, T<sub>CM</sub>, and T<sub>EM</sub> cells from CD8<sup>+</sup> and CD4<sup>+</sup> lineages were FACS separated from the blood of four human donors using CCR7 and CD45RA as distinguishing cell surface markers (Fig. 1a, b, and Supplemental Fig. 1a). Naïve T cells were defined as CCR7<sup>+</sup>CD45RA<sup>+</sup>; T<sub>CM</sub> as CCR7<sup>+</sup>CD45RA<sup>-</sup>; and T<sub>EM</sub> as CCR7<sup>-</sup>CD45RA<sup>-</sup>. CD8<sup>+</sup> T<sub>EMRA</sub> T cells (CD45RA<sup>+</sup>CCR7<sup>-</sup>), a terminal effector subset was also included for comparison (Supplemental Fig. 1b-c). To elucidate transcriptional differences between naïve and memory subsets RNA-sequencing (RNA-seq) was conducted on sorted populations. The mRNA content of each sorted subset group was assessed for shared and unique transcripts defining MTC from naïve T cells. Relative to their naïve counterparts, T<sub>CM</sub> and T<sub>EM</sub> cells exhibited 805 and 1,486 differentially expressed genes (DEG), respectively (Fig. 1c, Supplemental Data 1). Comparing subsets between lineages (CD4<sup>+</sup> vs. CD8<sup>+</sup>) showed that T<sub>EM</sub> and T<sub>CM</sub> cells each shared approximately 30% of their DEG, highlighting conserved relationships between the CD4<sup>+</sup> and CD8<sup>+</sup> lineages. Gene ontology analysis of the lineage specific and shared DEG showed that genes shared between lineages were highly enriched for pathways related to T-cell activation, co-stimulation, lymphocyte homeostasis, and cytokine response/production in both T<sub>CM</sub> (Supplemental Fig. 2a) and T<sub>EM</sub> subsets (Supplemental Fig. 2b). T<sub>EM</sub> DEG shared between CD4<sup>+</sup> and CD8<sup>+</sup> T cells were also more highly enriched for antigen processing and presentation, as well as GTPase signaling pathways.

Examples of genes expressed in both lineages included the transcription factors *LEF1* and *TCF7*, which were more highly expressed in both CD4<sup>+</sup> and CD8<sup>+</sup> T<sub>CM</sub> and naïve T cells. *PDCD1*, encoding PD-1, and MHC class II genes encoding HLA-DR were expressed higher in T<sub>EM</sub> cells of both lineages (Fig. 1d). Other genes exhibited lineage-specific expression patterns, including integrins (e.g., *ITGAL/CD11a*, *ITGA4/VLA4*), which were expressed more highly in CD8<sup>+</sup> MTC; whereas the chemokine receptors *CCR4* and *CCR6*, which have been implicated in homing to specific peripheral tissues such as the skin<sup>17,18</sup>, were more expressed in CD4<sup>+</sup> MTC (Fig. 1e). Lineage-specific DEG encoding components of the interleukin-2 receptor (IL2R)<sup>19</sup> were also observed, with CD4<sup>+</sup> MTC expressing higher levels of the  $\alpha$  chain (*IL2RA*, Fig. 1f), while CD8<sup>+</sup> MTC expressed considerably higher levels of *IL2RB* (Fig. 1f). These differences emphasize the fact that while all MTC share important gene pathways, which make them unique from naïve cells, transcriptional differences can also distinguish each lineage and/or subset underlying their distinct immunological functions.

**Resting T<sub>EM</sub>/T<sub>EMRA</sub> cells exhibit progressively greater transcriptional differentiation from naïve progenitors than T<sub>CM</sub>.** A total of 4,943 DEG distinguished naïve cells from CD4<sup>+</sup> and CD8<sup>+</sup> MTC subsets. Principal component analysis (PCA) of these DEG indicated that the bulk of this variation separated naïve T cells from T<sub>CM</sub>, T<sub>EM</sub>, and T<sub>EMRA</sub> subsets (Fig. 2a).



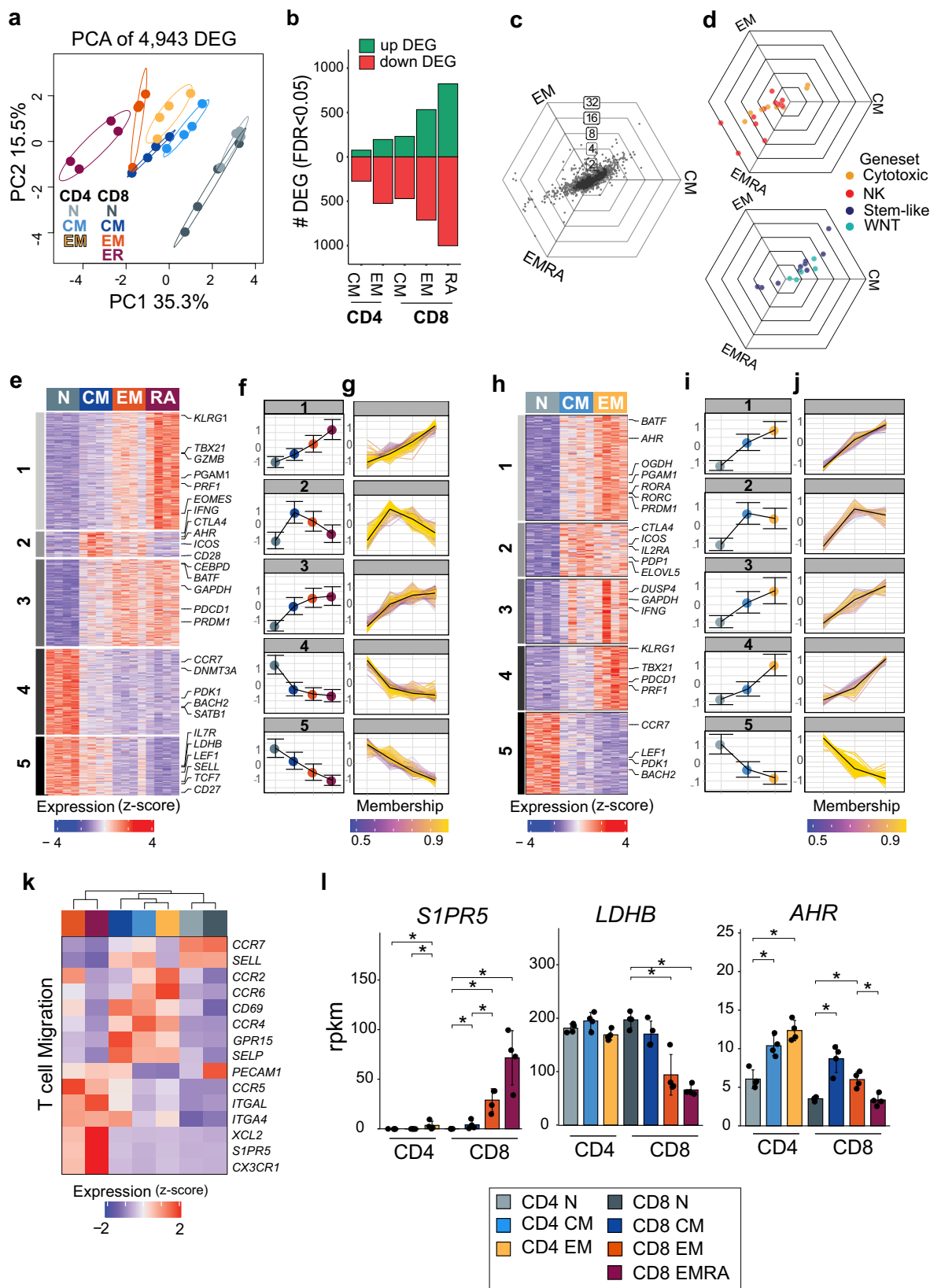
**Fig. 1** MTC display distinct patterns of shared and differentially expressed genes that define their subsets and lineage. **a** Experimental workflow for isolation, ex vivo stimulation, and sequencing analysis of MTC from human blood samples. **b** Color code of naïve T cell and MTC lineages and subsets used throughout are shown. **c** Venn diagrams representing total number of DEG for  $T_{CM}$  and  $T_{EM}$  compared to naïve T cells. **d** Heatmap and hierarchical clustering of gene expression for select genes commonly expressed in either  $T_{CM}$  or  $T_{EM}$  regardless of lineage. **e** Heatmap of lineage-specific gene expression in MTC. **f** Gene expression bar plots (reads per kilobase million, rpkm) for indicated IL-2 receptor genes. Data are plotted as mean  $\pm$  SD (standard deviation); asterisks indicate DEG (FDR  $\leq$  0.05) as detected by DESeq2 algorithm. Cell type color codes are shown in (b).

Differentiation from naïve cells showed progressive increases in both up and down DEG from  $T_{CM}$  to  $T_{EM}$  in both  $CD4^+$  and  $CD8^+$  T cells and finally to  $T_{EMRA}$  in  $CD8^+$  T cells subsets (Fig. 2b). Fold-change levels of DEG in  $CD8^+$  T-cell memory subsets were compared using hexagonal “tri-wise” visualization (Fig. 2c)<sup>20</sup>. These plots contain three axes corresponding to each subset and each radiating axis represents fold-change levels. DEG that fall directly on each axis represent those expressed exclusively by that subset, while those sharing high fold-change levels across two groups relative to the third subset fall midway between spokes. This analysis revealed that many of the DEG with highest fold change were shared between  $T_{EM}$  and  $T_{EMRA}$  subsets but absent in  $T_{CM}$  cells. Overlaying genes from several gene ontology gene sets relevant to MTC showed that genes involved in pathways related to cell cytotoxicity as well as NK-related genes involved in cell killing were exclusively upregulated in  $T_{EM}$  and  $T_{EMRA}$  subsets (Fig. 2d). Alternatively, a custom gene set representing genes expressed by stem-like T cells derived from Hudson et al.<sup>21</sup>, as well as gene ontology gene sets representing the WNT-beta catenin pathway were highly upregulated in the  $T_{CM}$  subset relative to either effector subset (Fig. 2d).

Fuzzy c-means clustering of the DEG was used to identify gene modules across the  $CD8^+$  and  $CD4^+$  MTC subsets. Fuzzy c-means clustering allows genes to be assigned membership to multiple overlapping clusters before an ultimate single cluster classification is determined by ranking the magnitude of membership score for that gene. This analysis identified five

distinct modules of gene expression programs within the DEG of  $CD8^+$  T cells across all memory subsets (Fig. 2e, Supplemental Data 2). Modules 1 and 3 corresponded to genes that were upregulated in  $CD8^+$   $T_{EM}$  and  $T_{EMRA}$  cells compared to naïve T cells. Gene ontology analysis showed that genes in module 1 were most highly expressed in  $T_{EMRA}$  cells and enriched for pathways related to cell killing, cytolysis, and innate immune response (Supplemental Fig. 3a). Module 3 genes were expressed equally in  $CD8^+$   $T_{EM}$  and  $T_{EMRA}$  cells and enriched for pathways related to apoptosis, MAPK activity, and tyrosine kinases (Supplemental Fig. 3b). Modules 4 and 5 represented genes repressed in  $CD8^+$   $T_{EM}$  and  $T_{EMRA}$  subsets but more highly expressed in naïve or  $T_{CM}$  cells (Fig. 2f) with individual genes plotted and colored according to cluster membership (Fig. 2g). These were enriched for DNA methyltransferases such as DNMT3A or pathways related to T-cell differentiation, respectively (Supplemental Figs. 3c, d). Module 2 contained fewer genes than any of the other modules and included those expressed in both  $CD8^+$   $T_{CM}$  and  $T_{EM}$  memory subsets with low expression in both naïve and  $T_{EMRA}$  cells. This module was highly enriched for pathways related to T-cell receptor activation, co-stimulation, and IL-2 production (Supplemental Fig. 3e).

Applying the same clustering method to  $CD4^+$  subsets showed a similar pattern of clustering with five total gene modules (Fig. 2h, i, Supplemental Data 3) with individual genes plotted (Fig. 2j). Modules 1, 3, and 4 were most highly expressed in  $CD4^+$   $T_{EM}$  cells (Fig. 2i) but showed differences in gene ontology



pathway enrichment (Supplemental Fig. 3f-h). These included secretory granules and carbohydrate metabolism (module 1), cytokine secretion, and T-cell activation (module 3), and MAP Kinase activity or cytolysis (module 4). The expression of genes in module 5 was highest in CD4<sup>+</sup> naïve T cells and contained similar gene set enrichments to module 5 in CD8<sup>+</sup> MTC

(Supplemental Fig. 3i), including self-renewal related transcription factors, such as *LEF1*. Module 2 was expressed in both CD4<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub> subsets as in the CD8<sup>+</sup> lineage. Pathway enrichment for this module showed enrichment for several metabolism-related pathways including cofactor and lipid biosynthesis (Supplemental Fig. 3j). As with CD8<sup>+</sup> MTC,

**Fig. 2 Resting MTC subsets in peripheral blood show progressive levels of differentiation from naïve.** **a** Principal component (PC) analysis plot of DEG from RNA-sequencing data showing each sample separated by first two components. **b** Bar plot representing number of up and downregulated DEG for each memory subset compared to naïve T cells. **c** Three-way representation of DEG expression between CD8<sup>+</sup> T<sub>CM</sub>, T<sub>EM</sub>, and T<sub>EMRA</sub>. Each dot represents a single DEG; dots falling directly on an axis represent exclusive upregulation in each group, and concentric rings represent degree of log<sub>2</sub>-fold-change differences between groups. **d** Overlay of C with genes belonging to two gene sets (NK cell and cytotoxicity) upregulated in CD8<sup>+</sup> T<sub>EM</sub> and T<sub>EMRA</sub> cells, or genes representing WNT-Beta catenin and stem-like T-cell expression are upregulated in CD8<sup>+</sup> T<sub>CM</sub> cells. **e, h** Heatmaps showing DEGs in CD8<sup>+</sup> or CD4<sup>+</sup> naïve and memory subsets clustered by fuzzy c-means clustering. Highlighted examples for each module are listed. **f, i** Line plots showing mean expression of all genes in each cluster by subset (error bars represent  $\pm 1$  SD). **g, j** Line plots showing expression of individual genes in cluster colored by membership score for respective cluster. Average across modules shown by black line. **k** Heatmap of selected genes representing differences in expression of migration-related genes. Data represents the mean expression of each cell subset. **l** Gene expression bar plots showing reads per kilobase million (rpkm) for the indicated genes. Data are plotted as mean  $\pm$  SD; asterisks indicate DEG as detected by DESeq2 algorithm.

important co-stimulatory/co-inhibitory molecules, such as CTLA4 and ICOS, were also present in CD4<sup>+</sup> module 2 and expressed by both T<sub>CM</sub> and T<sub>EM</sub> subsets. Thus, both CD4<sup>+</sup> and CD8<sup>+</sup> lineages showed a progression of increasing gene expression changes relative to naïve in T<sub>CM</sub>, T<sub>EM</sub>, and finally T<sub>EMRA</sub> MTC subsets.

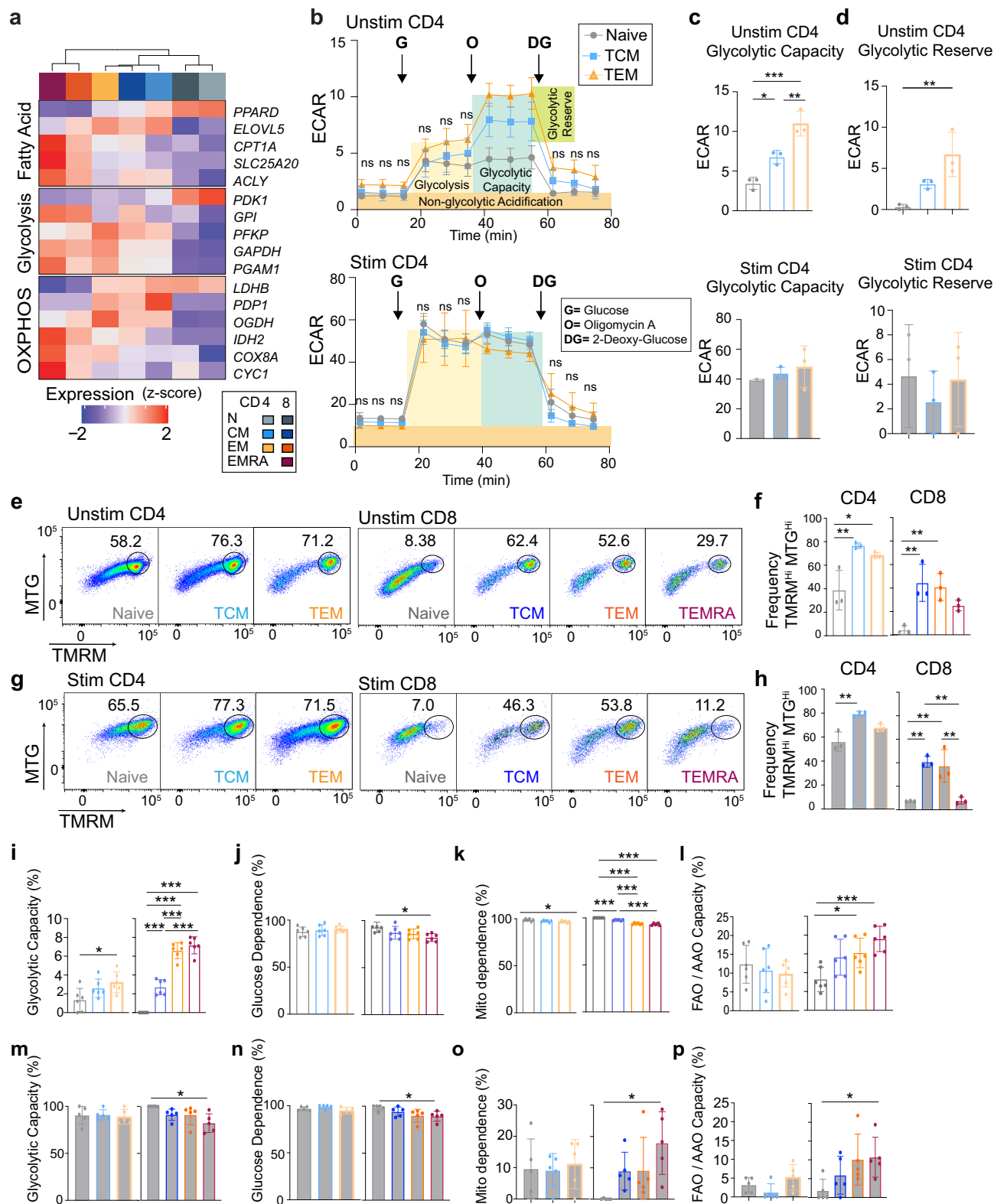
**MTC subsets exhibit distinct migration and metabolism characteristics.** Contained within the above modules were genes representing the potential for important functional differences across MTC subsets, including genes involved in T-cell migration and metabolism. For example as expected<sup>3</sup>, CD4<sup>+</sup> and CD8<sup>+</sup> naïve and T<sub>CM</sub> cells expressed high levels of both *CCR7* and *SELL* (L-selectin). Whereas CD4<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub> exclusively expressed *CCR4* and *CCR2* (Fig. 2k). CD8<sup>+</sup> T<sub>EM</sub>/T<sub>EMRA</sub> cells had the highest expression of *SIPR5*, which has been shown to be associated with promoting egress of lymphocytes from secondary lymphoid organs or bone marrow (Fig. 2l)<sup>22</sup>. MTC in general were found to have upregulated a greater number of metabolism-related genes compared to resting naïve T cells (Fig. 3a). Subset-specific differences were observed in expression of genes responsible for fatty acid metabolism, glycolysis, and oxidative phosphorylation (Fig. 3a). Many of these genes relate to the regulation of acetyl CoA or lactate metabolism as exemplified by expression differences of the genes *PDK1*, *PDPI*, and *LDHB*. For example, in CD8<sup>+</sup> T<sub>EM</sub> and T<sub>EMRA</sub> cells (Fig. 2j), *LDHB*, which encodes lactate dehydrogenase enzyme subunit B has been previously associated with aerobic glycolytic metabolism in effector T cells, as well as in cancer<sup>23,24</sup>.

Several assays to measure and compare the metabolic states of naïve and MTC were performed. Using the Seahorse-based Glycolysis Stress Test assay on unstimulated CD4<sup>+</sup> T-cell subsets, an extracellular acidification rate (ECAR) was determined (Fig. 3b). In response to glucose, CD4<sup>+</sup> MTC did not exhibit a significant difference in glycolysis; however, T<sub>CM</sub> and T<sub>EM</sub> populations showed significantly higher glycolytic capacity (Fig. 3c) and glycolytic reserve (Fig. 3d) compared to naïve CD4<sup>+</sup> T cells, suggesting that they have higher potential to increase ATP production via glycolysis under stress or other physiologically energy-demanding conditions. Anti-CD3/CD28 bead stimulated CD4<sup>+</sup> T-cell subsets showed an elevated level of glycolysis compared to unstimulated cells, which was consistent with metabolic reprogramming phenomena in response to stimulation (Fig. 3b). No significant difference between the stimulated MTC subsets regarding glycolysis, glycolytic capacity, or glycolytic reserve was observed, suggesting that regardless of subset, CD4<sup>+</sup> MTCs can achieve similar glycolytic metabolic rates.

Due to the limitations in obtaining purified subset cell numbers, we could not perform a similar set of assays on CD8<sup>+</sup> T cells. Instead, two flow cytometry-based assays were used to assess oxidative phosphorylation (OXPHOS) and fatty acid

metabolic states of both the CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Fig. 3e–h). In the first assay, cells were stained with MitoTracker Green (MTG), a mitochondrial specific dye that provides a relative assessment of mitochondrial mass, and tetramethylrhodamine methyl (TMRM), an OXPHOS marker that accumulates in functional mitochondria caused by differential membrane potential<sup>25,26</sup>. Unstimulated T<sub>CM</sub> and T<sub>EM</sub> populations had higher frequencies of cells with highly functional mitochondria (MTG<sup>Hi</sup> and TMRM<sup>Hi</sup>) compared to naïve populations in both CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Fig. 3f). Stimulated CD4<sup>+</sup> and CD8<sup>+</sup> T<sub>CM</sub> and CD8<sup>+</sup> T<sub>EM</sub> subsets also had higher frequency of cells with functional mitochondria than naïve populations (Fig. 3h). In contrast, a lower percentage of CD8<sup>+</sup> T<sub>EMRA</sub> had functional mitochondria and this number decreased when the cells were stimulated, which is consistent with previous studies showing mitochondrial impairment of this population<sup>9,27</sup>.

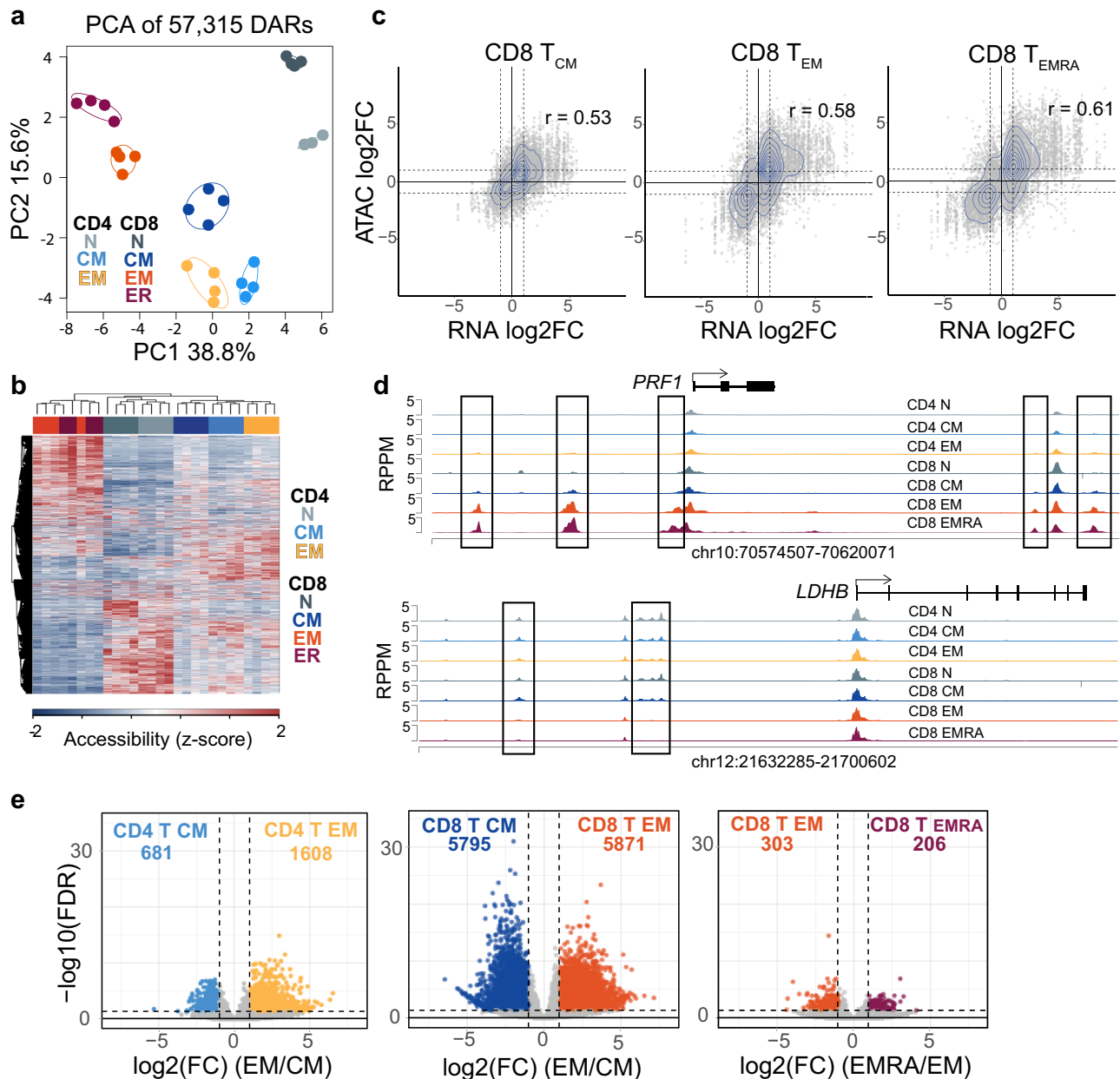
In a second assay, we performed the single cell energetic metabolism by profiling translation inhibition (SCENITH) assay that measures metabolism of individual cells based on relative translation rates and the incorporation of the puromycin into elongating ribosomes in the presence and absence of metabolic pathway inhibitors to assess glycolysis, OXPHOS, or fatty acid metabolism<sup>28</sup> (Fig. 3i–p). The SCENITH data confirmed the Seahorse results showing that unstimulated CD4<sup>+</sup> T<sub>EM</sub> had higher glycolytic capacity than the naïve CD4<sup>+</sup> T cells (Fig. 3i). In addition, all CD8<sup>+</sup> MTC showed higher glycolytic capacity than the naïve CD8<sup>+</sup> T cells. Glucose dependence; however, was significantly lower in only CD8<sup>+</sup> T<sub>EMRA</sub> (Fig. 3j). Overall, unstimulated naïve T cells exhibited more mitochondrial dependence than the MTC (Fig. 3k). Together with the mitochondria staining data, these results suggest that the unstimulated MTC subsets have elevated OXPHOS metabolism compared to naïve populations, a finding that is also consistent with the RNA-seq data; however, resting naïve T cells depended on mitochondria more than the MTC subsets. CD8<sup>+</sup> T<sub>EM</sub> and T<sub>EMRA</sub> cells exhibited an elevated fatty acid and amino acid oxidation capacity (FAO and AAO) compared to naïve CD8<sup>+</sup> T cells (Fig. 3l). In general, when these cells were stimulated they reprogrammed their metabolism to be less dependent on mitochondria and instead increased glycolytic capacity to its maximum (Fig. 3m, n). We did not detect any significant difference in the metabolisms of stimulated naïve, T<sub>CM</sub> and T<sub>EM</sub> populations of CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Fig. 3m–p). However, CD8<sup>+</sup> T<sub>EMRA</sub> cells showed less glycolytic capacity and glucose dependence, but higher mitochondrial dependence and FAO and AAO capacity than naïve CD8<sup>+</sup> T cells. Taken together these data show the distinct transcriptional differences between naïve and MTC, the resulting functional consequences to cell metabolism and phenotype, and that these differences appear to increase in both number and magnitude as cells differentiate towards an effector-memory phenotype.



**The chromatin landscape of memory subsets correlates with transcriptional differentiation.** To determine the extent to which the differences in gene expression were coordinated with changes to the chromatin organization in MTC subsets, the assay for transposase accessible chromatin-sequencing (ATAC-seq)<sup>29</sup> was performed on the MTC subsets described above. This analysis resulted in the identification of 57,315 differentially accessible regions (DAR) between naïve, T<sub>CM</sub>, T<sub>EM</sub>, and T<sub>EMRA</sub> cells of

both CD4<sup>+</sup> and CD8<sup>+</sup> T-cell lineage groups (Supplemental Data 4). As with the RNA-seq data, principal component (PC) 1 separated effector-memory subsets (T<sub>EM</sub>, T<sub>EMRA</sub>) from T<sub>CM</sub> and naïve T cells in a progressive manner (Fig. 4a). Hierarchical clustering of the DAR also revealed CD8<sup>+</sup> effector-memory and naïve T cells as being the most distinct, with central memory and all CD4 MTC subsets sharing more similarity to naïve cells, and CD8 T<sub>EM</sub>/T<sub>EMRA</sub> cells clustering closely together (Fig. 4b).

**Fig. 3 MTC subsets exhibit distinct metabolic features.** **a** Heatmap of selected genes representing differences in metabolism. Data represent the mean expression of each cell subset. **b** Line plot showing Seahorse-derived ECAR data (mpH / min /  $10^5$  cells) over time for unstimulated and stimulated naïve,  $T_{CM}$ , and  $T_{EM}$   $CD4^+$  T cells. Shaded regions indicate regions of metabolic relevance as labeled. Arrows indicate time of inhibitor addition (G, glucose; O, oligomycin; and DG: 2-deoxy-glucose). **c, d** Bar plots showing glycolytic capacity and glycolytic reserve calculated for unstimulated and stimulated  $CD4^+$  T cells, respectively. Error bars indicate  $\pm$ SD. **e** Representative flow cytometry plots showing gating strategy for unstimulated cells using MitoTracker green (MTG) and tetramethylrhodamine methyl (TMRM). **f** Bar plots showing frequency of  $MTG^{HI}TMRM^{HI}$  cells in naïve and MTC subsets. **g** Flow cytometry plots as in (e) for stimulated cells. **h** Bar plots showing frequencies as in (f). **i** Bar plots showing metabolic attributes calculated from SCENITH data as indicated for unstimulated (i-l) and stimulated (m-p) cell types. Error bars indicate  $\pm$ SD. For all experiments, 3–6 independent samples were analyzed. One-way ANOVA, with multiple comparisons were used to determine significance. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .



**Fig. 4 Effector-memory subsets exhibit a greater number of changes to chromatin accessibility compared to central memory subsets.** **a** PCA plot of DAR from ATAC-seq data. **b** Heatmap displaying DAR z-score normalized across samples. **c** Scatter plots showing log<sub>2</sub>-fold changes (FC) of MTC subsets vs. naïve T cells of DEG (x-axis) mapped to DAR (y-axis) for  $CD8^+$   $T_{CM}$ ,  $CD8^+$   $T_{EM}$ , and  $CD8^+$   $T_{EMRA}$ . Dotted lines represent  $\pm 1$  log<sub>2</sub>FC. **d** Genome plots showing the average accessibility levels in each of the memory subsets and naïve T cells at the indicated locus. Data represent the mean of each cell type. DAR are highlighted by boxes. **e** Volcano plots showing DAR that are more accessible in  $CD4^+$   $T_{EM}$  vs  $CD4^+$   $T_{CM}$ ,  $CD8^+$   $T_{EM}$  vs  $CD8^+$   $T_{CM}$ , and  $CD8^+$   $T_{EMRA}$  vs  $CD8^+$   $T_{EM}$ . Number of DAR in each direction are totaled at the top of each plot.

Differential accessibility at regions that mapped to DEG was found to be positively correlated for each of the memory subsets (Fig. 4c, Supplemental Fig. 4a-b). Comparing the magnitude of these differences by fold change of both DEG and nearby DAR showed that gene expression and chromatin accessibility increased in both quantity and intensity in effector-memory cells compared to smaller differences from naïve T cells in central MTC (Fig. 4c). Some examples of DAR near DEG include the *PRF1* locus (Fig. 4d), which was found to have gained accessibility only in effector-memory subsets. Alternatively, chromatin at two regions upstream of the *LDHB* gene, which is downregulated in these subsets (Fig. 2j), was found to have decreased in accessibility in the effector-memory T cells (Fig. 4d). In a similar fashion, accessibility around gene loci expressed more highly in naïve or  $T_{CM}$  cells, such as genes encoding *LEF1* and the G protein coupled receptor *GPR15* (involved in T-cell homing), showed lower chromatin accessibility in  $T_{EM}$  subsets (Supplemental Fig. 4c and d). Overall  $CD4^+$  MTC exhibited relatively fewer DAR between  $T_{CM}$  and  $T_{EM}$  subsets compared to the large number of accessibility changes between  $CD8^+$  MTC subsets (Fig. 4e), and chromatin accessibility differences between  $CD8^+$   $T_{EM}$  and terminally differentiated  $T_{EMRA}$  were the fewest in number (Fig. 4e). Collectively these data demonstrate that the transcriptional differentiation found in memory subsets is highly correlated with changes to chromatin accessibility.

### Differentiated chromatin between MTC subsets is enriched for bZIP, HMG, T-box, and bHLH transcription factor motifs.

Transcription factors play key roles in orchestrating the global gene expression changes involved in establishing and maintaining cell type differentiation and establishing distinct gene regulatory states<sup>30</sup>. To identify potential transcription factors regulating differentiation of MTC subsets, we analyzed the variation of chromatin accessibility between subsets at sites with known binding motifs for transcription factors in the genome using chromVAR<sup>31</sup>. ChromVAR identifies motifs around which chromatin accessibility varies the most in a given set of samples (in this case cell subsets). We found that the chromatin accessibility around binding motifs for AP-1 (Jun, FOS, BATF), T-box (T-BET, EOMES), and HMG family transcription factors (*LEF1*) was the most highly variable across all MTC and naïve T cells (Fig. 5a). Visualization of the ChromVAR analysis data using tSNE projection distinguished samples by MTC subtype (Fig. 5b). Overlaying the ChromVar deviation score for individual transcription factor motifs onto the tSNE projection showed that  $T_{EM}$  cells had higher scores for T-box factors such as T-BET, encoded by *TBX21* (Fig. 5c), while naïve T cells and  $T_{CM}$  samples showed higher scores for *LEF1* binding sites (Fig. 5d). Higher ChromVar scores were also found at AP-1 and IRF family binding motifs in the effector-memory samples (Supplemental Figs. 5a-f).

To further interrogate the various transcription factor regulatory networks involved in differentiating the MTC subsets the PageRank algorithm<sup>32</sup> was used to correlate the presence of transcription factor binding motifs in accessible regions with changes in target gene expression, ultimately calculating a rank of transcription factor importance to the network. PageRank identified several transcription factors that increased in network rank in at least one MTC subset relative to naïve T cells and recapitulated naïve-associated transcription factors such as *LEF1* and *TCF7* (Fig. 5e). To compare subset differences, the fold change in PageRank scores from naïve cells was plotted against RNA expression for each factor in  $CD8^+$   $T_{CM}$  and  $T_{EM}$  cells (Fig. 5f-g, Supplemental Fig. 5g-i). These data also corroborated the ChromVAR analysis by indicating both higher relative expression and PageRank scores for AP-1 and T-box factors,

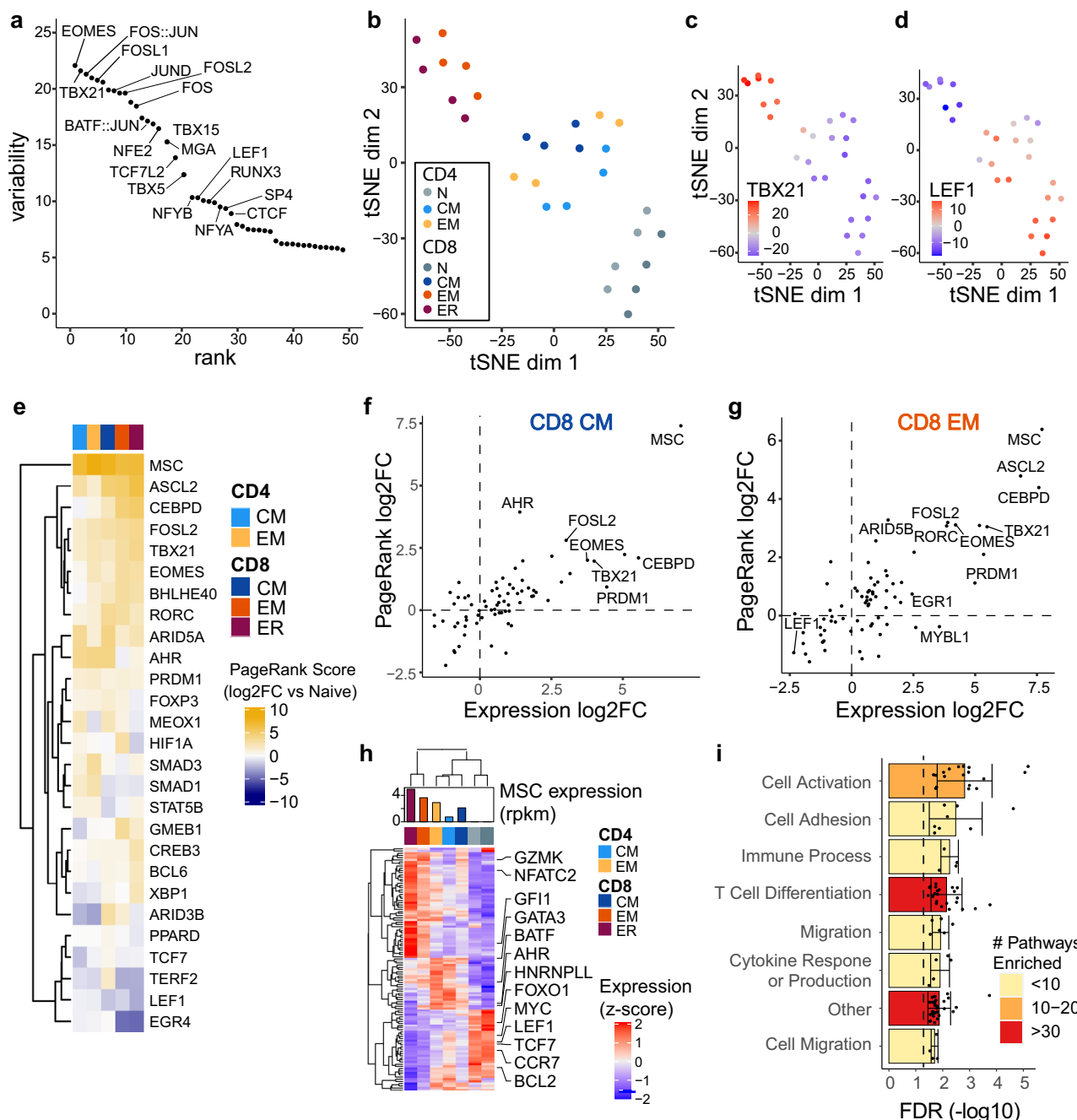
such as *FOSL2*, T-BET, and *EOMES* in the effector-memory subsets.

One transcription factor, *MSC* (encoding for musclin or *ABF-1*), was observed to have the highest rank score in all MTC subsets while being absent in naïve T cells. *MSC* has been previously identified as a transcriptional repressor capable of binding to E-box elements and implicated in the activation pathways of B cells, as well as the differentiation of peripheral  $CD4^+$  Treg cells<sup>33,34</sup>. Intracellular staining of *MSC* showed expression in all subsets with significant increases in stimulated T-cell subsets (Supplemental Fig. 6a), suggesting a potential role in T-cell activation. Analysis of the PageRank-generated regulatory network for *MSC* identified its potential for regulating many important genes in MTC subsets, including repression of *LEF1* and *TCF7* (Fig. 5h). The *MSC* target gene network was enriched for several key pathways such as cell activation, adhesion, cytokine production, and differentiation (Fig. 5i).

### bHLH Family Factors AHR and HIF1A potentially regulate different environmental responses in $CD8^+$ $T_{CM}$ and $T_{EM}$ MTC subsets.

Reconstructing transcription factor regulatory networks via PageRank-predicted interactions emphasized the centrality of key transcription factors *LEF1*, *TCF7*, T-BET, and *EOMES*, as well as differences in TF importance by subset. *MSC* was predicted by PageRank to regulate *LEF1* and *TCF7*, potentially leading to higher repression of these factors in  $CD4^+$   $T_{EM}$  and other effector-memory subsets (Fig. 6a). The bHLH transcription factor *AHR* was highly induced as seen by intracellular protein staining in all T-cell subsets (Supplemental Fig. 6a). Intriguingly, *AHR* also showed both high relative expression and PageRank score exclusively in the  $CD8^+$   $T_{CM}$  subset, while bHLH transcription factor, *HIF1A* (hypoxia-inducible factor 1 $\alpha$ ), was more highly ranked by PageRank in  $CD8^+$   $T_{EM}$  cells (Fig. 6b, Supplemental Fig. 5k). *AHR* protein levels were increased in all subsets upon stimulation. *HIF1A* was increased after stimulation in all T-cell subsets except for  $T_{EM}$ , which maintained more constant concentrations of this protein (Supplemental Fig. 6a). *AHR* has been implicated in sensing xenobiotics and T-cell homing to tissues<sup>35</sup>, while *HIF1A* has previously been associated with sensing hypoxic conditions and modulating metabolism in these circumstances<sup>36</sup>. *HIF1A* and *AHR* are known to have an antagonistic relationships with respect to target genes and their own expression<sup>37,38</sup>. GSEA using the list of genes putatively regulated by *AHR* showed significant enrichment of these genes in those upregulated in  $CD8^+$   $T_{CM}$  vs  $CD8^+$   $T_{EM}$  (Fig. 6c). Expression of several of these genes in the leading edge of GSEA enrichment were positively correlated with *AHR* expression in MTC and were exclusively expressed in  $T_{CM}$  (Fig. 6d). Two of these genes, *VBPI* and *PVT1*, along with *HIF1A* are part of the hypoxia response<sup>39,40</sup>. Two DAR with greater accessibility in  $T_{CM}$ , were found near the promoter region of the *VBPI* gene, one of which contains a binding motif known to bind *AHR* (Fig. 6e). *VBPI* encodes for the protein VHL which is the substrate recognition subunit of an E3 ligase known to target *HIF1A* for degradation<sup>39</sup>. Another DAR with higher accessibility in  $T_{CM}$  and naïve T cells was found at the promoter of the *INPP4B* gene, encoding inositol polyphosphate-4-phosphatase type II B (Fig. 6e). This region contains three *AHR* binding motifs as well as a motif specific for *HIF1A* binding, suggesting potentially competitive regulation between these two factors. In total, these data suggest that bZIP, HMG, and T-box family transcription factors are important for memory-subset differentiation, and additionally bHLH family factors, such as *MSC*, *AHR*, and *HIF1A* may play key roles in regulating distinct subsets and their transcriptional programs before and following stimulation.

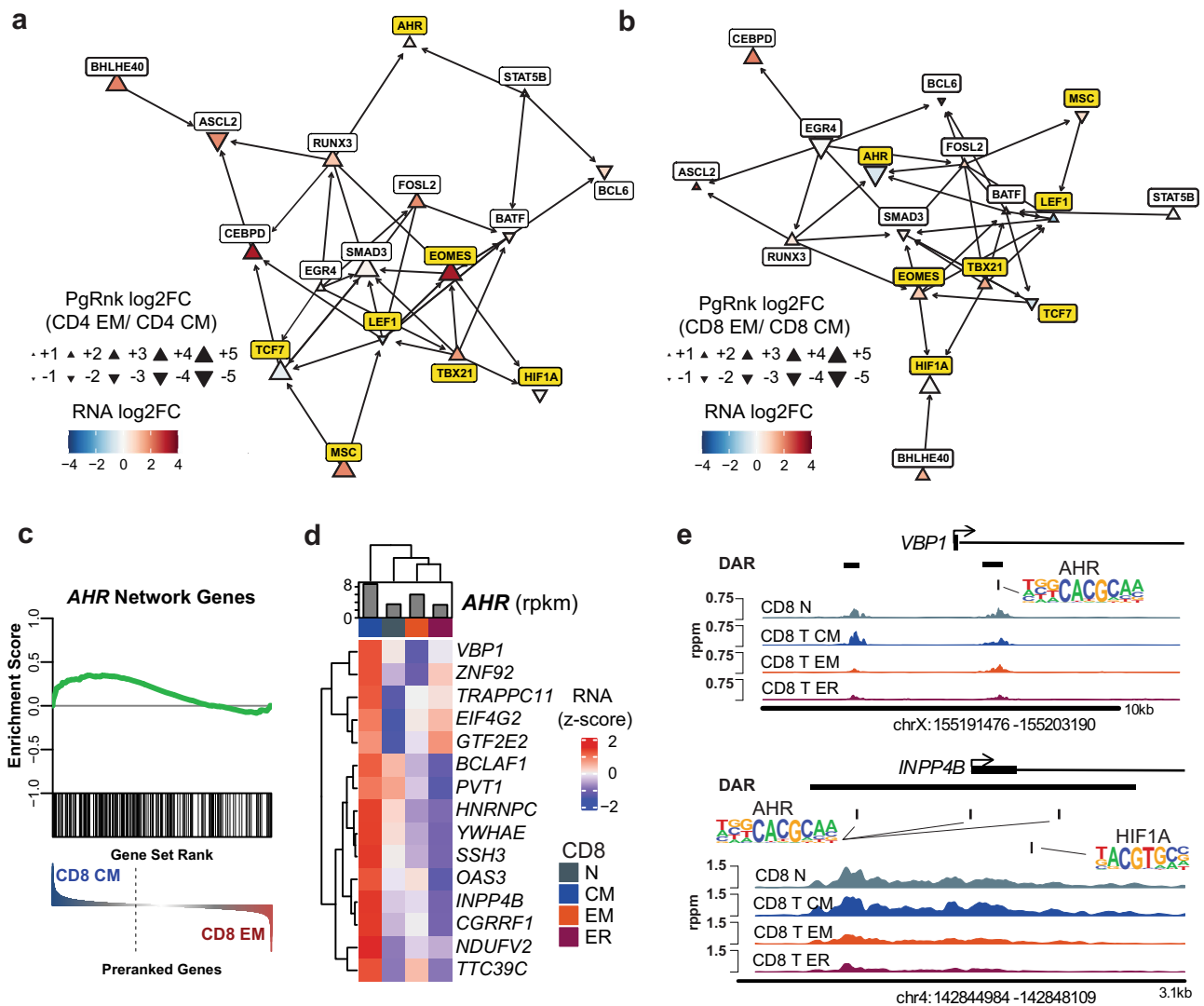




**Fig. 5** bZIP, HMG, T-box, and bHLH factors differentiate MTC subsets by chromatin accessibility. **a** Plot showing the top 50 transcription factors ranked (x-axis) according to variability (ChromVar) in chromatin accessibility (y-axis) around their corresponding binding motifs. **b** tSNE plot showing reduced dimensional visualization of variability of chromatin accessibility data for each memory-subset sample at all differential sites. **c, d** tSNE plot from **b** colored by ChromVar variability around *TBX21* (T-BET) and *LEF1* binding motifs, respectively. **e** Heatmap of the  $\log_2$ -transformed fold change in PageRank statistics for each MTC subset (compared to naive). **f, g** Scatter plots showing  $\log_2$ FC of RNA expression (x-axis) and PageRank statistic  $\log_2$ FC (y-axis) between memory and naive samples for labeled CD8<sup>+</sup> MTC subsets. **h** Heatmap of RNA expression from gene network of MSC transcription factor with bar plot showing expression of MSC in each subset. Data represent the mean expression of each cell subset. **i** Bar plot showing average gene set enrichment FDR corrected *p*-value from MSC network genes binned by T-cell-relevant categories. Error bars represent one standard deviation. Color corresponds to the number of pathways significantly enriched in each bin.

**Genes that are uniquely upregulated in stimulated memory-subset cells include Induced and Augmented transcripts.** To better understand the transcriptional and epigenetic properties that allow memory T-cell subsets to rapidly respond to stimulation, all five MTC subsets, as well as naive T cells were stimulated using anti-CD3/CD28 beads *ex vivo* for 42–48 h. RNA-seq analysis of the stimulated T cells showed a profound change in

transcriptional programs after stimulation and much of this response was homogeneous across the different MTC subsets. This is highlighted by PCA wherein PC1 (52.59% variation) separated resting cells from all those that have been activated (Fig. 7a). GSEA showed that activated naive cells more strongly upregulated pathways such as MYC target genes to a greater extent than in MTC, while activated MTC more strongly

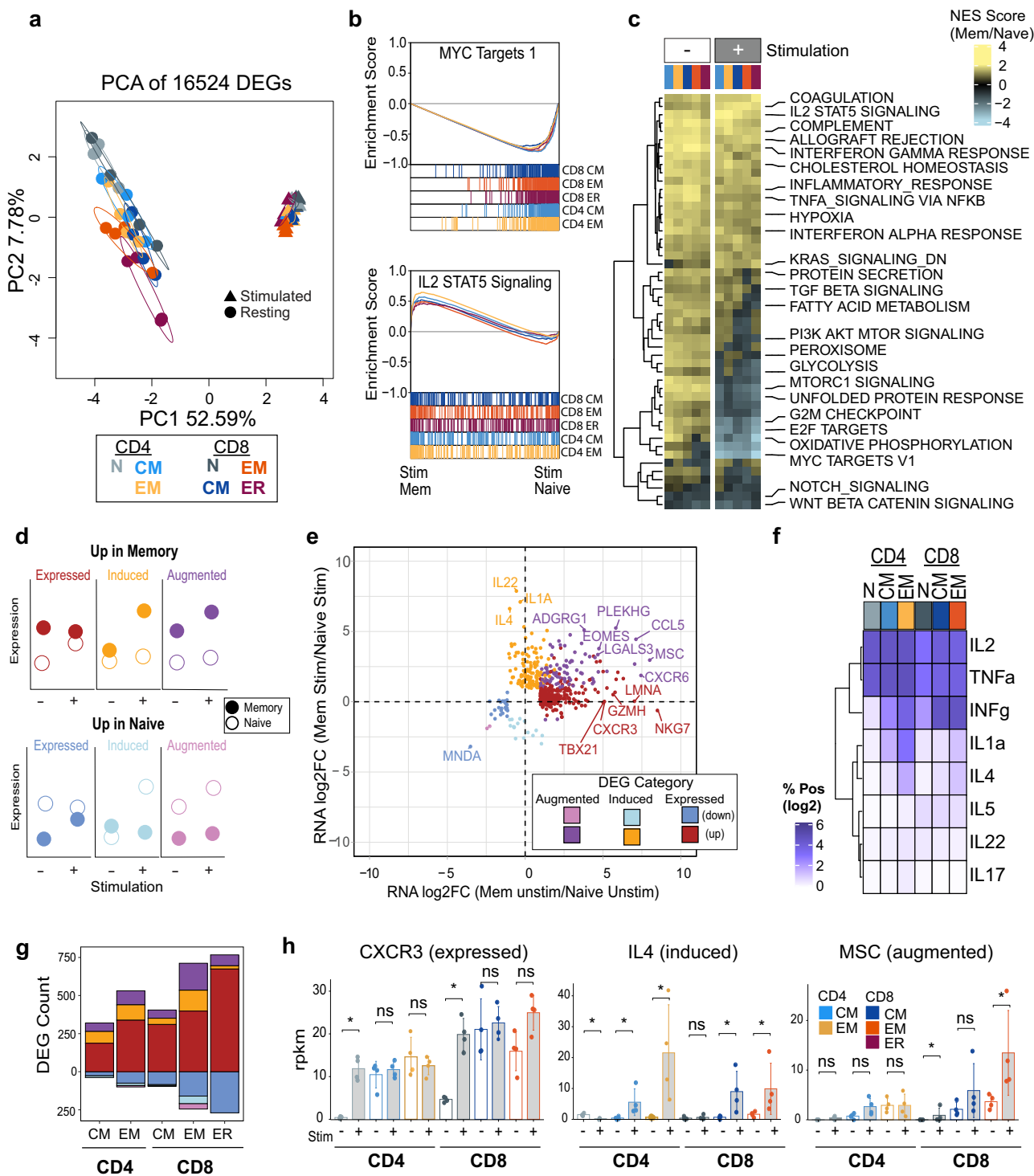


**Fig. 6 bHLH transcription factors AHR and HIF1A distinguish CD8<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub>.** **a** Network diagram showing transcription factors in CD4<sup>+</sup> MTC highly ranked by PageRank and associated connections to other TF. Triangles represent log<sub>2</sub>FC (size) and direction (orientation) of PageRank as indicated in the legend. Arrows represent putative regulation via nearby binding motifs. AHR, HIF1A, MSC, and other memory-subset defining TFs are highlighted in yellow. **b** Network diagram as in (a) for CD8<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub>. **c** GSEA plot using gene set made up of genes putatively regulated by AHR showing enrichment in CD8<sup>+</sup> T<sub>CM</sub> vs. T<sub>EM</sub> DEG. **d** Heatmap of leading-edge genes from GSEA most highly upregulated in T<sub>CM</sub>. Bar plot indicates normalized expression of AHR in rpkm. Data represent the mean expression of each cell subset. **e** Genome plots showing the average accessibility levels in each of the CD8<sup>+</sup> memory and naïve T cells at the indicated locus. DAR are highlighted by black horizontal bars, with AHR and HIF1A binding motifs at indicated locations. Data represent the mean of each cell type.

upregulated genes associated with IL-2, STAT5 signaling, and other immune response signaling pathways (Fig. 7b). Despite a large degree of consistency in response to stimulation in the different memory subsets, gene sets related to fatty acid metabolism, glycolysis, and MTOR signaling were enriched to varying degrees in GSEA between T<sub>CM</sub> and T<sub>EM</sub> cell subsets after stimulation (Fig. 7c).

To identify genes that were differentially expressed upon stimulation in MTC subsets, expression patterns were examined that were specific to MTC. Overall, the data revealed the presence of three main groups of genes that change in expression compared to naïve T cells before and after stimulation as hypothetically illustrated in Fig. 7d. Each group has two expression states: (a) its constitutive expression in unstimulated MTC vs. naïve; and (b) its change in expression following stimulation. Expressed genes had high fold-change differential expression between naïve and MTC, which naïve T cells induced

only after stimulation. Induced genes were similarly expressed in both resting naïve and MTC and are induced in only the MTC group following stimulation. Augmented genes showed expression differences between MTC and naïve cells in both the resting and stimulated states (Fig. 7d). Overlaying these categories on gene expression differences of naïve versus memory resting and stimulated T cells showed that the majority of such DEG were positively upregulated in MTC (Fig. 7e). The memory Induced gene category was highly enriched for cytokine genes including *IL4* (Fig. 7h), *IL1A*, and *IL22*. These genes were uniquely expressed by MTC subsets to varying degrees after stimulation but repressed in activated naïve T cells at 48 h after the same stimulation. To confirm these results, we measured a panel of common cytokines using intracellular staining coupled with flow cytometry (Supplemental Fig. 7). Some cytokines, such as IL-2 and TNF $\alpha$  were expressed in a significant number of cells by all activated T cells studied. In contrast, the production of effector



**Fig. 7 MTC induce or augment unique gene expression programs after stimulation.** **a** PC analysis plot of DEG for resting (circle) and stimulated (triangle) memory and naïve T-cell samples. **b** GSEA plots for indicated gene sets showing the enrichment score for each stimulated MTC subset as compared to stimulated naïve T cells. **c** Heatmap of normalized enrichment scores (NES) from GSEA of both stimulated and unstimulated MTC compared to respective naïve T cells. **d** Schematic describing examples of DEG categorization; before and after stimulation (Stim). **e** Scatter plot showing log<sub>2</sub>FC of DEG in unstimulated vs. stimulated MTC. Dots are colored according to DEG category as indicated. **f** Heatmap showing log<sub>2</sub> transformation of the percent of cells positive for respective cytokine protein expression. **g** Bar plot showing number of DEG compared to naïve T cells belonging to Expressed, Induced, or Augmented categories for each memory subset. **h** Gene expression bar plots showing rpkm for indicated genes. Error bars represent ±1 SD from mean in each group. Asterisks indicate relevant significant differences as detected by DESeq2 algorithm.

T-cell-associated cytokines (IFN $\gamma$ , IL-1 $\alpha$ , IL-4, IL-17, and IL-5) was differentially increased in various MTC after stimulation compared to stimulated naïve T cells (Fig. 7f). IL-22 protein was significantly induced in naïve and T<sub>EM</sub> CD4<sup>+</sup> MTC. The T<sub>EM</sub> MTC subsets of both CD4<sup>+</sup> and CD8<sup>+</sup> lineage produced the highest level of cytokines after stimulation in comparison to naïve or T<sub>CM</sub> cells (Supplemental Fig. 7). The stimulation-independent Expressed gene category contained the most DEG overall, but each memory subset also exhibited Induced and Augmented genes (Fig. 7g). Genes found within the Expressed category for T<sub>EM</sub> cells included important effector T-cell molecules, such as granzyme H, encoded by *GZMH*; T-BET; and *CXCR3*, a chemokine receptor associated with T<sub>HI</sub> CD4<sup>+</sup> T cells and effector CD8<sup>+</sup> T cells<sup>41</sup> (Fig. 7h). Both CD8<sup>+</sup> and CD4<sup>+</sup> T<sub>EM</sub> cells showed higher numbers of Induced and Augmented genes than T<sub>CM</sub>. Intriguingly, *MSC* was also found to exhibit an Augmented expression pattern in these cells after stimulation, while remaining repressed in stimulated naïve T cells (Fig. 7h). Thus, despite both MTC and naïve T cells activating several similar gene programs upon TCR stimulation, MTC subsets have unique induction/expression profiles upon stimulation, a property that may provide them with a greater degree of efficiency and capacity upon rechallenge.

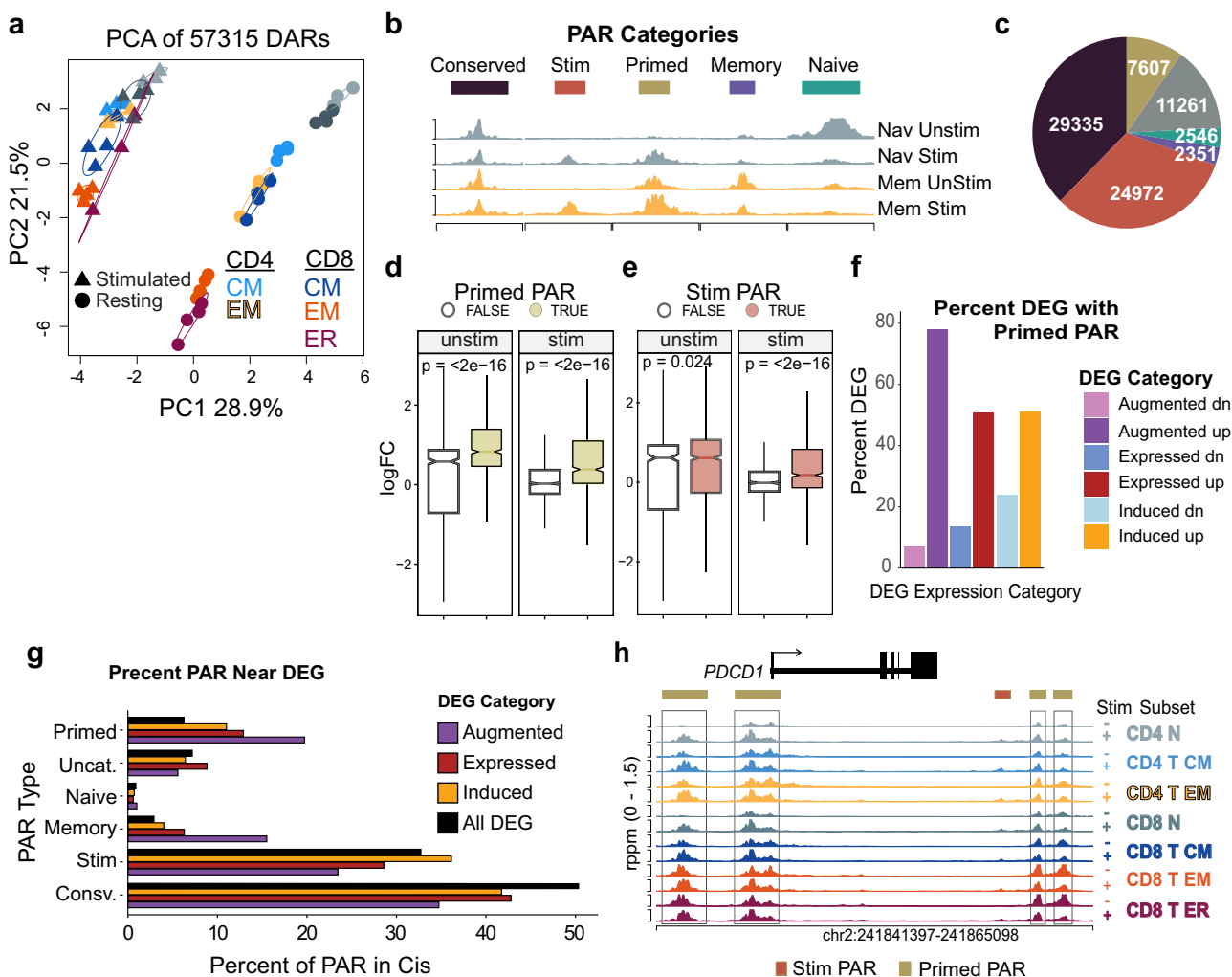
**Augmented gene expression is correlated with epigenetic changes introduced by earlier activation of naïve cells.** To determine the role of chromatin accessibility in the unique response to stimulation described above, ATAC-seq was performed on the stimulated memory and naïve T cells. Similar to the global transcriptional response to stimulation, PCA of the ATAC-seq datasets showed that the greatest amount of variation separated stimulated from unstimulated cells, and that differences between memory subsets were diminished in the stimulated samples (Fig. 8a). Individual peaks of accessibility in the ATAC-seq analysis followed a set of patterns that we have termed patterned accessibility regions (PAR) (Supplemental Data 5). PAR are defined by chromatin state before and after stimulation. Five categories of PAR emerged from the analysis: conserved, stimulated, primed, memory, and naïve (Fig. 8b). Conserved-PAR were unchanged across all subsets and stimulation states, whereas stimulation-PAR were present only after stimulation. Each of these two PAR groups occurred equally in both naïve and MTC and made up the majority of the accessible chromatin regions found in the ATAC-seq data in all memory subsets (Fig. 8c, Supplemental Fig. 8a). Conserved-PAR were also found to be more highly enriched in promoter regions than other PAR categories (Supplemental Fig. 8b). PAR specific to MTC occurred in one of three ways: memory-PAR were specific to MTC and were unchanged by stimulation; naïve-PAR were present in resting naïve T cells, but neither their stimulated counterparts nor in MTC; and finally primed-PAR were accessible in resting and stimulated MTC and became accessible upon activation in naïve T cells. Primed-PAR were more than twice as abundant as memory- or naïve-PAR.

In examining the relationships between PAR types and gene expression, we observed that the presence of at least one primed-PAR was strongly correlated with upregulation of DEG in MTC for both resting and stimulated states (Fig. 8d). A similar influence of stimulated-PAR was found in the stimulated but not the resting MTC transcriptional data (Fig. 8e). The presence of memory-PAR was found to be correlated with both resting and stimulated expression in MTC, while the presence of naïve-PAR biased gene expression in most nearby genes towards higher expression in naïve T cells (Supplemental Figs. 9a, b). Comparing these nearby DEG to the MTC-specific gene programs as defined

by unique expression after stimulation (Expressed, Augmented, or Induced) revealed that nearly 80% of the upregulated genes in the augmented category had at least one primed peak in *cis* (Fig. 8f). Additionally, compared to the genome-wide prevalence, primed- and memory-PAR made up a significantly higher proportion of all the accessible regions surrounding DEG that were augmented in MTC (Fig. 8g). This is exemplified by the *PDCDI* locus, a gene showing augmented expression in MTC, which is also composed primarily with primed-PAR (Fig. 8h). These data support an epigenetic mechanism of enhanced recall in which primed-PAR are inscribed in the epigenome upon previous activation of naïve T cells<sup>42</sup> and occur within regions of the genome that are associated with higher expression levels of important genes in MTC.

**PAR contain distinct sets of transcription factor motifs that segregate T cells by lineage and memory subtype.** To identify transcription factors associated with each of the identified PAR categories, motif discovery, and enrichment analysis was performed using HOMER<sup>43</sup>. To control for uneven power, driven by unequal numbers of individual peaks in each category (Fig. 8c), enriched transcription factor binding motifs in DAR were compared across all five peak categories by relative ranking of enrichment *p*-values (Fig. 9a). Enrichments of HMG family transcription factors (LEF1 and TCF7) were more highly ranked within naïve-PAR, and CTCF-motif enrichment was ranked highest in the constitutively accessible conserved-PAR (Fig. 9b). Interestingly, several transcription factor motifs were enriched in both primed- and stimulation-PAR, including the AP-1 family factors (Fig. 9b). Motifs specific for other transcription factors were enriched in both primed- and memory-PAR, including the T-box factors EOMES and T-BET.

Enrichment rank scores of transcription factor motifs associated with primed-PAR across the different memory subsets in CD8<sup>+</sup> and CD4<sup>+</sup> T cells clustered according to cell lineage (Fig. 9c); whereas transcription factor motifs enriched in memory-PAR clustered according to memory cell subtype (Fig. 9d). Differences between memory subsets were also observed, with CD4<sup>+</sup> T cells showing slightly higher percentages of primed-PAR containing AP-1 factor motifs (Fig. 9e). The percentage of memory-PAR containing T-BET motifs was highest in CD8<sup>+</sup> T<sub>EM</sub> cells and lowest in CD4<sup>+</sup> T<sub>CM</sub>, a finding that is consistent with the role of T-box factors in effector T-cell programming (Fig. 9f)<sup>44</sup>. The PageRank algorithm was used to compare the relative importance of different transcription factors in terms of their effect on target DEG after stimulation. EOMES was found to have the highest PageRank score in stimulated CD8<sup>+</sup> MTC (compared to stimulated naïve), as well as the greatest fold change in augmented gene expression (Fig. 9g, i), a finding consistent with its role in maintaining CD8<sup>+</sup> T-cell memory<sup>44,45</sup>. PageRank scores for EOMES and T-BET were similar. T-BET expression was lowest in naïve cells but varied across MTC subsets and was induced to similar levels irrespective of subset (Fig. 9h). Protein measurements of both intracellular EOMES and T-BET via flow cytometry showed that each were increased after stimulation (Supplemental Fig. 6a); however, EOMES levels were already significantly high in resting CD8 T<sub>EM</sub> and remained unchanged at these high levels after stimulation. T-BET levels were induced to the highest degree in naïve T cells in both CD4<sup>+</sup> and CD8<sup>+</sup> lineage groups. Likewise, the PageRank score and expression of *IRF8* were highest in stimulated naïve T cells (Fig. 9g). Expression data showed that stimulation induced high levels of *IRF8* in naïve cells, but all memory subsets induced significantly lower levels of *IRF8* after the 48 h of stimulation (Fig. 9h) compared to naïve T cells. Collectively these data show



**Fig. 8 Accessible chromatin regions primed from previous activation are associated with augmented gene expression of MTC.** **a** PC analysis plot of DAR between unstimulated and stimulated memory and naïve T-cell samples. **b** Schematic illustrating examples of PAR categorized by accessibility in both unstimulated (Unstim) and stimulated (Stim) states in memory (Mem) and naïve (Nav) T cells. **c** Pie chart showing number of DAR peaks in each category. **d** Box plot showing  $\log_2FC$  of DEG found to be significant in either unstimulated memory vs unstimulated naïve T cells or stimulated memory vs stimulated naïve T cells. Boxes colored in gold represent  $\log_2FC$  distribution of genes with at least one primed-PAR mapped to a differential gene locus. Notches represent median values, box limits represent upper and lower quartiles. Lines represent range. **e** Box plot as in (d). Boxes colored in red represent  $\log_2FC$  distribution of genes with at least one stimulation-primed-PAR mapped to a differential gene locus. **f** Bar plot showing the percent of genes belonging to either expressed, induced, or augmented expression categories with at least one primed-PAR nearby. **g** Bar plot showing the proportion of each PAR type compared to all peaks surrounding DEG belonging to each of the different memory-specific expression categories. The *All DEG* category represents all genes differentially expressed compared to naïve T cells. **h** Genome plot showing accessibility data around the *PDCD1* locus. DAR are indicated by horizontal bars colored according to PAR category above tracks. Primed-PAR are highlighted by boxes.

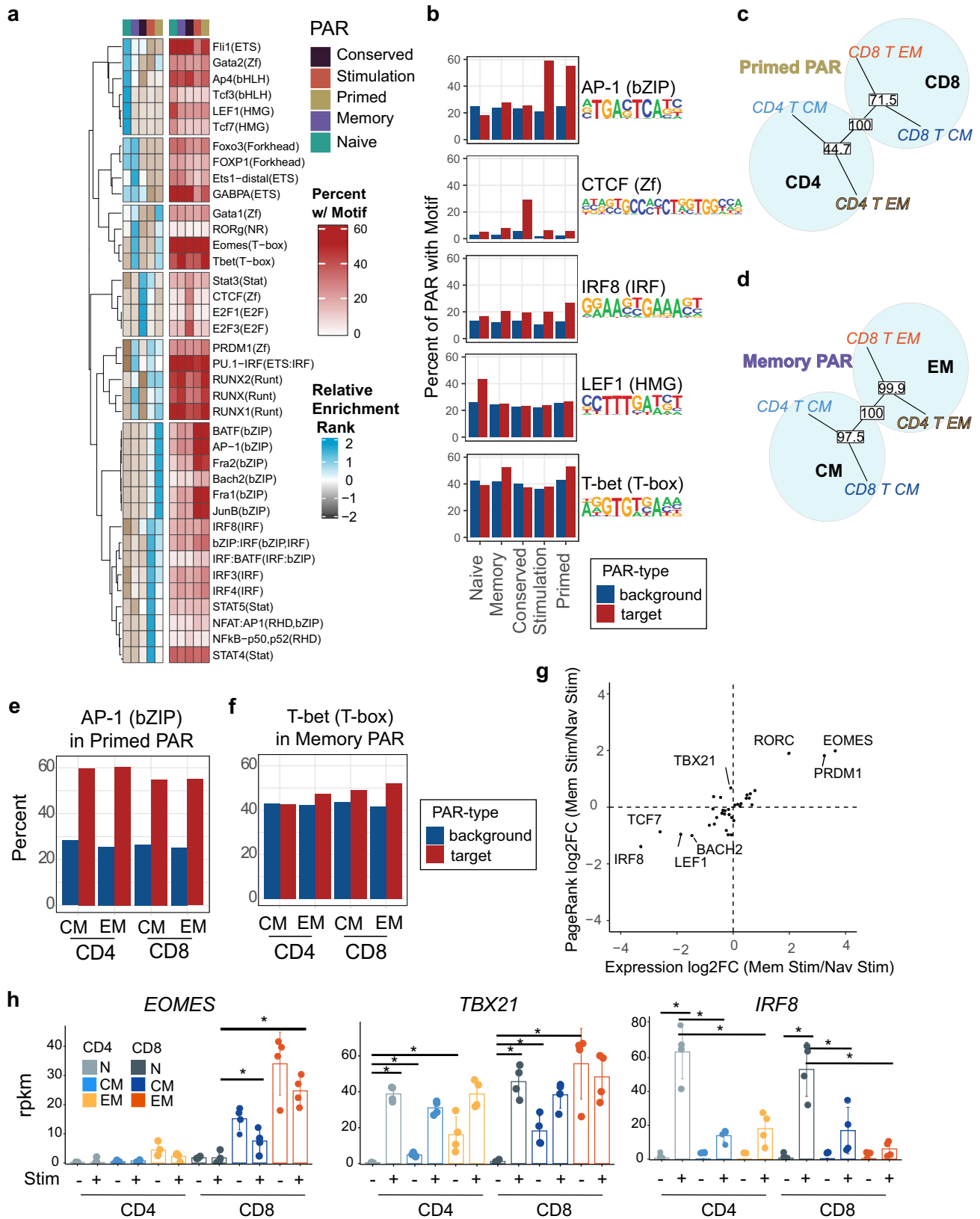
that both distinct and overlapping sets of transcription factor binding motifs are enriched in each of the identified PAR categories. Moreover, the expression of many of the transcription factors known to bind such motifs (IRF8, EOMES, T-BET, and LEF1) are coordinated to regulate and enable unique aspects of MTC differentiation following stimulation.

**Discussion**

In this study we used an integrated transcriptomic and epigenetic sequencing approach to understand the cellular programming of the largest memory-subset groups found in human blood and how these programs change in response to ex vivo stimulation. In addition to specific and differential relationships between subsets and lineages, a series of regulatory modalities, pathways, and transcription factors were found to be associated with specific MTC genes. States of chromatin accessibility in key genes within

MTC revealed a robust mechanism to control gene expression in response to secondary activation that allow MTC to respond more efficiently and with fewer epigenetic reprogramming steps. Regardless of cell lineage, MTC subsets shared nearly one third of the DEGs that distinguish them from naïve T cells, suggesting a common memory signature. However, expression differences between each MTC subset and naïve T cells were also observed in gene pathways related to important functions such as cytotoxicity, metabolism, and self-renewal.

It is known that MTC exploit distinct metabolic pathways based on their differentiation and memory status<sup>46–48</sup>. Here, in addition to the transcriptional evidence, we provided additional characterization of the metabolic states of naïve and MTC subsets. Consistent with previous studies<sup>49</sup>, we showed that resting MTC populations predominantly utilize OXPHOS and fatty acid oxidation for their main or primary metabolic needs and switch



their metabolism to aerobic glycolysis when they are stimulated ex vivo. In addition, some of the MTC subsets use the other pathways to different degrees as a secondary means of metabolism. Previous studies revealed that although both resting naïve and MTC rely on OXPHOS, naïve T cells harbor less mitochondrial mass<sup>50</sup>. In addition, mitochondria of MTC display

more elongated structures with tight cristae compared to naïve counterparts, indicating that OXPHOS is highly efficient in these cells<sup>51,52</sup>. Supporting these studies, our mitochondria analysis of T-cell subsets demonstrates distinct distribution of populations with MTG<sup>Hi</sup> and TMRM<sup>Hi</sup> among the subsets according to their activation state. In particular, T<sub>EM</sub> and T<sub>CM</sub> subsets exhibit

**Fig. 9 Transcription factor motif and expression patterns differentiate between memory and naïve T-cell response to stimulation.** **a** Heatmaps of enrichment rankings for transcription factor binding motifs in each of the PAR categories for CD8<sup>+</sup> T<sub>EM</sub> cells and percent of PAR with the target motif for the selected example motifs. **b** Bar plots showing percent of input peaks containing indicated motifs (red) and percent background regions containing motif (blue) for five motifs representing enrichment specific to one or more of the PAR categories. **c, d** Phylogenetic dendrogram of transcription factor binding motif enrichment ranks found within Primed (**c**) and Memory (**d**) PAR compared across MTC subset groups. The reproducibility of tree structures was tested using bootstrapping analysis with the percent reproducibility shown in white boxes for each node. **e, f** Bar plots showing examples of differences in percent input PAR containing the indicated motifs across MTC subsets for primed (**e**) and memory (**f**) PAR. **g** Scatter plot showing log<sub>2</sub>FC of RNA expression (x-axis) vs PageRank statistic (y-axis) of CD8<sup>+</sup> T<sub>EM</sub> compared to naïve T cells for transcription factors known to bind to motifs found to be enriched in primed or memory-PAR. **h** Bar plot showing rpkms values for indicated genes in both resting and stimulated samples. Error bars indicate ±SD. Asterisks indicate relevant significant differences as detected by DESeq2 algorithm.

significantly higher frequency of cells with greater mitochondria mass and potential than naïve T cells. By contrast, CD8<sup>+</sup> T<sub>EMRA</sub> cells did not show a significant difference compared to their naïve counterpart, and stimulation of these cells further decreased the frequency of MTG<sup>Hi</sup> and TMRM<sup>Hi</sup>, suggesting that these cells are metabolically distinct than the other MTC subsets. CD4<sup>+</sup> MTC subsets have a similar fatty acid oxidation capacity; however, CD8<sup>+</sup> T<sub>EM</sub> and T<sub>EMRA</sub> subsets show increased capacity for this metabolic pathway compared to naïve counterparts. Consistent with previous studies<sup>53</sup>, all activated T cells decreased fatty acid utilization.

Recent studies showed that MTC shift towards aerobic glycolysis more rapidly than naïve T cells to facilitate a rapid secondary response<sup>54</sup>. We showed that T<sub>EM</sub> and T<sub>CM</sub> subsets differ in expression of genes related to the regulation of metabolism and in particular to acetyl CoA production with differential expression of glycolysis-promoting enzymes such as LDH. Our metabolic data confirm that all resting MTC subsets have higher glycolytic capacity than their naïve counterparts, suggesting a potential contribution to more efficient secondary responses. Moreover, the glycolytic capacity of the effector MTC is significantly higher than central memory and naïve subsets that might be required for their rapid effector function.

Overall, our data show evidence of a gradient of differentiation from naïve T-cell progenitors in MTC subsets which is highest in T<sub>EM</sub>. We show that this subset exhibits the greatest numbers of changes from the naïve T-cell transcriptome as well as higher numbers of changes to chromatin accessibility. T<sub>CM</sub>, on the other hand, maintained the highest expression levels of genes that were also expressed in naïve T cells, a result consistent with other recent findings<sup>55</sup>. In addition shared expression of genes in naïve T cells and T<sub>CM</sub> MTC includes the transcription factors downstream of the WNT-beta catenin signaling pathway (LEF1 and TCF1), which have been previously associated with self-renewal and stem-like properties in mature CD8<sup>+</sup> T cells<sup>56</sup>. It is important to note that the naïve CCR7 + CD45RA + T cells isolated in this study may also contain a small percent of stem-cell MTC: however, this population has been shown to make up less than 10% of total CD8<sup>+</sup> T cells and so likely had a small effect on our data<sup>57</sup>.

One mechanism by which MTC may be able to alter the way in which they respond to antigen stimulation compared to naïve T cells is through modification of chromatin structure. Large numbers of chromatin accessibility and epigenetic modifications such as DNA methylation changes have been observed previously in mouse CD8<sup>+</sup> MTC after acute or chronic viral infections<sup>13,58</sup>. Here, we found that all human CD8<sup>+</sup> and CD4<sup>+</sup> MTC contain large numbers of DAR compared to naïve T cells. These DAR fell within loci that were correlated with differential gene expression, and which progressively increased in both number and magnitude from central memory to effector MTC subsets. It remains unclear whether this variation in epigenetic changes between central and effector MTC reflects a one-way linear differentiation

path in CD8<sup>+</sup> T cells as previously suggested by some models<sup>59</sup>, or results from a more complicated plastic development trajectory<sup>58,60</sup>. However, fewer differences were found when comparing the chromatin of CD8<sup>+</sup> T<sub>EM</sub> to T<sub>EMRA</sub>, suggesting a closely shared differentiation pathway for these cells. Interestingly, accessibility differences between circulating CD4<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub> were subtle, reflecting their close relationship. This may reflect a heterogeneous population not as well captured by traditional T<sub>CM</sub> and T<sub>EM</sub> definitions, as the diverse roles CD4<sup>+</sup> effector T cells play in the immune response is likely to lead to a memory population that is equally as diverse. Upon activation, however, CD4<sup>+</sup> T<sub>EM</sub> show a greater degree of transcription and chromatin accessibility changes (greater than CD4<sup>+</sup> T<sub>CM</sub>), which may be related to the activation requirement of effector functions of different CD4<sup>+</sup> T-cell effector subsets.

Transcription factor motifs associated with DAR and DEG identified sets of potential factors that could be responsible for defining MTC transcriptional networks. T-box, AP-1, and HMG family transcription factor binding motifs were enriched within DAR that separate memory from naïve T cells, particularly those that distinguish effector memory. Supporting these results, the T-box factors T-BET and EOMES are known to be important in the formation of CD8<sup>+</sup> effector T-cell programming and memory cell populations respectively<sup>44,61</sup>. The HMG factor LEF1 has also been implicated in early T-cell differentiation<sup>62</sup>. Members of the bHLH family of transcription factors, such as AHR, HIF1A, and MSC were found to be important for distinguishing between the memory subsets. AHR and HIF1A are capable of sensing changes in environmental conditions<sup>38,63</sup>. An AHR binding motif found within a DAR near the promoter of *VBPI* (encoding VHL) suggests one potential mechanism by which AHR modulates metabolic differences of CD8<sup>+</sup> T<sub>CM</sub> and T<sub>EM</sub>, as conditional deletion of this gene has been shown to promote constitutive glycolysis in CD8<sup>+</sup> T cells reminiscent of that exhibited by T<sub>EM</sub><sup>64</sup>. Our data suggests that AHR plays a greater role in the CD4<sup>+</sup> MTC and CD8<sup>+</sup> T<sub>CM</sub> subsets, including the highest protein expression of AHR in T<sub>CM</sub> after stimulation. The integrated assay data also suggested that HIF1A might be more important in the CD8<sup>+</sup> T<sub>EM</sub> population. However, interestingly, protein data for this factor showed that T<sub>EM</sub> populations were the only subset studied for which HIF1A concentration did not increase after stimulation. Ultimately other mechanisms (such as dynamic degradation by AHR-regulated genes like *VBPI* (VHL)) could be a part of differential environmental sensing mechanisms between MTC subsets<sup>65</sup>. The evidence for unique regulation of HIF1A and its targets in T<sub>EM</sub> shown here suggests a potentially important role in this MTC subset which will require further experimentation to elucidate.

MSC was the most highly ranked transcription factor in all MTC by the PageRank algorithm, but its expression was absent in naïve T cells. MSC is a repressor capable of binding E-box elements as either a homodimer or heterodimer with E2A (TCF3) and was initially found to be highly expressed in activated B cells

where it has been shown to play a role in promoting memory B cell differentiation<sup>33,66</sup>. We showed that it is also highly expressed in activated T cells as well. The potential interactions of MSC with E2A is particularly interesting in the context of memory T cells as motifs specific for this factor were found to surround several genes repressed in activated CD8<sup>+</sup> T cells in mice<sup>13</sup>. MSC is expressed by several CD4<sup>+</sup> effector T-cell subsets, including Tfh, Th17, and Treg cells in mice<sup>34,67</sup>. In human Th17 cells, MSC inhibits cellular response to IL-2 via STAT5B signaling<sup>68</sup>. Studies using genetic mouse models have shown that MSC deficiency leads to spontaneous gut and lung inflammation with age<sup>34</sup> while also enhancing inflammation and IL-22 secretion in inflammatory bowel disease models<sup>69</sup>. Despite its clear role in a variety of settings and evidence of its importance for MTC gene regulatory networks shown here, it is largely unclear exactly how MSC participates in regulating MTC, and what role it plays in activation of naïve and MTC respectively.

Stimulation of T cells led to dramatic changes in both gene expression and chromatin accessibility. Various cytokines genes (*IL4*, *IL1A*, and *IL22*) were induced after 48 h of stimulation in MTC but not in naïve T cells, potentially reflecting a mechanism for expediting effector gene expression that is unique to antigen-experienced cells. We showed that induction of these cytokine transcripts translates directly to higher levels of cytokine protein in MTC relative to naïve cells under similar stimulation conditions. This effect is most prominent in T<sub>EM</sub> subsets of both CD4<sup>+</sup> and CD8<sup>+</sup> MTC which express the highest levels of their respective effector cytokines. MTC also constitutively express a large number of genes that are not expressed in naïve T cells but are upregulated following stimulation. Continuous expression of effector genes in this category reflects the ability of MTC to maintain certain effector functions, even in the absence of active infection, as exemplified by increased expression of the gene *CXCR3* known to play a role in trafficking CD4<sup>+</sup> and CD8<sup>+</sup> T cells to peripheral sites of inflammation<sup>41</sup>. Several MTC-specific genes were augmented in their expression, having high levels in the resting state that were increased further following stimulation. These genes were correlated with accessible chromatin loci that were present only in MTC prior to stimulation, suggesting that these genes were poised for expression upon rechallenge. The transcription factor EOMES exemplifies this set, with higher expression in MTC potentially facilitating its central role in the maintenance/programming of these cells<sup>44,70</sup>.

Binding motifs specific for known activation-induced transcription factors such as AP-1 family members (e.g., BATF) and their binding partners NFAT and IRF were highly enriched in both stimulation- and primed-PAR, suggesting that these regions may also be maintained in an open state in MTC in order to more rapidly respond to TCR stimulation. Binding site motifs for transcription factors at primed-PAR separated the samples by lineage (CD4<sup>+</sup> vs CD8<sup>+</sup>) suggesting distinct epigenetic control of activation for each lineage. Interestingly, binding motifs within memory-PAR separated the samples by subset (T<sub>CM</sub> vs T<sub>EM</sub>) irrespective of lineage. The exact cues leading to specific MTC subset differentiation remain elusive; however, it is likely that epigenetic mechanisms targeting these memory-subset-specific PAR play a role when paired with simultaneous expression of master regulator transcription factors.

Transcription factor motifs may play multiple roles depending on whether they are maintained in an accessible state within naïve or MTC, and when matched with expression of their corresponding TF. For example, in all MTC, IRF8 motifs are enriched in primed- and stimulation-PAR, but network analysis suggests that this factor plays a greater role in expression changes of activated naïve T cells than in activated MTC, due perhaps to its lowered expression in MTC. Coupled with IRF8's known role in

driving effector T-cell generation in mice<sup>71,72</sup>, this may suggest repression of IRF8 as an important aspect of T-cell memory. Conversely, as noted above, the expression of EOMES is augmented in CD8<sup>+</sup> MTC before and after stimulation while expression of the competing factor T-BET is similar between naïve and MTC after stimulation. Expression of EOMES after stimulation in CD8<sup>+</sup> MTC is slightly decreased after stimulation, but still expressed to a greater extent than in stimulated naïve cells, perhaps leading to a higher overall ratio of EOMES/T-BET expression in stimulated MTC. This along with the presence of EOMES/T-BET binding motifs within both memory- and primed-PAR highlights a potential mechanism by which EOMES helps drive memory-specific gene expression, particularly in CD8<sup>+</sup> T<sub>EM</sub> cells, which express high levels of this factor in a resting state. In its absence, T-BET controls the effector cell transcriptional response after activation<sup>61</sup>. Accordingly, higher resting expression of T-BET and greater accessibility of its target sites in T<sub>EM</sub> may drive the effector-like phenotype of this subset.

In summary, MTC display lineage and subset-specific gene expression and chromatin accessibility patterns. These provide MTC with a unique epigenetic context driven by a history of their previous activation during encounters with antigen. In total, these memory-specific features enable MTC to adopt expression profiles of effector T cells more rapidly during secondary immune challenge.

## Methods

**Human subjects.** Whole blood samples from four deidentified individuals were obtained with informed consent in accordance with Emory University School of Medicine Institutional Review Board protocols, IRB00045821. PBMCs were separated by density gradient centrifugation (Ficoll-Paque, GE Healthcare), treated with ACK lysing buffer to remove red blood cells, and washed in PBS.

**MACS isolation and ex vivo stimulation.** CD4<sup>+</sup> or CD8<sup>+</sup> T cells were isolated from PBMC samples using MACS microbead (Miltenyi Biotec) isolation kits (CD4: #130-096-533, CD8: #130-096-495) via negative selection of non-target cells. Briefly,  $1 \times 10^7$  cells were resuspended in 40  $\mu$ l MACS buffer and incubated with CD4<sup>+</sup> or CD8<sup>+</sup> biotin-antibody cocktail and then incubated with 20  $\mu$ l of T-cell MicroBead cocktail for 10 min. The flow-through was collected from a MACS separator column and washed with 3 mL MACS buffer. Collected cells were separated for either immediate flow sorting (resting cells) or incubated for 42–48 h in complete Roswell Park Memorial Institute (RPMI) media with anti-CD3/CD28 beads (Gibco: #11131D) for ex vivo stimulation.  $10^6$  cells were added to cell suspensions in a 2:1 bead-to-cell ratio before being removed prior to flow staining and sorting.

**Flow cytometry isolation of human memory T-cell subsets.** Cells were resuspended at  $1 \times 10^6/100 \mu$ l in FACS buffer (PBS, 1% BSA, and 2 mM EDTA), stained with CCR7-BB515 (BD Biosciences: 565870) for 30 min at 37 °C, and then a cocktail of the following: CD3-V450 (Tonbo Biosciences: 75-0038), CD4-PE/Cy7 (Biolegend; 300511), CD8-FITC (Life technologies: MHCD08014), CD45RA-PE/TxRed (Biolegend: 304145), Zombie Yellow Fixable Viability Kit (Biolegend; 423104) for 30 min at 4 °C and then washed with 1 ml of FACS buffer. The following gating strategy was used to define memory subsets: lymphocytes were gated based on SSC-A / FSC-A, single cells by FSC-H / FSC-A, and live cells were based on exclusion of Zombie Yellow Fixable Viability Kit. T cells of the appropriate lineage were selected using the markers CD3, CD4, and CD8. Memory and naïve T-cell subsets were isolated using the markers CCR7 and CD45RA. Cell sorting was performed at the Emory Flow Cytometry Core using a FACSAria II (BD Biosciences) and BD FACSDiva software (BD Biosciences). Data were analyzed and figures generated using FlowJo v10.6.2. Supplementary Table 1 contains a list of all antibodies and the concentrations used.

**RNA sequencing.** One thousand cells were sorted directly into RLT buffer (79216; Qiagen) containing 1% 2-mercaptoethanol. RNA was isolated using the Quick-RNA Microprep kit (Zymo Research; R1050). The SMART-Seq v4 Ultra Low Input RNA Kit (634894; Takara Bio) was used for cDNA synthesis, and 400 pg of cDNA was used as input for the NexteraXT kit (Illumina) to create sequence libraries. DNA libraries were sequenced at the University of Alabama at Birmingham's Heflin center for genomics using a NextSeq500.

**RNA-sequencing data analysis.** Raw sequencing data were mapped to hg38 using STAR v.2.5.3<sup>73</sup>. Duplicate reads were identified and removed using PICARD



(<http://broadinstitute.github.io/picard/>). Reads per kilobase per million (rpkm) normalized gene expression counts were derived by analyzing coverage across all exons that fall within unique ENTREZ genes using the GenomicRanges package<sup>74</sup>. The Bioconductor package DESeq2<sup>75</sup> was used to determine differentially expressed genes (DEG) which were defined as having an absolute  $\log_2$  fold-change of  $\geq 1$  and a false discovery rate (FDR) of  $\leq 0.05$ . Differentially expressed genes are listed in Supplemental Data 1. All detected transcripts were pre-ranked for gene set enrichment analysis (GSEA)<sup>76</sup> by multiplying the sign of the fold change ( $\pm$ ) by  $-\log_{10}$  of the  $p$ -value. Heatmaps were generated using the ComplexHeatmap R package<sup>77</sup>. Three dimensional comparisons of differential expression used the normalized read counts which were averaged across MTC subset groups,  $\log_2$  transformed, and converted to barycentric coordinates for visualizing using the Tri-wise R package<sup>20</sup>. T-cell-relevant gene sets were obtained from the MSigDB v7.4 database gene ontology collection (GOBP-GO\_NATURAL\_KILLER\_CELL\_MEDIATED\_IMMUNITY; GOBP-LEUKOCYTE\_MEDIATED\_CYTOTOXICITY), or derived by taking the top 100 upregulated genes in CD8<sup>+</sup>CD101<sup>-</sup>Tim3<sup>-</sup> cells as described by Hudson, et al.<sup>21</sup>. Gene modules were discovered by analyzing genes with detected counts, filtering for DEG, and then filtering for genes which had an expression value  $\geq 3$  rpkm across all samples in any one sample group. Clustering was performed using fuzzy c-means clustering (using the 'cmeans' function of the e1071 R package) after first estimating the fuzzifier parameter<sup>78</sup> and a selecting a  $c$  value based on analysis of within sum of squared error. Module assignments for each gene are listed in Supplemental Data 2 (CD8<sup>+</sup>DEG) and Supplemental Data 3 (CD4<sup>+</sup>DEG).

**ATAC-seq.** For each sample, 1000–20,000 cells were sorted into FACS buffer and Tn5 transposition was performed<sup>79</sup>. Briefly, cells were resuspended in 12.5  $\mu$ l 2 $\times$  tagmentation DNA Buffer, 2.5  $\mu$ l Tn5, 2.5  $\mu$ l 1% Tween-20, 2.5  $\mu$ l 0.2% Digitonin, and 5  $\mu$ l H<sub>2</sub>O and incubated at 37 °C for 1 h. Cells were then lysed with the addition of 2  $\mu$ l 10 mg/ml Proteinase-K, 23  $\mu$ l Tagmentation Clean-up buffer (326 mM NaCl, 109 mM EDTA, 0.63% SDS), and incubated at 40 °C for 30 min. Tagmented DNA was purified and size selected for small fragments using AMPure XP beads (Beckman Coulter, A63881) and PCR amplified (Roche, KK2602) with dual indexing primers (Illumina, FC-131-2004) to generate a sequencing library. Final libraries were again purified, and size selected using AMPureXP beads, quantitated by QuBit (Life Technologies, Q33231), size distributions determined by bioanalyzer (Agilent 2100), pooled at equimolar ratios, and sequenced at the Emory Non-human Primate Genomics Core on a NovaSeq6000 using a PE100 run.

**ATAC-seq data analysis.** Raw sequencing data was mapped to the hg38 genome using Bowtie v1.1.1<sup>80</sup>. Peaks of accessibility enrichment were called using MACS2 v2.1.0<sup>81</sup> and annotated to the nearest gene using HOMER<sup>43</sup>. Differential testing of accessible peak regions was performed using DESeq2 using the cutoffs of FDR  $\leq 0.05$  and  $>1.5 \log_2$  fold-change to establish significance. Differentially accessible regions are listed in Supplemental Data 4. Count-based motif enrichment analysis was performed using the chromVAR<sup>31</sup> R package. For this analysis, peaks identified by MACS2 were restricted to a fixed width of 250 bp using the 'resize' function from the GenomicRanges<sup>74</sup> package in R. A matrix of counts for fragment insertions within these peaks was then generated using the 'getCounts' function from the ChromVAR package using the previously mapped sequencing reads. Overlapping peaks or those containing no fragment counts were excluded from the analysis and the motifs for transcription factor binding sites were sourced from the JASPAR<sup>82</sup> database. Discovery of patterned accessibility regions (PAR) was performed using a custom R script, which used FDR cutoffs in either resting or stimulated comparisons. For example, primed-PAR were defined as significantly differential between stimulated and unstimulated naïve cells with a non-zero  $\log_2$ FC as well as significant in differential comparison between unstimulated naïve and MTC. A list of all PAR assignments for each locus can be found in Supplemental Data 5. The HOMER 'findMotifsGenome.pl' function was used for de novo motif enrichment analysis or known motif enrichment from this database. Relative enrichment rank change values were calculated using HOMER by normalizing the enrichment  $p$ -values for individual motifs in each set of peaks by the total number of enriched motifs found in that peak set. Resulting values were then  $z$ -score scaled for relative comparison across groups.

**Integrated analysis and statistics.** For principal component analysis, normalized count data was mean scaled by row (across samples) and then analyzed using the 'princomp' function from the stats package in R. Phylogenetic analysis was performed by computing a Euclidean distance matrix between enrichment rank values and clustering using the 'hclust' function in R. Trees were plotted with APE v3.4<sup>83</sup> as unrooted trees. Bootstrapping was used to assess the reproducibility of clustering using the 'boot.phylo' function with 10,000 permutations. PageRank<sup>84</sup> analysis was performed using both normalized differential expression values derived from the RNA-seq data as well as raw ATAC-seq data. All statistical analyses were performed with R using DESeq2 for large-scale statistical analysis of RNA-seq or ATAC-seq data or using Wilcoxon Rank sum tests on individual genes with  $p \leq 0.05$  considered significant.

For all metabolic assays, a one-way ANOVA with multiple comparisons was used to determine significance. Three to six independent samples were used in these assays as indicated in the legends. Values  $p \leq 0.05$  were considered significant.

**Metabolic flux analysis (Seahorse assay).** ECAR was measured with XF96 Extracellular Flux Analyzer (Seahorse Bioscience). Human PBMCs isolated from three different donors. T cells were enriched by using human Pan T-cell isolation kit (Miltenyi Biotec). Cells were stained with antibodies to CD3 (BV450), CD4 (APC), CD8 (APC-Cy7), CD45RA (BV650), CCR7 (AF488) and ghost viability dye (BV510). Then, CD4<sup>+</sup> (naïve, T<sub>EM</sub>, T<sub>CM</sub>) populations were sorted by FACS. The cells were cultured in RPMI media supplemented with IL-7 and IL-15 for 2 days with or without anti-CD3/CD28 beads (1:2 ratio). Cells were harvested, washed, and then resuspended in XF RPMI media. Cells (180,000–250,000) were transferred into a poly-D-lysine coated 96-well plate as three technical repeats for each cell type and centrifuged at 400xg for 5 min to allow the cells to collect into monolayer at the bottom of the plate. The plate was incubated at 37 °C non-CO<sub>2</sub> incubator for 1 h and then placed into the XF96 Extracellular Flux Analyzer. Cells were monitored under basal conditions and in response to 10 mM glucose, 1  $\mu$ M oligomycin, 100 mM 2-Deoxy-Glucose. ECAR values were extracted from Agilent Seahorse Wave Desktop software and normalized according to the initial cell number. Glycolysis, glycolytic capacity and glycolytic reserve values were calculated according to the equations provided at Agilent Report Generator User Guide.

**Mitochondrial measurements.** T cells were enriched from human PBMCs by using a human Pan T-cell isolation kit (Miltenyi Biotec). The cells were cultured in RPMI media supplemented with IL-7 and IL-15 for 2 days with or without anti-CD3/CD28 beads (1:2 ratio). For assessment of mitochondrial mass, the cells were incubated with 25 nM MitoTracker Green FM (MTG, Invitrogen) for 1 h. To investigate the mitochondrial membrane potential of the cells, TMRM dye (Invitrogen) was added to a final concentration of 100 nM for the last 30 min of the above MTG incubation. Cells were stained with anti-CD3 (BV450), CD4 (PE-Cy7), CD8 (APC-Cy7), CCR7 (APC), and CD45RA (AF700) antibodies. Unfixed samples were immediately analyzed, and all samples were analyzed on a BD Fortessa flow cytometer.

**SCENITH assays.** T cells were enriched from freshly isolated human PBMCs. The cells were plated in a 96-well round bottom plate and rested at 37 °C, 5% CO<sub>2</sub> for 2 h. Some samples were stimulated with anti-CD3/CD28 beads (2:1 ratio) for 1 day. If the cells were not stimulated, the SCENITH protocol was performed after the 2 h resting step. In the SCENITH protocol, cells were untreated or treated with 2-deoxy-glucose (250 mM, Sigma-Aldrich), oligomycin (1.5  $\mu$ M, Sigma-Aldrich), or a combination of both drugs for 20 min at 37 °C. Puromycin (10  $\mu$ g/ml, Sigma-Aldrich) was added, and the cells were incubated at 37 °C for another 25 min. The cells were immediately washed with FACS buffer, then the surface staining was performed. The cells were incubated with anti-CD3 (V450), CD4 (PE-Cy7), CD8 (APC-Cy7), CCR7 (FITC), CD45RA (AF700) antibodies and ghost viability dye (BV510) in FACS buffer at 4 °C for 20 min. After washing with FACS buffer, the cells were fixed and permeabilized using Foxp3 intracellular staining kit (Invitrogen eBioscience) for 1 h at room temperature. Intracellular staining of puromycin was performed by using 1:1100 dilution of anti-puromycin (AF647) antibodies (Clone 12D10, Sigma-Aldrich) in kit supplied perm/wash buffer. The cells were incubated at room temperature for one hr, and then washed with perm/wash buffer twice. Finally, the cells were resuspended in FACS buffer and analyzed by flow cytometry. The mitochondrial dependence, glucose dependence, glycolytic capacity, fatty acid, and amino acid oxidation capacity (FAO and AAO) were calculated from the MFI of puromycin in the different treatments<sup>28</sup> with adjustment of range values across samples within a group if a value within the group was negative. Briefly, percent mitochondrial dependence is calculated as the difference between control and oligomycin-treated cells divided by the difference between control and 2-deoxy-glucose plus oligomycin-treated cells. Percent glucose dependence was calculated as difference between control and 2-deoxy-glucose-treated cells divided by the difference between control and 2-deoxy-glucose plus oligomycin-treated cells. Glycolytic capacity and FAO and AAO capacity were calculated by subtracting the percent mitochondrial capacity or percent glycolytic capacity from 100, respectively.

**Detection of intracellular proteins by flow cytometry.** For intracellular cytokine detection experiments, T cells were enriched from PBMC as above and stimulated for 48 h with anti-CD3/CD28 beads (2:1 ratio). The cells were incubated with both Golgi Stop (BD Biosciences) and Brefeldin A (Biolegend) for 5 h at 37 °C. Surface staining was performed with washed cells using by anti-CD3 (BV605), CD4 (BV785), CD8 (APC-Cy7), CCR7 (AF488), CD45RA (AF700) antibodies and ghost viability dye (BV510). After washing with FACS buffer, cells were fixed and permeabilized for 20 min using the Cytofix/Cytoperm kit (BD Bioscience). The cells were again washed with permeabilization buffer. Intracellular staining was performed to detect the cytokines by using IL-1 $\alpha$  (PE), IL-4 (APC), IL-2 (PE-Cy7), IL-5 (eFlour450), IFN $\gamma$  (BV711), TNF $\alpha$  (BV650), IL-22 (BUV737), and IL-17 (PerCP) antibodies in permeabilization buffer. The isotype controls used were Mouse IgG1,  $\kappa$  Isotype (PE), Rat IgG1,  $\kappa$  Isotype (APC), Rat IgG2a,  $\kappa$  Isotype (PE/Cy7), Rat IgG1

$\kappa$  Isotype (eFlour450), Mouse IgG1,  $\kappa$  Isotype (BV711), Mouse IgG1,  $\kappa$  Isotype (BV650), Mouse IgG1  $\kappa$  Isotype (BUV737), Mouse IgG1 (PerCP). The cells were incubated at 4 °C for 30 min, and then washed with permeabilization buffer and resuspended in FACS buffer. The samples were analyzed by flow cytometry.

To detect the transcription factors levels in unstimulated and stimulated T cells (with anti-CD3/CD28 beads for 2 days), surface staining was performed with anti-CD3 (BV650), CD4 (BV711), CD8 (APC-Cy7), CCR7 (AF488), and CD45RA (AF700) for 20 min at 4 °C. After washing the cells, they were fixed and permeabilized using the Fopx3 intracellular staining kit as above. Intracellular staining was done using antibodies to EOMES (PE-Cy5.5), T-BET (BV421), HIF1 $\alpha$  (APC), and AHR (PE-Cy7). Following staining for MSC, cells were restained with an anti-rabbit secondary antibody (PE). The isotype controls used were Mouse IgG1  $\kappa$  Isotype (PE/Cy5.5), Mouse IgG1,  $\kappa$  Isotype (BV421), Mouse IgG1 (APC), Mouse IgG2b  $\kappa$  Isotype (PE/Cy7), Rabbit IgG. The cells were washed with permeabilization buffer and resuspended in FACS buffer for flow cytometry analysis.

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All sequencing data have been deposited in NCBI Gene Expression Omnibus (GEO) under the following accession numbers GSE186463 for RNA-seq and GSE186462 for ATAC-seq. Raw data used to create Fig. 1f, 2b, l, 3c, d, f, h-p, 5i, 7g, h, 8d-g, 9b, e, f, and h are included in Supplemental Data 6.

### Code availability

Code and data processing scripts are available from the corresponding author upon request and at [https://github.com/cdschar/Rose\\_MTC\\_genomics](https://github.com/cdschar/Rose_MTC_genomics) or at <https://doi.org/10.5281/zenodo.7692608>.

Received: 24 March 2022; Accepted: 22 March 2023;

Published online: 03 April 2023

### References

- Ahmed, R. & Gray, D. Immunological memory and protective immunity: understanding their relation. *Science* **272**, 54–60 (1996).
- Hammarlund, E. et al. Duration of antiviral immunity after smallpox vaccination. *Nat. Med.* **9**, 1131–1137 (2003).
- Sallusto, F., Lenig, D., Förster, R., Lipp, M. & Lanzavecchia, A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* **401**, 708–712 (1999).
- Jameson, S. C. & Masopust, D. Understanding subset diversity in T cell memory. *Immunity* **48**, 214–226 (2018).
- Hope, J. L., Stairiker, C. J., Bae, E.-A., Otero, D. C. & Bradley, L. M. Striking a Balance—cellular and molecular drivers of memory T cell development and responses to chronic stimulation. *Front. Immunol.* **10**, 1596 (2019).
- Michie, C. A., McLean, A., Alcock, C. & Beverley, P. C. L. Lifespan of human lymphocyte subsets defined by CD45 isoforms. *Nature* **360**, 264–265 (1992).
- Hamann, D. et al. Phenotypic and functional separation of memory and effector human CD8+ T cells. *J. Exp. Med.* **186**, 1407–1418 (1997).
- Wills, M. R. et al. Human virus-specific CD8+ CTL clones revert from CD45ROhigh to CD45RAhigh in vivo: CD45RAhighCD8+ T cells comprise both naive and memory cells. *J. Immunol.* **162**, 7080–7087 (1999).
- Henson, S. M. et al. p38 signaling inhibits mTORC1-independent autophagy in senescent human CD8+ T cells. *J. Clin. Invest.* **124**, 4004–4016 (2014).
- Srivastava, R. et al. Phenotypic and functional signatures of Herpes Simplex virus-specific effector Memory CD73+CD45RAhighCCR7lowCD8+ TEMRA and CD73+CD45RAlowCCR7lowCD8+ TEM cells are associated with asymptomatic ocular Herpes. *J. Immunol.* **201**, 2315–2330 (2018).
- Araki, Y. et al. Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8+ T cells. *Immunity* **30**, 912–925 (2009).
- Best, J. A. et al. Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. *Nat. Immunol.* **14**, 404–412 (2013).
- Scharer, C. D., Bally, A. P. R., Gandham, B. & Boss, J. M. Cutting edge: chromatin accessibility programs CD8 T cell memory. *J. Immunol.* **198**, 2238–2243 (2017).
- Crompton, J. G. et al. Lineage relationship of CD8+ T cell subsets is revealed by progressive changes in the epigenetic landscape. *Cell. Mol. Immunol.* **13**, 502–513 (2016).
- Rodriguez, R. M. et al. Epigenetic networks regulate the transcriptional program in memory and terminally differentiated CD8+ T cells. *J. Immunol.* **198**, 937–949 (2017).
- Barski, A. et al. Rapid recall ability of memory T cells is encoded in their epigenome. *Sci. Rep.* **7**, 39785 (2017).
- Griffith, J. W., Sokol, C. L. & Luster, A. D. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu. Rev. Immunol.* **32**, 659–702 (2014).
- Mueller, S. N., Gebhardt, T., Carbone, F. R. & Heath, W. R. Memory T cell subsets, migration patterns, and tissue residence. *Annu. Rev. Immunol.* **31**, 137–161 (2013).
- Smith, K. A. The structure of IL2 bound to the three chains of the IL2 receptor and how signaling occurs. *Med. Immunol.* **5**, 3 (2006).
- van de Laar, L. et al. Yolk Sac macrophages, fetal liver, and adult monocytes can colonize an empty niche and develop into functional tissue-resident macrophages. *Immunity* **44**, 755–768 (2016).
- 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**, 1043–1058.e4 (2019).
- Drouillard, A. et al. S1PR5 is essential for human natural killer cell migration toward sphingosine-1 phosphate. *J. Allergy Clin. Immunol.* **141**, 2265–2268.e1 (2018).
- Cui, J. et al. Suppressed expression of LDHB promotes pancreatic cancer progression via inducing glycolytic phenotype. *Med. Oncol.* **32**, 143 (2015).
- Hermans, D. et al. Lactate dehydrogenase inhibition synergizes with IL-21 to promote CD8+ T cell stemness and antitumor immunity. *Proc. Natl. Acad. Sci. USA* **117**, 6047–6055 (2020).
- Presley, A. D., Fuller, K. M. & Arriaga, E. A. MitoTracker Green labeling of mitochondrial proteins and their subsequent analysis by capillary electrophoresis with laser-induced fluorescence detection. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **793**, 141–150 (2003).
- Perry, S. W., Norman, J. P., Barbieri, J., Brown, E. B. & Gelbard, H. A. Mitochondrial membrane potential probes and the proton gradient: a practical usage guide. *BioTechniques* **50**, 98–115 (2011).
- Mitri, D. D. et al. Reversible senescence in human CD4+CD45RA+CD27– memory T cells. *J. Immunol.* **187**, 2093–2100 (2011).
- Argüello, R. J. et al. SCENITH: a flow cytometry-based method to functionally profile energy metabolism with single-cell resolution. *Cell Metab.* **32**, 1063–1075.e7 (2020).
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. *Nat. Methods* **10**, 1213–1218 (2013).
- Maston, G. A., Evans, S. K. & Green, M. R. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genom. Hum. Genet.* **7**, 29–59 (2006).
- Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
- Yu, B. et al. Epigenetic landscapes reveal transcription factors regulating CD8+ T cell differentiation. *Nat. Immunol.* **18**, 573–582 (2017).
- Massari, M. E. et al. Characterization of ABF-1, a novel basic helix-loop-helix transcription factor expressed in activated B lymphocytes. *Mol. Cell. Biol.* **18**, 3130–3139 (1998).
- Wu, C. et al. The transcription factor musclin promotes the unidirectional development of peripheral Treg cells by suppressing the TH2 transcriptional program. *Nat. Immunol.* **18**, 344–353 (2017).
- Wang, H., Wei, Y. & Yu, D. Control of lymphocyte homeostasis and effector function by the aryl hydrocarbon receptor. *Int. Immunopharmacol.* **28**, 818–824 (2015).
- Palmer, C. S., Ostrowski, M., Balderson, B., Christian, N. & Crowe, S. M. Glucose metabolism regulates T cell activation, differentiation, and functions. *Front. Immunol.* **6**, 1 (2015).
- Semenza, G. L. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annu. Rev. Pathol. Mech. Dis.* **9**, 47–71 (2014).
- Tsai, C.-H. et al. NcoA2-dependent inhibition of HIF-1 $\alpha$  activation is regulated via AhR. *Toxicol. Sci.* **148**, 517–530 (2015).
- Maranchie, J. K. et al. The contribution of VHL substrate binding and HIF1- $\alpha$  to the phenotype of VHL loss in renal cell carcinoma. *Cancer Cell* **1**, 247–255 (2002).
- Wang, J., Dong, Z., Sheng, Z. & Cai, Y. Hypoxia-induced PVT1 promotes lung cancer chemoresistance to cisplatin by autophagy via PVT1/miR-140-3p/ATG5 axis. *Cell Death Discov.* **8**, 1–10 (2022).
- Groom, J. R. & Luster, A. D. CXCR3 in T cell function. *Exp. Cell Res.* **317**, 620–631 (2011).
- Bevington, S. L., Cauchy, P. & Cockerill, P. N. Chromatin priming elements establish immunological memory in T cells without activating transcription. *BioEssays* **39**, 1600184 (2017).
- 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).

44. Intlekofer, A. M. et al. Effector and memory CD8 + T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* **6**, 1236–1244 (2005).
45. Moskowitz, D. M. et al. Epigenomics of human CD8 T cell differentiation and aging. *Sci. Immunol.* **2**, eaag0192 (2017).
46. Almeida, L., Lochner, M., Berod, L. & Sparwasser, T. Metabolic pathways in T cell activation and lineage differentiation. *Semin. Immunol.* **28**, 514–524 (2016).
47. Gerriets, V. A. & Rathmell, J. C. Metabolic pathways in T cell fate and function. *Trends Immunol.* **33**, 168–173 (2012).
48. Quinn, K. M., Palchadhuri, R., Palmer, C. S. & La Gruta, N. L. The clock is ticking: the impact of ageing on T cell metabolism. *Clin. Transl. Immunol.* **8**, e01091 (2019).
49. van der Windt, G. J. & Pearce, E. L. Metabolic switching and fuel choice during T-cell differentiation and memory development. *Immunol. Rev.* **249**, 27–42 (2012).
50. van der Windt, G. J. W. et al. Mitochondrial respiratory capacity is a critical regulator of CD8+ T cell memory development. *Immunity* **36**, 68–78 (2012).
51. Buck, M. D. et al. Mitochondrial dynamics controls T cell fate through metabolic programming. *Cell* **166**, 63–76 (2016).
52. Corrado, M. & Pearce, E. L. Targeting memory T cell metabolism to improve immunity. *J. Clin. Invest.* **132**, e148546 (2022).
53. Raud, B., McGuire, P. J., Jones, R. G., Sparwasser, T. & Berod, L. Fatty acid metabolism in CD8+ T cell memory: challenging current concepts. *Immunol. Rev.* **283**, 213–231 (2018).
54. Menk, A. V. et al. Early TCR signaling induces rapid aerobic glycolysis enabling distinct acute T cell effector functions. *Cell Rep.* **22**, 1509–1521 (2018).
55. Giles, J. R. et al. Human epigenetic and transcriptional T cell differentiation atlas for identifying functional T cell-specific enhancers. *Immunity* **55**, 557–574.e7 (2022).
56. Zhao, X., Shan, Q. & Xue, H.-H. TCF1 in T cell immunity: a broadened frontier. *Nat. Rev. Immunol.* <https://doi.org/10.1038/s41577-021-00563-6> (2021).
57. Gattinoni, L. et al. A human memory T-cell subset with stem cell-like properties. *Nat. Med.* **17**, 1290–1297 (2011).
58. Youngblood, B. et al. Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature* **552**, 404–409 (2017).
59. Restifo, N. P. & Gattinoni, L. Lineage relationship of effector and memory T cells. *Curr. Opin. Immunol.* **25**, 556–563 (2013).
60. Xu, T., Pereira, R. M. & Martinez, G. J. An updated model for the epigenetic regulation of effector and memory CD8+ T cell differentiation. *J. Immunol.* **207**, 1497–1505 (2021).
61. Knox, J. J., Cosma, G. L., Betts, M. R. & McLane, L. M. Characterization of T-Bet and Eomes in peripheral human immune cells. *Front. Immunol.* **5**, 217 (2014).
62. Staal, F. J. T. & Clevers, H. Tcf/Lef transcription factors during T-cell development: unique and overlapping functions. *Hematol. J.* **1**, 3–6 (2000).
63. Rothhammer, V. & Quintana, F. J. The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. *Nat. Rev. Immunol.* **19**, 184–197 (2019).
64. Phan, A. T. et al. Constitutive glycolytic metabolism supports CD8+ T cell effector memory differentiation during viral infection. *Immunity* **45**, 1024–1037 (2016).
65. Kim, J. A. et al. VBP1 represses cancer metastasis by enhancing HIF-1 $\alpha$  degradation induced by pVHL. *FEBS J.* **285**, 115–126 (2018).
66. Chiu, Y.-K. et al. Transcription factor ABF-1 suppresses plasma cell differentiation but facilitates memory B cell formation. *J. Immunol.* **193**, 2207–2217 (2014).
67. Debuissson, D., Mari, N., Denanglaire, S., Leo, O. & Andris, F. Myor/ABF-1 Mrna expression marks follicular helper T cells but is dispensable for Tfh cell differentiation and function in vivo. *PLoS ONE* **8**, e84415 (2013).
68. Santarlasci, V. et al. Muscadin inhibits human T-helper 17 cell response to interleukin 2 by controlling STAT5B activity. *Eur. J. Immunol.* **47**, 1427–1442 (2017).
69. Yu, J. et al. Muscadin deficiency aggravates colonic injury and inflammation in mice with inflammatory bowel disease. *Inflammation* **43**, 1455–1463 (2020).
70. Intlekofer, A. M. et al. Anomalous type 17 response to viral infection by CD8+ T cells lacking T-bet and eomesodermin. *Science* **321**, 408–411 (2008).
71. Miyagawa, F. et al. Interferon regulatory factor 8 integrates T-cell receptor and cytokine-signaling pathways and drives effector differentiation of CD8 T cells. *Proc. Natl Acad. Sci. USA* **109**, 12123–12128 (2012).
72. Sun, L. et al. Interferon regulator factor 8 (IRF8) limits ocular pathology during HSV-1 infection by restraining the activation and expansion of CD8+ T cells. *PLoS ONE* **11**, e0155420 (2016).
73. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
74. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
75. 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).
76. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
77. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
78. Schwämmle, V. & Jensen, O. N. A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics* **26**, 2841–2848 (2010).
79. Guo, M. et al. EZH2 represses the B cell transcriptional program and regulates antibody-secreting cell metabolism and antibody production. *J. Immunol.* **200**, 1039–1052 (2018).
80. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
81. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
82. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
83. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinforma. Oxf. Engl.* **20**, 289–290 (2004).
84. 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).

### Acknowledgements

We thank members of the Boss and Scherer laboratories for their scientific contribution and critical reading of the manuscript, Tian Mi for initial bioinformatic analysis and the Emory Flow Cytometry Core for FACS. This work was supported by the National Institute of Allergy and Infectious Diseases grants RO1 AI113021 to J.M.B., T32 GM0008490 to J.R.R., and F32 AI161857 to M.D.P.

### Author contributions

A.R., C.S., B.A., and J.M.B. designed the experiments, and J.R. analyzed the omics data. B.A. and A.R. performed the experiments, and analyzed the cellular data, and S.H. created the sequencing libraries. B.A. and M.P. analyzed the flow cytometry data. J.M.B. and C.S. conceived and supervised the study. J.R., J.M.B., and B.A. wrote the manuscript. All authors discussed and commented on the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-04747-9>.

**Correspondence** and requests for materials should be addressed to Jeremy M. Boss.

**Peer review information** *Communications Biology* thanks Jianxun Song, Jörg Goronzy, Abhinav Jain and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editors: Guideng Li and Zhijuan Qiu.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023