HHS Public Access

Author manuscript

Nat Immunol. Author manuscript; available in PMC 2017 September 13.

Published in final edited form as:

Nat Immunol. 2017 May ; 18(5): 573–582. doi:10.1038/ni.3706.

Epigenetic landscapes reveal transcription factors regulating CD8+ T cell differentiation

Bingfei Yu[1,11], **Kai Zhang**[2,11], **J. Justin Milner**[1], **Clara Toma**[1], **Runqiang Chen**[3,4], **James P. Scott-Browne**[5], **Renata M. Pereira**[5,10], **Shane Crotty**[3,6], **John T. Chang**[7], **Matthew E. Pipkin**[4], **Wei Wang**[2,8,9], and **Ananda W. Goldrath**[1]

[1]Division of Biological Sciences, University of California, San Diego, La Jolla, CA

[2]Bioinformatics and Systems Biology Graduate Program, University of California, San Diego, La Jolla, CA

[3]Division of Vaccine Discovery, La Jolla Institute for Allergy and Immunology, La Jolla, CA

[4]Department of Immunology and Microbial Science, The Scripps Research Institute, Jupiter, FL

[5]Division of Signaling and Gene Expression, La Jolla Institute for Allergy and Immunology, La Jolla, CA

[6]Division of Infectious Diseases, Department of Medicine, University of California, San Diego, La Jolla, CA

[7]Department of Medicine, University of California, San Diego, La Jolla, CA

[8]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA

[9]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA

Abstract

Dynamic changes in the expression of transcription factors (TFs) can influence specification of distinct CD8+ T cell fates, but the observation of equivalent expression of TF among differentially-fated precursor cells suggests additional underlying mechanisms. Here, we profiled genome-wide histone modifications, open chromatin and gene expression of naive, terminal-effector, memory-precursor and memory CD8+ T cell populations induced during the *in vivo* response to bacterial

Correspondence should be addressed to A.W.G. (agoldrath@ucsd.edu) and W.W. (wei-wang@ucsd.edu).

[10]Present address: Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

[11]These authors contributed equally to this work

AUTHOR CONTRIBUTIONS

B.Y. designed and performed experiments, analyzed the data and wrote the paper; K.Z. performed computational analysis and wrote the paper; B.Y., J.J.M., R.C. and C.T. performed shRNA knockdown; R.M.P and J.P.S-B. provided ATAC-seq datasets for polyclonal CD8+ T cell populations. M.E.P. and S.C. provided reagents, advice for the design of experiments and analysis of experiments and assisted in writing the paper; J.T.C. provided advice and assisted in writing the paper. W.W. supervised the computational analysis and wrote the paper; and A.W.G. supervised the project, designed the experiments, analyzed the data, and wrote the paper.

COMPETING FINANCIAL INTERESTS STATEMENT

The authors declare no competing financial interests.

ACCESSION CODES

GEO: ATAC-seq and ChIP-seq: GSE89036, microarray: GSE89037

infection. Integration of these data suggested that TF expression and binding contributed to establishment of subset-specific enhancers during differentiation. We developed a new bioinformatics method using the PageRank algorithm to reveal novel TFs influencing the generation of effector and memory populations. The TFs YY1 and Nr3c1, both constitutively expressed during $CD8^+$ T cell differentiation, regulated the formation of terminal-effector and memory-precursor cell-fates, respectively. Our data define the epigenetic landscape of differentiation intermediates, facilitating identification of TFs with previously unappreciated roles in $CD8^+$ T cell differentiation.

INTRODUCTION

In response to infection, naive $CD8^+$ T cells differentiate into a heterogeneous population of pathogen-specific effector $CD8^+$ T cells. While the majority of these T cells undergo apoptosis after resolution of infection, a small fraction persists as memory cells, providing lasting protection from re-infection[1]. Recent studies demonstrate that commitment of $CD8^+$ T cell fate occurs early after infection, and the differential expression of killer cell lectin-like receptor (KLRG1) and interleukin-7 receptor (IL-7R) may be used to distinguish two effector subsets with distinct memory potential: terminally-differentiated effector (TE, $KLRG1^{hi}IL\text{-}7R^{lo}$) and memory-precursor effector (MP, $KLRG1^{lo}IL\text{-}7R^{hi}$) $CD8^+$ T cells[2,3]. Numerous TFs have been identified as critical regulators of $CD8^+$ T cell fate including T-bet, Blimp-1, Id2, IRF4, BATF, and Zeb2 for TE and effector populations; TCF-1, Eomes, Id3, E proteins, Bcl-6, and FOXO1 for MP and memory populations[2–5]. Notably, not all these factors exhibit differential expression between the TE and MP subsets, suggesting that additional mechanisms contribute to their activity in promoting cell fates. Further, how these TFs function within a coherent regulatory network is unknown, and additional TFs relevant in $CD8^+$ T cell differentiation remain unidentified.

We reasoned that integrated analysis of TF expression, binding, and the expression of their gene targets would provide additional insights to identify previously unappreciated TFs involved in $CD8^+$ T cell differentiation. Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) has recently been used to globally probe open chromatin to map TF binding regions with high genomic resolution requiring minimal material[6,7]. By scanning TF binding motifs on accessible chromatin regions, it is possible to infer the binding of hundreds of TFs and identify potential gene targets of these TFs simultaneously, which has previously been technically impossible to achieve[8]. ATAC-seq proves powerful for pinpointing TF binding sites within regulatory elements characterized by active epigenetic marks such as: promoters marked by trimethylation of histone H3 lysine 4 (H3K4me3); enhancers associated with monomethylation of histone H3 lysine 4 (H3K4me1) and acetylation of histone H3 lysine 27 (H3K27ac)[9–11]. Additionally, trimethylation of histone H3 lysine 27 (H3K27me3) is associated with gene repression[10]. Recent studies combining ATAC-seq and histone modifications have facilitated the prediction of TFs and enhancers that define tissue-specific macrophages and of lineage-determining TFs in hematopoiesis[12,13]. In naive $CD8^+$ T cells, co-deposition of H3K4me3 and H3K27me3 at promoter regions is a signature of genes important for cellular differentiation, suggesting an epigenetic mechanism underlying $CD8^+$ T cell

differentiation[14,15]. However, these studies focused exclusively on promoters. Accumulating evidence suggests that enhancers also play a key role in fine-tuning gene expression, providing better specificity compared with promoters[12,16]. However, enhancer landscapes important for effector and memory CD8$^+$ T cell differentiation remain largely unknown.

Here, we characterized the epigenetic landscapes of naive, TE, MP, and memory CD8$^+$ T cells generated during bacterial infection to identify both enhancer and promoter regions important for CD8$^+$ T cell differentiation. Using ATAC-seq to identify accessible regulatory regions, we predicted TF candidates and further constructed a transcriptional regulatory network for each subset. To facilitate the identification of key TFs, we developed a new bioinformatics method using the PageRank algorithm to rank the importance of TF in each regulatory network. We identified TFs known to be central to CD8$^+$ T cell differentiation and TFs not previously associated with CD8$^+$ T cell fate specification. Among these, we experimentally validated that Yin and Yang-1 (YY1) and Nuclear Receptor Subfamily 3 Group C member 1 (Nr3c1) promote TE cell and MP cell phenotypes respectively. Taken together, our results yielded a comprehensive catalog of the regulatory elements of CD8$^+$ T cells, revealing unexpected regulators controlling CD8$^+$ T cell fate. Furthermore, our computational framework can be applied generally to any cell or tissue type to decipher regulatory networks and identify biologically-important TFs.

RESULTS

Differential gene expression by TE and MP CD8$^+$ T cells

The effector CD8$^+$ T cell population is characterized by extensive phenotypic and functional heterogeneity, including TE and MP subsets[2]. Microarray analysis of TE and MP subsets revealed differentially expressed genes between these two subsets on day 8 of infection, and when compared to gene-expression data for total effector and memory CD8$^+$ T cell populations, genes upregulated in the TE versus MP subsets were enriched in total effector versus memory CD8$^+$ T cells, respectively (Supplementary Fig. 1a,b)[17]. This result indicates the unique transcriptional identities of effector and memory CD8$^+$ T cells can be captured by analysis of TE and MP subsets. Interestingly, the differences in abundance of mRNA and protein for the majority of TFs known to control differentiation of TE versus MP subset were subtle (Supplementary Fig. 1c,d), suggesting that expression differences alone do not account for the differential dependence on TF in distinct subsets. RNA-seq for TE and MP subsets were consistent with our microarray analyses; TE-specific and MP-specific gene signatures were enriched in total effector and memory CD8$^+$ T cells respectively, and many of the key TFs were similarly expressed between the TE and MP subsets at these time points (Supplementary Fig. 1e–g and Supplementary Table 1). Thus, besides TF expression, additional regulatory mechanisms, such as the control of TF binding, may contribute to differentiation of these two subsets and subsequent formation of long-lived memory cells.

CD8$^+$ T cell subsets exhibit distinct enhancer repertoires

Spatial and temporal regulation of gene expression requires the specific binding of TFs at regulatory elements, which is affected by chromatin state and accessibility. We combined ChIP-seq of histone modifications (H3K4me1, H3K4me3, H3K27ac, and H3K27me3) for

characterization of potential enhancer and promoter elements and ATAC-seq to integrate chromatin state and accessibility for each CD8$^+$ T cell subset, allowing prediction of TF binding at specific regulatory elements. OT-I TCR-transgenic CD8$^+$ T cells that specifically recognize a peptide fragment of ovalbumin (OVA) were transferred to host mice, followed by infection with *Listeria monocytogenes* engineered to express recombinant OVA (Lm-OVA)[17]. Naive, TE, MP and memory CD8$^+$ T cell populations were sorted for ChIP-seq and ATAC-seq (Supplementary Fig. 2a,b). Notably, OT-I and polyclonal CD8$^+$ T cells responding to infection show highly correlated gene expression throughout the immune response[17], and OT-I and polyclonal effector and memory CD8$^+$ T cells display similar ATAC-seq profiles (data not shown).

Previous studies showed that bivalent chromatin domains, comprising H3K4me3 and H3K27me3 modifications, exist in the promoters of effector genes in naive cells, and H3K27me3 occupancy at these promoters was reduced upon differentiation into effector CD8$^+$ T cells[14,15]. We also observed this pattern in the change of bivalent modification of effector genes including *Tbx21* (Fig. 1a). Conversely, we found that genes enriched in naive T cells, such as *Tcf7*, became repressed in effector CD8$^+$ T cells, concomitant with increased H3K27me3 occupancy at promoters (Fig. 1a). Further, the percentage of genes with H3K27me3 occupancy at promoters was higher during differentiation to effector compared to memory CD8$^+$ T cells, suggesting that epigenetic repression of genes enriched in naive CD8$^+$ T cells may be essential for terminal differentiation of effector CD8$^+$ T cells (Fig. 1b).

Focusing on distal regulatory regions of well-characterized genes in CD8$^+$ effector and memory T cells, we found both gains and losses of enhancer and repressive H3K27me3 marks. For example, *Gzma*, a characteristic effector gene highly expressed in TE cells, was associated with increased H3K4me1 and H3K27ac upon differentiation from naive CD8$^+$ T cells to the TE subset (Fig. 1c). Conversely, *Il7r* exhibited higher H3K4me1 and H3K27ac in MP and memory CD8$^+$ T cells compared to the TE subset, consistent with its role promoting long-term survival of memory CD8$^+$ T cells (Fig. 1c)[3,18]. Alternatively, *Id2* and *Id3*, encoding established transcriptional regulators of CD8$^+$ T cell differentiation, exhibited high occupancy of H3K4me1 in all CD8$^+$ T cells, but were associated with differential intensities of H3K27ac and H3K27me3 during differentiation (Fig. 1c)[19,20]. Thus, as expected, combinatorial epigenetic marks set the stage for gene expression.

To systematically identify putative enhancers, we applied a machine learning algorithm called RFECS[21]. RFECS identified 27,236, 26,561, 23,302, and 21,883 enhancers in naive, TE, MP, and memory CD8$^+$ T cells, respectively, comprising a non-redundant set of 52,331 putative enhancers. Upon differentiation of naive CD8$^+$ T cells during infection, TE gain a greater number of newly-formed enhancers relative to MP and memory cells, while all populations lose a similar number of enhancers (Fig. 2a). To understand the dynamic usage of enhancers during differentiation, we performed *k*-means clustering on 52,331 enhancers according to their H3K4me1 intensity. Enhancers were separated into five distinct clusters (Fig. 2b). In cluster V, H3K4me1 intensity was maintained equivalently across the CD8$^+$ T cell subsets, and genes associated with this cluster were highly expressed in all subsets (*Cd8a* and *Lck*). H3K4me1 intensity was increased in clusters I and II during differentiation,

showing enrichment in the TE subset compared to MP and memory CD8$^+$ T cells. Genes associated with these clusters were associated with TE subset differentiation (*Klrg1* and *Tbx21*) (Fig. 2b)[2]. In cluster III, H3K4me1 intensity was higher in all differentiated subsets compared to naive CD8$^+$ T cells, and genes associated this enhancer cluster were involved in CD8$^+$ T cell activation such as *Prf1*. Conversely, for cluster IV, H3K4me1 intensity was decreased during differentiation of naive CD8$^+$ T cells to the TE subset, and increased in MP and memory CD8$^+$ T cells compared to the TE subset. Enhancers of canonical regulators of memory potential and homeostasis were found in cluster IV (*Il7r* and *Cxcr4*) (Fig. 2b)[3,22].

To test if differential establishment of enhancers regulate subset-specific gene expression, we assigned enhancers to the nearest genes and compared gene expression during CD8$^+$ T cell differentiation. Cluster I, II, and III enhancers were associated with genes upregulated in activated CD8$^+$ T cells and cluster IV enhancers were associated with genes enriched in naive CD8$^+$ T cells (Supplementary Fig. 3a). Notably, the expression of genes with cluster I and II enhancers were more enriched in the TE subset, while genes with cluster IV enhancers were more enriched in the MP subset (Supplementary Fig. 3b). We performed Gene Ontology (GO) analysis using the GREAT tool[23] and found that the IL-12 signaling pathway was enriched in cluster I and II, consistent with the role of IL-12 in promoting the TE subset differentiation[2]. In addition, TGF-β and EGF signaling pathways were enriched in cluster IV, suggesting that these signaling pathways may favor the naive and/or memory T cell state, consistent with data showing TGF-β signaling is required for memory T cell differentiation (Fig. 2c)[24]. We further observed that genes associated with increased number of enhancers correlated with higher expression compared with those associated with a single enhancer (Fig. 2d).

TF motif enrichment at subset-specific regulatory regions

We reasoned that TF binding motifs would be enriched in accessible regulatory regions, which could be used to discover TFs important for CD8$^+$ T cell differentiation. Thus, we identified subset-specific open enhancers and promoters and then scanned 761 unique known TF-binding motifs at the center of the ATAC-peaks of these regulatory regions (Supplementary Fig. 4a,b). For example, the T-bet binding motif was enriched in a TE-specific accessible enhancer near the *Zeb2* gene, which was exclusively expressed in the TE subset, supporting previous findings that T-bet directly regulates *Zeb2* to promote TE subset differentiation (Fig. 3a)[25,26]. Our motif enrichment analysis predicted putative binding of known TFs at promoters and enhancers[4,27–31]. T-bet, BATF, SREBP2 and AP-1 binding motifs were depleted in naive and enriched in all differentiated CD8$^+$ T cell subsets, consistent with their crucial roles in CD8$^+$ T cell activation and effector function (Fig. 3b, Supplementary Fig. 4c)[27–29,32]. Tcf7, Lef1 and E2A binding motifs were depleted in TE and enriched in naive, MP and memory CD8$^+$ T cells, corresponding with well-characterized roles in regulating differentiation of memory populations (Fig. 3b)[4,30,31]. Enrichment of some TF binding motifs (Tcf7 and T-bet) was highly correlated with gene expression and function[2,4]; in contrast, the enrichment of other TF binding motifs (SREBP2 motif enriched in effector T cells and E2A motif enriched in MP and memory cells) was consistent with their demonstrated roles (SREBP2 maintaining effector T cell activation and E2A promoting MP and memory cell differentiation), yet they were not differentially expressed (Fig.

3c)[28,31]. These data indicated that subset-specific enhancers and promoters might be established by key TFs and that putative TF binding, in addition to differential expression, must be considered when identifying TF involvement.

Construction of CD8$^+$ T cell subset TF regulatory networks

To elucidate TF-mediated regulatory mechanisms underlying CD8$^+$ T cell differentiation, we sought to construct a TF regulatory network in different CD8$^+$ T cell subsets. Previous studies have applied gene co-expression correlation to construct regulatory networks[33,34]; however, this does not consider direct TF-binding. We combined TF-binding motifs, chromatin states and chromatin accessibility information to predict and link TF-binding sites to their potential gene targets (Supplementary Fig 5). We reconstructed TF regulatory networks and uncovered critical regulatory circuits responsible for CD8$^+$ T cell differentiation. For example, we identified a substantial number of putative targets regulated by T-bet in both the TE and MP subsets (Fig. 4a, Supplementary Table 2). We compared predicted T-bet-regulated genes between the TE and MP subsets and found that 61.4% of candidate genes were shared, such as *Ifng* and *Cxcr3,* which are well-established T-bet–regulated targets important for effector function (Fig. 4b)[35,36]. Interestingly, the subset-specific T-bet regulatory circuits predicted that T-bet uniquely controls the expression of *Zeb2*, *Gzma* and *Klrb1c* in TE cells and *Bcl2*, *Crtam* and *Pou6f1* in MP cells. To validate our analyses, we co-transferred *Tbx21*$^{+/+}$ and *Tbx21*$^{-/-}$ OT-I CD8$^+$ T cells into hosts followed by Lm-OVA infection. Given the loss of the TE subset with T-bet deficiency, we sort-purified total donor CD8$^+$ T cells or the MP subset from *Tbx21*$^{+/+}$ and *Tbx21*$^{-/-}$ populations and compared the mRNA expression of candidate genes. In the absence of T-bet, there was a ~200-fold decrease of *Zeb2* expression in total donor CD8$^+$ T cells compared to a 5-fold decrease in the MP subset, indicating the regulation of *Zeb2* expression by T-bet in the TE rather than the MP subsets (Fig. 4c,d)[25,26]. To avoid the bias of a complete loss of TE cells in T-bet-deficient T cells, we compared mRNA abundance for TE subsets derived from *Tbx21*$^{+/+}$ and *Tbx21*$^{+/-}$ populations and confirmed a decreased expression of *Zeb2*, *Gzma* and *Klrb1c* in the TE subset with loss of T-bet (Fig. 4e,f). Interestingly, the loss of T-bet impacted the expression of *Bcl2*, *Crtam* and *Pou6f1* in the MP subset, suggesting that T-bet regulates these genes in a MP-specific manner (Fig. 4d,f). Importantly, the absence of T-bet resulted in a defect in accumulation of MP cells over the course of infection, consistent with T-bet also regulating memory differentiation (Fig. 4g)[27]. Thus, we demonstrated that T-bet positively regulates different genes in distinct CD8$^+$ T cell subsets, highlighting that this approach allows prediction of potential gene targets unique to different CD8$^+$ T cell subsets.

Identification of key TFs from PageRank-based TF ranking

Constitutively-expressed TFs can exert cell-type-specific functions via regulating distinct gene expression, but incorporating this knowledge to identify key TF remains challenging because the TF targets are largely unknown. To overcome this limitation, we leveraged the TF regulatory network and developed a new bioinformatics method using the personalized PageRank algorithm[37] to assess the importance of each TF in the regulatory network (See Online Methods) (Fig. 5a). The TF ranks determined by our method are hence influenced by the number of genes and the importance (determined from their expression) of genes

regulated by the query TF. Thus, TFs that regulate more important genes would receive higher ranks.

Using PageRank analysis, we predicted the top 100 key TFs important for $CD8^+$ T cell differentiation. We compared our PageRank analysis with motif enrichment analysis used by prior studies[12,13] to investigate how many TFs reported previously as essential regulators of $CD8^+$ T cell differentiation can be recovered from predicted TF pools. We found that approximately half of the predicted TFs were shared by both analyses, and 25% of these shared TFs were identified in previous studies (Supplementary Fig. 6a,b). PageRank analysis revealed more known TFs compared with motif-enrichment analysis (22% in PageRank-specific compared to 5% in motif-enrichment-specific) (Fig. 5b). For example, PageRank analysis scored STAT3 higher in memory compared to the TE subset (Fig. 5c). This was consistent with the role of STAT3 in promoting mature and self-renewing memory $CD8^+$ T cells[38]. Additionally, more TFs consistent with known roles in $CD8^+$ T cell differentiation were recovered by PageRank analysis compared to another method, TF activity (TFA) analysis[39] (Supplementary Fig. 6c,d). These data highlighted the robustness of PageRank analysis, suggesting that unknown TFs predicted by PageRank analysis might be critical for $CD8^+$ T cell differentiation.

Validation of PageRank-predicted novel TFs

To highlight the power of PageRank analysis, we focused on YY1 and Nr3c1, two regulators identified by PageRank analysis but not by the motif-enrichment analysis. Although the expression of *Yy1* and *Nr3c1* did not change during $CD8^+$ T cell differentiation (Supplementary Fig. 6e), YY1 was ranked highly in the TE subset while Nr3c1 ranked highly in the MP subset (Fig. 5c). YY1 is a TF involved in transcriptional activation and repression, important in immune cell development including the differentiation of B, T helper 2 (T_H2) and regulatory T cells[40–42]. *Nr3c1* encodes glucocorticoid receptor, which translocates into the nucleus to regulate gene expression after binding to glucocorticoids in the cytosol. Nr3c1 plays a critical role in development, metabolism, and the immune response[43–45]. The role of YY1 and Nr3c1 in effector or memory $CD8^+$ T cell differentiation in the response to infection is unknown.

Based on the PageRank predictions, we hypothesized abrogated expression of YY1 and Nr3c1 would affect the formation of TE or MP subsets, respectively. To test if YY1 is essential for TE subset differentiation, we co-transferred $CD8^+$ T cells infected with retrovirus encoding shRNA targeting *Yy1* (sh*Yy1*) or control shRNA targeting *Cd19* into recipient mice followed by Lm-OVA infection and followed effector T cell differentiation (Supplementary Fig. 7a). RT-qPCR confirmed knockdown of *Yy1* expression resulting in a 54% reduction in mRNA abundance (Fig. 6a). Flow cytometric analysis of $CD8^+$ T cell subsets on day 7 of infection showed a significant reduction in both the frequency and number of the TE subset after knockdown of *Yy1* (Fig. 6b,c). In addition, the expression of MP-associated molecules including CD27, CXCR3, and TCF-1 were significantly increased after knockdown of *Yy1* (Fig. 6d)[4]. Furthermore, cytokine production by sh*Yy1*- and control-shRNA transduced cells showed that mean expression and the number of IFN-γ-

producing cells were reduced in the absence of YY1 (Fig. 6e). Together, these data confirmed that YY1 is important for TE subset differentiation.

Similarly, we determined how lowering *Nr3c1* expression impacted the MP subset differentiation. Knockdown of *Nr3c1* resulted in 86% reduction compared to control shRNA (Fig. 7a). Importantly, both frequency and number of MP cells were significantly decreased after knockdown of *Nr3c1* (Fig. 7b,c). Consistent with a loss of IL-7R-expressing cells, expression of MP-associated molecules including CD27, CXCR3, and TCF-1 was significantly reduced after knockdown of *Nr3c1* (Fig. 7d), supporting a functional role for Nr3c1 in MP subset differentiation. We monitored the percentage of MP cells from day 8 to day 30 of infection and observed a decrease of the percentage of MP subset on day 30 after the loss of Nr3c1 (Fig. 7e,f). Nr3c1 has been shown to interact with co-factors such as Nuclear receptor co-repressor 1 (Ncor1) to modulate hormone-response gene expression[46]. Notably, *Ncor1* shRNA knockdown similarly affected MP subset differentiation; the frequency of MP cells was decreased after knockdown of *Ncor1* (Supplementary Fig. 7b,c). To further confirm the role of Nr3c1 in MP subset differentiation, we treated mice with synthetic glucocorticoid dexamethasone (Dex) for 7 days and observed that the frequency of the MP subset was significantly increased after Dex treatment (Supplementary Fig. 7d,e). Collectively, these data demonstrated that glucocorticoid receptor Nr3c1 promotes MP subset differentiation.

DISCUSSION

The function and differentiation state of immune cells are controlled by TFs that relay environmental cues through regulation of gene expression. Efficient transcriptional regulation requires the interplay between TFs and chromatin remodelers to control TF binding with high fidelity. Key information is encoded in regulatory elements that contain TF binding sequences and are associated with specific histone modifications that influence accessibility, structure, and location of those elements[16]. To identify the TF-mediated regulatory circuits critical for $CD8^+$ T cell differentiation, we systematically characterized the epigenome of $CD8^+$ T cell subsets during pathogen infection. Our global map of regulatory elements revealed a dynamic pattern of enhancer establishment that foreshadows specific gene-expression programs. Our network analysis of T-bet regulatory circuits in distinct effector subsets revealed overlapping and distinct T-bet-targets between TE and MP subsets. This analysis suggests a novel function for T-bet in maintaining MP cell accumulation, potentially through regulating anti-apoptotic protein Bcl-2 and additional targets. Studies of distinct targets will further elucidate nuanced functions of T-bet in driving effector and memory fates.

Numerous crucial TFs modulating $CD8^+$ T cell differentiation have been identified based on differential gene expression and TF-gene co-expression correlation[17,33,34]. However, alterations of TF binding without changes of expression also result in differential expression of downstream gene targets, making it clear that identification of relevant TFs based exclusively on gene-expression analysis provides only a partial understanding of the TF networks involved. Indeed, our data demonstrated that gene expression alone does not fully explain the mechanisms behind cell-fate determination and support the idea that TF binding

and gene expression should be considered together to facilitate the identification of important TFs. Differential TF binding can be achieved via numerous mechanisms including: variable chromatin state and accessibility, TF localization, availability of co-factors, and post-translational modification of TFs. Our approach represents an advance in the efforts to achieve a comprehensive view of the regulatory networks that establish $CD8^+$ effector and memory T cell fates by integrating data describing mRNA expression as well as chromatin states and accessibility.

To prioritize these data, it is essential to develop new methods that rank the potential importance of TFs based on the quantity and quality of TF-regulated genes. Here, we applied the personalized PageRank algorithm to rank the absolute importance of TFs in each subset and relative importance across cell types by considering both TF-binding and gene expression. Importantly, our method ranks TFs by integrating two features, distinct weights for TF-regulated genes assessed by differential expression and a hierarchy of TF-to-TF circuitry. This strategy allows the identification of TFs regulating relatively few but important genes, which are often overlooked by other analyses. Future modifications of gene weights using gene ontology could facilitate identification of TFs important in specific functions or pathways.

We also validated the functions of two TFs identified by PageRank, YY1 and Nr3c1, demonstrating their essential roles in TE and MP subset differentiation, respectively. YY1 has been shown to modulate long-range chromatin interaction of cytokine loci in T_H2 cells[42]. How YY1 regulates TE subset differentiation and if YY1 controls chromatin interactions in the TE subset remains to be determined. Glucocorticoid receptor (Nr3c1) has been shown to regulate thymocyte apoptosis and inflammation response[43–45]. Here, we showed that Nr3c1 promotes MP differentiation, consistent with a role of glucocorticoids in inducing IL-7R expression[44]. Treatment with dexamethasone increased the proportion of MP subset during differentiation, demonstrating a novel role of glucocorticoid hormone in modulating $CD8^+$ T cell differentiation and a provocative potential strategy for manipulating memory cell differentiation. Thus, using our framework, we can both identify critical TFs and predict microenvironmental signals involved in regulating $CD8^+$ T cell differentiation.

Despite successful validation of TFs predicted by our computational framework, additional factors could be integrated to refine our results. Global investigation of TF binding motifs using new approaches, such as protein binding microarrays, may be beneficial in broadening the database of known TF binding motifs[47]. Moreover, TFs function with cofactors to regulate specific gene expression; co-binding analyses could be incorporated to improve our network construction[48]. Furthermore, the assignment of enhancers to the nearest genes is a limited heuristic, and being able to better associate long-range enhancers to gene targets would enhance the power of our approach considerably. Recent studies have shown the interaction of enhancers and promoters is confined in topologically associated domains[49], thus exploration of chromatin organization of enhancer marks as well as using new computational methods will facilitate the assignment of enhancers to their targets[50]. Here we provide evidence for involvement of many TFs previously overlooked in $CD8^+$ T cell immunity; our future studies will aim to refine and resolve the transcriptional networks by incorporating these additional approaches.

ONLINE METHODS

Mice, cell transfer, infection, and drug treatment

All mice were maintained in specific pathogen-free conditions according to the instructions of Institutional Animal Care and Use Committee (IACUC) of the University of California, San Diego (UCSD). OT-I TCR transgenic (specific for $OVA_{257–264}$)-MHC H2-K^b), *Tbx21*$^{-/-}$, CD45.1 congenic, and C57BL/6J mice were either bred at UCSD or received from The Jackson Laboratory. We transferred 5×10^3 OT-I TCR-transgenic $CD8^+$ T cells into congenically distinct mice by i.v. injection and then infected i.v. with 5×10^3 cfu *L. monocytogenes* expressing OVA (Lm-OVA) one day later. For T-bet-deficient experiments, we co-transferred 1×10^4 *Tbx21*$^{+/+}$ OT-I and *Tbx21*$^{-/-}$ OT-I $CD8^+$ T cells into host mice and then i.v. infected with 5×10^3 cfu Lm-OVA. For drug treatment, dexamethasone (Sigma-Aldrich) was dissolved in DMSO and diluted in PBS and then administrated to mice by i.p. injection at 10 mg/kg daily after i.v. infection with 5×10^3 cfu Lm-OVA.

Antibodies and flow cytometry

KLRG1 (2F1), CD127 (A7R34), CD8 (53-6.7), CD45.1 (A20-1.7), CD45.2 (104), CXCR3 (CXCR3-173), CD27 (LG-7F9), T-bet (4B10), Bcl-6 (K112-91), were purchased from eBioscience. FOXO1 (C29H4), TCF-1 (C63D9), IFN-γ (XMG1.2), TNF (MP6-XT22) were from Cell Signaling Technology. Antibodies for ChIP-seq, H3K4me3 (Ab8580), H3K4me1 (Ab8895) and H3K27ac (Ab4729) were from Abcam. H3K27me3 (07-449) was from Millipore. For intracellular staining of cytokines, splenocytes were *in vitro* restimulated with 1 μg/ml OVA peptide (SIINFEKL) with Protein Transport Inhibitor (eBioscience) for 4 h and then fixed and permeabilized using BD cytofix/cytoperm kit (BD Biosciences). Foxp3-transcription factor staining buffer kit (eBioscience) were used for intracellular staining of transcription factors. For intracellular staining of shRNA-transduced cells containing Ametrine-reporter, cells were fixed using freshly made 2% formaldehyde for 45 min on ice and then permeabilized. All flow cytometry data were acquired by BD LSRFortessa X-20 and all cell sorting was performed on BD FACS Aria.

shRNA knockdown by retroviral transduction

The detailed protocol was described previously[51]. PLAT-E cells were transfected with shRNAmir using TransIT-LT1 Reagent (Mirus). Retrovirus-containing supernatant was harvested after 48 h and mixed with 2-mercaptoethanol and polybrene (Millipore) for subsequent transductions. Purified naive OT-I $CD8^+$ T cells were *in vitro* activated by anti-CD3 (145-2C11) and anti-CD28 (37.51) (eBioscience) for at least 18 h and then spinfected at 805 × g with retrovirus for 1 h at 37 °C. After 4 h incubation, the retrovirus-containing medium was replaced by T cell medium. Transduction efficiency were measured by flow cytometric analysis of ametrine-reporter after 24 h and 1×10^4 shRNA transduced cells were transferred into host mice followed by Lm-OVA infection. For *Ncor1* shRNA knockdown, purified P14 $CD8^+$ T cells were *in vitro* activated and transduced by shRNA retrovirus similarly to OT-I $CD8^+$ T cells. Transduced P14 $CD8^+$ T cells (5×10^5) were transferred into host mice followed by 1.5×10^5 pfu LCMV-C13 i.p. infection, which results in an acute infection[51]. The full hairpin sequence for shRNA: sh *Yy1*: 5′-TGCTGTTGACAGTGAGCGCCCTCCTGATTATTCTGAATAATAGTGAAGCCACAGAT

GTATTATTCAGAATAATCAGGAGGTTGCCTACTGCCTCGGA-3′, sh*Nr3c1*: 5′-TGCTGTTGACAGTGAGCGAATGCATGATGTGGTTGAAAAATAGTGAAGCCACAGATGTATTTTTCAACCACATCATGCATGTGCCTACTGCCTCGGA-3′.

RT-PCR and qPCR

For RT-PCR, RNA was extracted using Trizol (Life Technologies) followed by precipitation of isopropanol. CDNA was synthesized using Superscript II kit (Life Technologies) following the manufacturer's instruction. For qPCR, cDNA was quantitatively amplified using Stratagene Brilliant II Syber Green master mix (Agilent Technologies). The abundance of transcripts was normalized to housekeeping gene *Hprt*. The following primers were used: *Zeb2* forward: 5′-CATGAACCCATTTAGTGCCA-3′, *Zeb2* reverse: 5′-AGCAAGTCTCCCTGAAATCC-3′, *Bcl2* forward: 5′-ACTTCGCAGAGATGTCCAGTCA-3′, *Bcl2* reverse: 5′-TGGCAAAGCGTCCCCTC-3′, *Gzma* forward: 5′-TGCTGCCCACTGTAACGTG-3′, *Gzma* reverse: 5′-GGTAGGTGAAGGATAGCCACAT-3′, *Klrb1c* forward: 5′-GACACAGCAAGTATCTACCT-3′, *Klrb1c* reverse: 5′-TACTAAGACTCGCACTAAGAC-3′, *Pou6f1* forward: 5′-GTCAGATCCTCACGAATGCTC-3′, *Pou6f1* reverse: 5′-GAGTCACGGCTTGGACCTG-3′, *Crtam* forward: 5′-CCTTTTCATCATCGTTCAGCTCT-3′, *Crtam* reverse: GGAGCCTGGCTGCTATTCTC-3′, *Yy1* forward: 5′-CATGTGGTCCTCGGATGAAA-3′, *Yy1* reverse: 5′-GGGAGTTTCTTGCCTGTCATA-3′, *Nr3c1* forward: 5′-CCGGGTCCCCAGGTAAAGA-3′, *Nr3c1* reverse: 5′-TGTCCGGTAAAATAAGAGGCTTG-3′, *Hprt* forward: 5′-GGCCAGACTTTGTTGGATTT-3′, *Hprt* reverse: 5′-CAACTTGCGCTCATCTTAGG-3′

Microarray analysis

The protocol was described previously[17]. KLRG1^{hi}IL-7R^{lo} TE and KLRG1^{lo}IL-7R^{hi} MP CD8^{+} T cells (2×10^4) were sorted into TRIzol on day 8 of Lm-OVA infection. RNA was amplified and labeled with biotin followed by hybridized to Affymetrix Mouse Gene ST 1.0 microarrays (Affymetrix). Microarray analysis was performed using GenePattern Multiplot Studio module. All data was generated in collaboration with the Immgen project (www.immgen.org) and passed ImmGen quality control pipeline. The gene-expression data of naive and memory CD8^{+} T cells were used from our previous study[17] and normalized with gene-expression data of TE and MP subsets by RMA normalization. Considering that the TE and MP subsets are highly similar "effector" populations on day 8 of infection and the fact that no genes show significant difference under the 1% false discovery rate using the Student's *t*-test, we used a 1.5-fold change cutoff to identify differentially expressed genes in TE and MP subsets.

Chromatin Immunoprecipitaion (ChIP), ChIP-seq library construction and sequence alignment

Cells were fixed in 1% formaldehyde for 10 min and then quenched with 125 mM glycine for 5 min. Cells were lysed for 5 min on ice and sonicated to generate 200–500 bp fragments using Bioruptor sonicator (Diagenode). Sonicated DNA was used as input control.

Magnetic-dynabeads (30 μl) were washed with blocking buffer twice and then mixed with 5 μg antibody in 500 μl blocking buffer and rotated at 4 °C. The sonicated lysates were first diluted to a final 0.1% SDS concentration. The diluted lysates were added to antibody-conjugated Dynabeads incubated at 4 °C. Beads were washed by Wash Buffer I, II and III for 5 min and then washed twice by TE buffer for 5 min. The beads were resuspended in 200 μl Elution Buffer and reverse-crosslinked at 65 °C overnight and then treated with RNAse for 30 min at 37 °C and Proteinase K at 55 °C for 1 h. DNA was purified by Zymo DNA Clean & Concentrator kit (Zymo Research). The purified DNA was end-repaired using End-it End-repair kit (Epicentre) and then added an "A" base to the 3′ end of DNA fragments using Klenow (NEB). Then DNA was ligated with adaptors using quick DNA ligase (NEB) at 25 °C for 15 min followed by size selection of 200–400 bp using AMPure SPRI beads (Beckman Coulter). The adaptor ligated DNA was amplified using NEBNext High-Fidelity 2X PCR master mix (NEB). To prevent PCR overamplification, 1 μl DNA was first quantitatively amplified using Syber Green I master mix to determine the best amplification cycle. Then the amplified library was size-selected as 200–400 bp using SPRI beads and quantified by Qubit dsDNA HS assay kit (ThermoFisher). Finally, the library was sequenced using Hiseq 2500 for single-end 50 bp sequencing to get around 20 million reads for each sample. We used BWA to map raw reads to the *Mus musculus* genome (mm10) with following parameters: "−q 5 −l 32 −k 2"[52]. Reads with low quality (MAPQ < 30) were filtered out. If multiple reads were mapped to the same location, only one read was kept.

ATAC-seq and peak calling

Cells were sorted (2.5×10^4) into 1 ml FACS buffer and spun down 500 × g for 20 min at 4°C. The cell pellet was resuspended in 25 μl lysis buffer and then spun down 600 × g for 30 min at 4 °C. The nuclei pellet was resuspended into 25 μl transposition reaction mixture containing Tn5 transposase from Nextera DNA Sample Prep Kit (Illumina) and incubated at 37 °C for 30 min. Then the transposase-associated DNA was purified using Zymo DNA clean-up kit. To amplify the library, the DNA was first amplified for 5 cycles using indexing primer from Nextera kit and NEBNext High-Fidelity 2X PCR master mix. To reduce the PCR amplification bias, 5 μl of amplified DNA after the first 5 cycles was used to do qPCR of 20 cycles to decide the number of cycles for the second round of PCR. Usually the maximum cycle of the second round of PCR is 5 cycles. Then the total amplified DNA was size selected to fragments less than 800 bp using SPRI beads. Quantification of the ATAC-seq library was based on KAPA library quantification kit (KAPAbiosystems). The size of the pooled library was examined by TapeStation. Finally, the library was sequenced using Hiseq 2500 for single-end 50 bp sequencing to get at least 10 million reads. To obtain confident peaks, we performed each ATAC-seq experiment at least twice and used the Irreproducibility Discovery Rate (IDR) framework to identify the reproducible peaks. In particular, we called peaks for each individual replicate as well as the pooled data from the two replicates using MACS2 with a relaxed threshold (*P*-value 0.01)[53]. These 3 sets of peaks were input for IDR analysis using a threshold of 0.05 to identify the confident set of peaks.

Predicting enhancers and putative TF binding sites

Enhancers were predicted by the RFECS algorithm using 3 histone marks (H3K4me1, H3K4me3 and H3K27ac). The RFECS model was trained on the active and distal P300

ChIP-seq peaks (at least 2 kb away from any TSS), which were taken as the representatives of enhancers (the positive set). For the non-enhancer class (the negative set), we chose promoters overlapping with DNase I hypersensitivity (DHS) peaks, and random 100-bp bins that are distal (2 kb away) to any P300 site and TSS. The data sets for model training were downloaded from ENCODE with following accession numbers: ENCSR000CCD (P300), ENCSR000CBF (H3K4me1), ENCSR000CBG (H3K4me3), ENCSR000CDE (H3K27ac) and DNase-seq (ENCSR000CMW)[9]. The trained model was used to scan the whole genome except the 2000 bp upstream of TSS and 500 bp downstream of TSS and classify each 100-bp bin as an enhancer or non-enhancer based on the histone modification pattern. To further reduce the false positives, we filtered the predicted enhancers using a false discover rate (FDR) of 1%. To identify putative binding sites of TFs, we first collected 761 unique motifs from two TF motif databases (JASPAR and UniPROBE) and one resource paper [54–56]. We then searched for TF binding sites in 150 bp regions centered around the ATAC-seq peak summits, using the algorithm described previously[57] with a *P*-value cutoff of 1×10^{-5}.

Motif enrichment analysis at open chromatin regions

To compute the enrichment of a TF motif over cell-type-specific open chromatin regions, we first identified the number of regions that contain at least one motif, denoted by m. Let N be the number of all regions, then $\frac{m}{N}$ is considered as the enrichment score of the query motif. To construct the null model for *P*-value calculation, we randomly selected 10,000 regions from all open chromatin sites and computed the fraction of those regions, denoted by p, containing at least one occurrence of the motif. The *P*-value for enrichment or depletion is then computed using the binomial test with p as the population proportion of null hypothesis.

Constructing TF regulatory networks

We selected active promoters as the 5-kb-regions around TSS (4 kb upstream and 1 kb downstream) that are marked by H3K4me3 peaks. Enhancers were predicted using the RFECS method based on enhancer-associated histone modification signatures. Enhancers were linked to the nearest genes. We connected a TF to a gene if the TF had any predicted binding site in the gene's promoter or linked enhancers. We assembled all the regulatory interactions between TFs and genes into a genetic network.

Personalized PageRank

The personalized PageRank algorithm measures global influence of each node in a network, used by Google and many other companies to order search engine results[37]. In an internet network, nodes are web pages and edges are links between websites. The PageRank algorithm was designed to find out how likely a specific web page is visited if web surfers who start on a random page sampled from a given distribution have α probability of choosing a random link from the page they are currently visiting and $1-\alpha$ probability of jumping to a random page chosen from all web pages. PageRank is the stationary distribution of a random walk which, at each step, with a certain probability α jumps to a random node, and with probability $1-\alpha$ follows a randomly chosen outgoing edge from the current node. Personalized PageRank is an extension of PageRank in which all the jumps are made by a pre-defined probability distribution[58]. To give a formal definition, let $G = (V, E)$

denote a directed graph, where V is a set of nodes and E contains a directed edge $\langle u, v \rangle$ if and only if node u links to node v. Let A be the transition matrix. We define $A_{ij}=\frac{1}{O(j)}$ if node j links to node i, and $A_{ij} = 0$ otherwise, where $O(j)$ is the out-degree of node j. Given a seed vector s, the personalized PageRank vector v is calculated by

$$v=(1-\alpha)Av+\alpha s$$

In a TF regulatory network, we set the weight of each gene to e^{z_i}, where z_i is the z-score of expression levels of gene i under different conditions or in different cell states. The weights of genes are then normalized and used as the seed vector for computing personalized PageRank.

For comparison between PageRank and TF activity (TFA) metric, TFA is a measurement of the activities of TFs[39], computed from the gene regulatory network (GRN) and genes' expression levels. Mathematically, TFA is defined by the following equation:

$$X_i=\sum_{k\in TFs} P_{i,k}A_k$$

where P is a matrix representing GRN, X is a vector containing the gene expression levels, and A is the TFA vector. The above equation can be written in matrix notation: $X = PA$, and TFA vector A can be solved by computing the pseudoinverse of matrix P: $A = P^{-1}X$.

To compare the performance of PageRank and TFA on predicting driver TFs, we used the gene expression profile and GRN as the input data to run both algorithm in each cell types. The gold standard is a set of 16 TFs that have known roles in TN, TE, MP and memory cells. 14 of them were identified from literature, and 2 of them were confirmed by experiments in this study. In Supplementary Figure 6, we showed PageRank successfully retrieved 12 out of 16 (75%) TFs whose pattern were consistent with previous reports, demonstrating a clear advantage over TFA metric.

Code Availability

The PageRank analysis pipeline can be downloaded at http://kzhang.org/Taiji/.

RNA-seq and GSEA analysis

For RNA-seq of TE and MP subset, we sorted 1×10^3 donor TE and MP $CD8^+$ T cells on day 8 of Lm-OVA infection. Isolation of $polyA^+$ RNA, RNA-Seq library preparation and RNA-seq analysis were performed as described in www.immgen.com/Protocols/11cells.pdf. The processed RNA-seq data was shown in supplementary table 1. Gene Set Enrichment Ananlysis (GSEA) was performed by GSEA module in GenePattern. TE- and MP-enriched gene signature were generated with 2-fold change cutoff by Multiplot module in GenePattern. Normalized enrichment score (NES) and FDR q values were calculated by permutation test.

Statistical analysis

All ChIP-seq and ATAC-seq were performed independently in two replicates. Microarray and RNA-seq were performed independently in three replicates except two replicates for TE $CD8^+$ T cells in microarray. The two tailed unpaired Student's *t*-test, two tailed paired Student's *t*-test, binomial test and Wilcoxon rank-sum test were applied as indicated. Statistical analysis for animal studies were performed using GraphPad Prism software.

Data Availability

The microarray, ATAC-seq and ChIP-seq data are available in the GEO database with the accession codes GSE89036 (ATAC-seq and ChIP-seq) and GSE89037 (microarray). Other public available datasets are described in subsections above. The source data published as supplementary items and data support the findings of this study are available from the corresponding author upon request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding of this work provided by UCSD NIH Cell and Molecular Genetics Training Grant (5T32GM007240-36) and Dr. Huang Memorial Scholarship for B.Y.; the National Institutes of Health (AI067545, A1072117 to A.W.G), (U19AI109976 to A.W.G, S.C., M.E.P.), (U54HG006997, AR070310 to W.W.); the Leukemia and Lymphoma Society and Pew Scholars Fund to A.W.G.; Pew Latin American Fellows Program in the Biomedical Sciences to R.M.P.; the Fraternal Order of Eagles Fellow of the Damon Runyon Cancer Research Foundation to J.P.S-B. The research performed by R.M.P. and J.P.S-B. was funded by NIH grants R01 AI109842 and AI40127 (to Anjana Rao). We thank the Immgen core team for help of gene-expression data processing. We thank C. Murre, A. Phan, K. Omilusik, L.A.Shaw, K. Brennan, and Goldrath lab members for critical discussions and review of the manuscript.

References

1. Ahmed R, Gray D. Immunological memory and protective immunity: understanding their relation. Science. 1996; 272:54–60. [PubMed: 8600537]
2. Joshi NS, et al. Inflammation Directs Memory Precursor and Short-Lived Effector CD8+ T Cell Fates via the Graded Expression of T-bet Transcription Factor. Immunity. 2007; 27:281–295. [PubMed: 17723218]
3. Kaech SM, et al. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. Nat Immunol. 2003; 4:1191–1198. [PubMed: 14625547]
4. Zhou X, et al. Differentiation and Persistence of Memory CD8(+) T Cells Depend on T Cell Factor 1. Immunity. 2010; 33:229–240. [PubMed: 20727791]
5. Chang JT, Wherry EJ, Goldrath AW. Molecular regulation of effector and memory T cell differentiation. Nat Rev Immunol. 2014; 15:1104–1115.
6. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013; 10:1213–1218. [PubMed: 24097267]
7. Winter D, et al. Making the case for chromatin profiling: a new tool to investigate the immune-regulatory landscape. Nat Rev Immunol. 2015; 15:585–94. [PubMed: 26272294]
8. Neph S, et al. Circuitry and dynamics of human transcription factor regulatory networks. Cell. 2010; 150:1274–1286.
9. Shen Y, et al. A map of the cis-regulatory sequences in the mouse genome. Nature. 2012; 488:116–120. [PubMed: 22763441]

10. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012; 13:613–626. [PubMed: 22868264]
11. Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? Mol Cell. 2013; 49:825–837. [PubMed: 23473601]
12. Lara-Astiaso D, et al. Chromatin state dynamics during blood formation. Science. 2014; 345:943–949. [PubMed: 25103404]
13. Lavin Y, et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell. 2014; 159:1312–1326. [PubMed: 25480296]
14. 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. 2009; 30:912–925. [PubMed: 19523850]
15. Russ BE, et al. Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8+ T cell differentiation. Immunity. 2014; 41:853–865. [PubMed: 25517617]
16. Heintzman ND, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009; 459:108–112. [PubMed: 19295514]
17. Best JA, et al. Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. Nat Immunol. 2013; 29:997–1003.
18. Rubinstein M, et al. IL-7 and IL-15 differentially regulate CD8+ T-cell subsets during contraction of the immune response. Blood. 2008; 112:3704–3712. [PubMed: 18689546]
19. Yang CY, et al. The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. Nat Immunol. 2011; 12:1221–1229. [PubMed: 22057289]
20. Miyazaki M, et al. The opposing roles of E2A and Id3 that orchestrate and enforce the naïve T cell fate. Nat Immunol. 2012; 12:992–1001.
21. Rajagopal N, et al. RFECS: A Random-Forest Based Algorithm for Enhancer Identification from Chromatin State. PLoS Comput Biol. 2013; 9
22. Chaix J, et al. Cutting edge: CXCR4 is critical for CD8+ memory T cell homeostatic self-renewal but not rechallenge self-renewal. J Immunol. 2014; 193:1013–1016. [PubMed: 24973450]
23. McLean CY, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501. [PubMed: 20436461]
24. Ma C, Zhang N. Transforming growth factor-β signaling is constantly shaping memory T-cell population. Proc Natl Acad Sci USA. 2015; 112:11013–11017. [PubMed: 26283373]
25. Omilusik KD, et al. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. J Exp Med. 2015; 212:2027–2039. [PubMed: 26503445]
26. Dominguez CX, et al. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. J Exp Med. 2015; 212:2041–2056. [PubMed: 26503446]
27. Intlekofer AM, et al. Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. Nat Immunol. 2005; 6:1236–1244. [PubMed: 16273099]
28. Kidani Y, et al. Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity. Nat Immunol. 2013; 14:489–499. [PubMed: 23563690]
29. Kurachi M, et al. The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells. Nat Immunol. 2014; 15:373–383. [PubMed: 24584090]
30. Zhou X, Xue H. Generation of memory precursors and functional memory CD8+ T cells depends on TCF-1 and LEF-1. J Immunol. 2012; 189:2722–2726. [PubMed: 22875805]
31. D'Cruz LM, Lind KC, Wu BB, Fujimoto JK, Goldrath AW. Loss of E protein transcription factors E2A and HEB delays memory-precursor formation during the CD8 + T-cell immune response. Eur J Immunol. 2012; 42:2031–2041. [PubMed: 22585759]
32. Rincón M, Flavell Ra. AP-1 transcriptional activity requires both T-cell receptor-mediated and co-stimulatory signals in primary T lymphocytes. EMBO J. 1994; 13:4370–4381. [PubMed: 7925281]
33. Doering T, et al. Network Analysis Reveals Centrally Connected Genes and Pathways Involved in CD8+ T Cell Exhaustion versus Memory. Immunity. 2012; 37:1130–1144. [PubMed: 23159438]

34. Hu G, Chen J. A genome-wide regulatory network identifies key transcription factors for memory $CD8^+$ T-cell development. Nat Commun. 2013; 4:2830. [PubMed: 24335726]
35. Szabo SJ, et al. Distinct Effects of T-bet in TH1 Lineage Commitment and IFN-γ Production in CD4 and CD8 T Cells. Science. 2002; 295:338–342. [PubMed: 11786644]
36. Lord GM, et al. T-bet is required for optimal proinflammatory CD4+ T-cell trafficking. Blood. 2005; 106:3432–3439. [PubMed: 16014561]
37. Page L, Brin S, Motwani R, Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. WORLD WIDE WEB. 1998; 54:1–17.
38. Cui W, Liu Y, Weinstein JS, Craft J, Kaech SM. An interleukin-21- interleukin-10-STAT3 pathway is critical for functional maturation of memory CD8+ T cells. Immunity. 2011; 35:792–805. [PubMed: 22118527]
39. Arrieta-Ortiz ML, et al. An experimentally supported model of the Bacillus subtilis global transcriptional regulatory network. Mol Syst Biol. 2015; 11:839. [PubMed: 26577401]
40. Hwang SS, et al. YY1 inhibits differentiation and function of regulatory T cells by blocking Foxp3 expression and activity. Nat Commun. 2016; 7:10789. [PubMed: 26892542]
41. Hwang SS, et al. Transcription factor YY1 is essential for regulation of the Th2 cytokine locus and for Th2 cell differentiation. Proc Natl Acad Sci USA. 2013; 110:276–281. [PubMed: 23248301]
42. Liu H, et al. Yin Yang 1 is a critical regulator of B-cell development. Genes Dev. 2007; 21:1179–1189. [PubMed: 17504937]
43. Herold MJ, McPherson KG, Reichardt HM. Glucocorticoids in T cell apoptosis and function. Cell Mol Life Sci. 2006; 63:60–72. [PubMed: 16314919]
44. Franchimont D, et al. Positive effects of glucocorticoids on T cell function by up-regulation of IL-7 receptor alpha. J Immunol. 2002; 168:2212–2218. [PubMed: 11859107]
45. Smoak KA, Cidlowski JA. Mechanisms of glucocorticoid receptor signaling during inflammation. Mech Ageing Dev. 2004; 125:697–706. [PubMed: 15541765]
46. Wang Q, et al. Equilibrium interactions of corepressors and coactivators with agonist and antagonist complexes of glucocorticoid receptors. Mol Endocrinol. 2004; 18:1376–1395. [PubMed: 15016838]
47. Tsankov AM, et al. Transcription factor binding dynamics during human ES cell differentiation. Nature. 2015; 518:344–349. [PubMed: 25693565]
48. Zhang K, Li N, Ainsworth R, Wang W. Systematic identification of protein combinations mediating chromatin looping. Nat Commun. 2016; 7:1–11.
49. Dixon JR, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015; 518:331–336. [PubMed: 25693564]
50. Zhu Y, et al. Constructing 3D interaction maps from 1D epigenomes. Nat Commun. 2016; 7:10812. [PubMed: 26960733]
51. Chen R, et al. In vivo RNA interference screens identify regulators of antiviral CD4+ and CD8+ T cell differentiation. Immunity. 2014; 41:325–338. [PubMed: 25148027]
52. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]
53. Zhang Y, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
54. Mathelier A, et al. JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2016; 44:D110–D115. [PubMed: 26531826]
55. Newburger DE, Bulyk ML. UniPROBE: An online database of protein binding microarray data on protein-DNA interactions. Nucleic Acids Res. 2009; 37:77–82.
56. Jolma A, et al. DNA-binding specificities of human transcription factors. Cell. 2013; 152:327–339. [PubMed: 23332764]
57. Grant CE, Bailey TL, Noble WS. FIMO: Scanning for occurrences of a given motif. Bioinformatics. 2011; 27:1017–1018. [PubMed: 21330290]
58. Jeh G, Widom J. Scaling personalized web search. Proc 2003 Int World Wide Web Conf (WWW'03). 2003; 12:271–279.

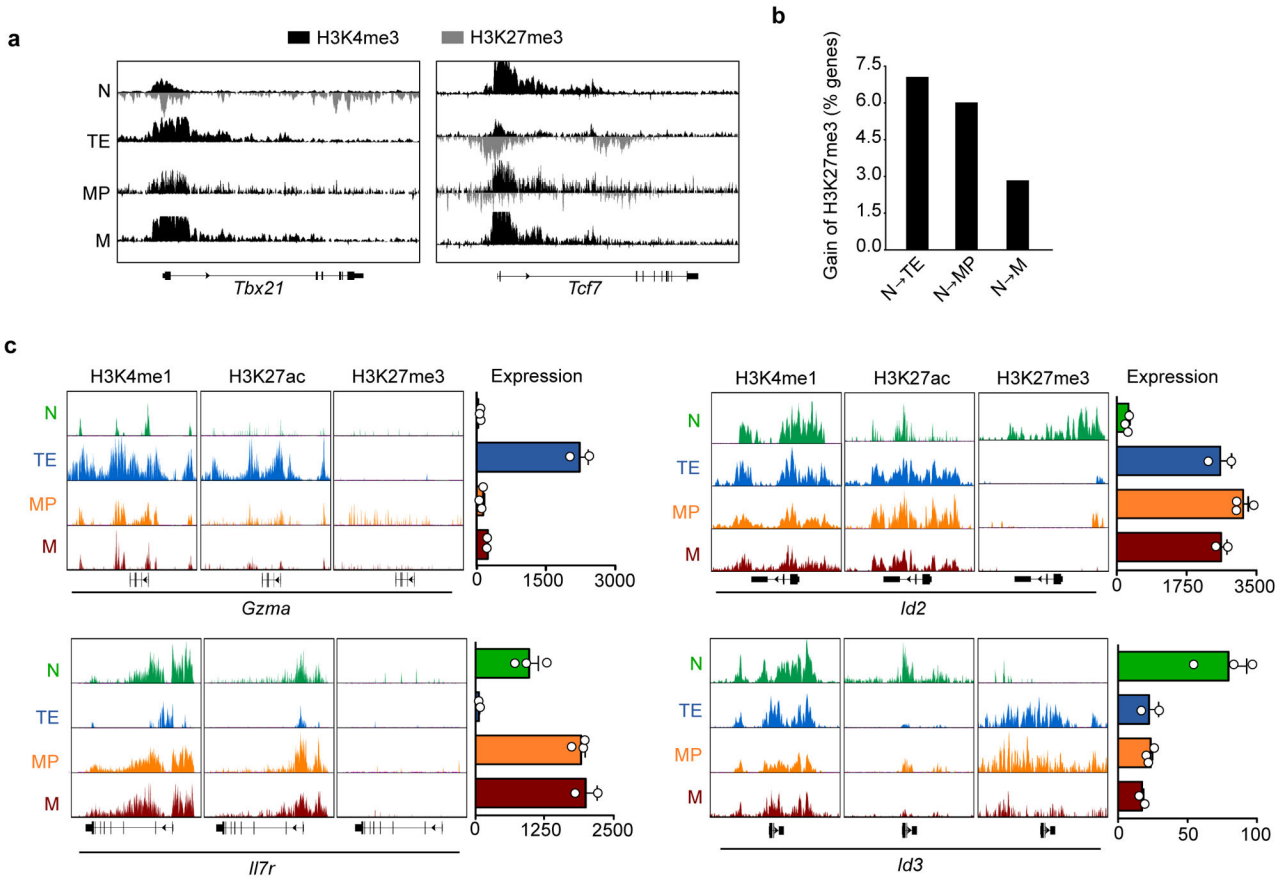

Figure 1.

Epigenetic landscapes of CD8$^+$ T cells in response to bacterial infection. (**a**) Representative genes displaying bivalent modification of H3K4me3 and H3K27me3 at promoter regions during CD8$^+$ T cell differentiation. (**b**) Comparison of the percentage of genes with increased H3K27me3 at promoter regions of genes with decreased expression upon differentiation. (**c**) Representative genes displaying dynamic change of enhancer H3K4me1 and active H3K27ac marks (left). Representative genes displaying a unique pattern of active H3K27ac and repressive H3K27me3 marks (right). Bar graphs showing the gene expression generated from microarray analysis. Data in (a,b) are representative of two independent experiments (n=10) and data in (c) are representative of three independent experiments (n=3, mean ± s.e.m.).

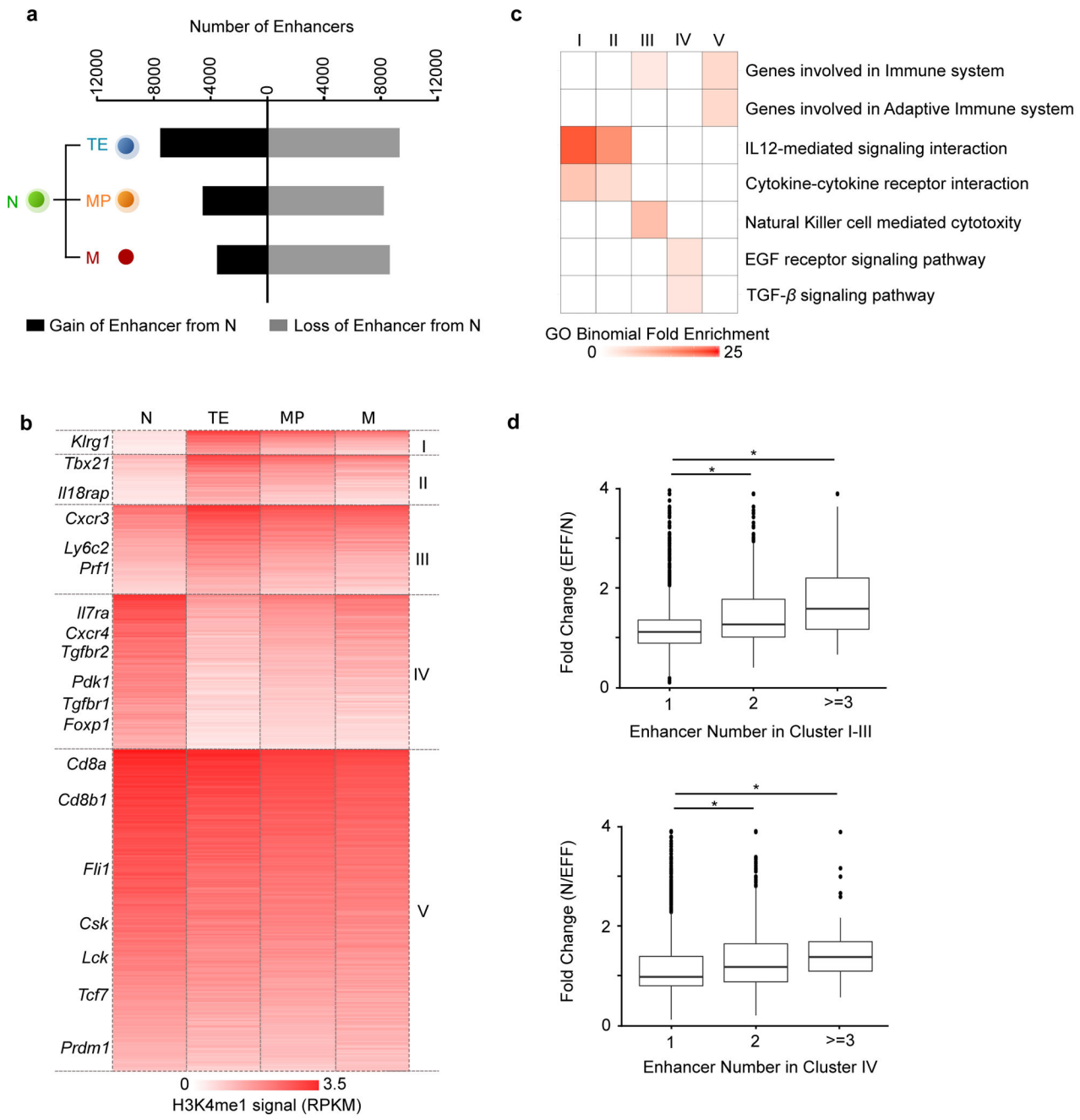

Figure 2.

Dynamic use of enhancers is associated with differentially expressed genes during CD8$^+$ T cell differentiation. (**a**) Bar plot showing the number of enhancers gained (black) and lost (gray) during differentiation from naive CD8$^+$ T cells to TE, MP and memory CD8$^+$ T cells. (**b**) Heatmap of *k*-means clustering (*k*=5) of total 52331 enhancers across CD8$^+$ T cell subsets. Bar plot showing the number of enhancers gained (black) and lost (gray) during differentiation from naive CD8$^+$ T cells to TE, MP and memory CD8$^+$ T cells. (**c**) Gene Ontology analysis of clusters in (a) using a binomial test with top 2 pathways shown (cut off as binomial *P* value < 0.001). (**d**) Box plots showing fold change of mRNA expression of

genes with the indicated number of enhancers in clusters I–III (left) and cluster IV (right) during differentiation of naive $CD8^+$ T cells to effector $CD8^+$ T cells. Data in (a–c) are representative of two independent experiments (n=10). Data in (d) are representative of three independent experiments (n=3, mean ± s.e.m.). *P* value was calculated by unpaired two tailed Student's *t*-test: *: $p<0.0001$

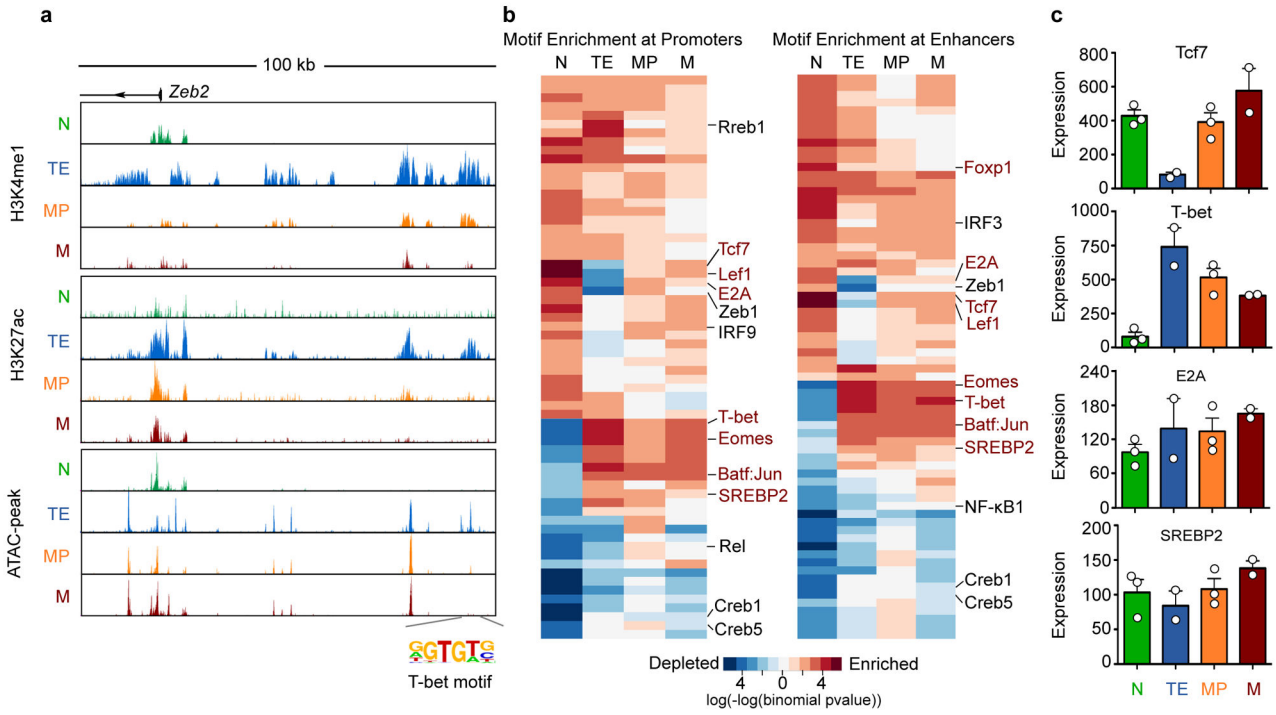

Figure 3.

Accessible regulatory regions allow prediction of TF regulators. (**a**) Representative gene displaying subset-specific accessible enhancer containing known TF motif. (**b**) Heatmap showing the *P*-value of TF motif enrichment at subset-specific enhancers or promoters calculated by binomial test using randomly-picked open chromatin regions as background. Motif enrichment or depletion are indicated as red or blue, respectively. Known TFs that are key to effector or/and memory $CD8^+$ T cell differentiation are highlighted in red. (**c**) Bar graphs showing mRNA expression of indicated TFs generated by microarray. Data in (a,b) are representative of two independent experiments (n=10). Data in (c) are representative of three independent experiments (n=3, mean ± s.e.m.).

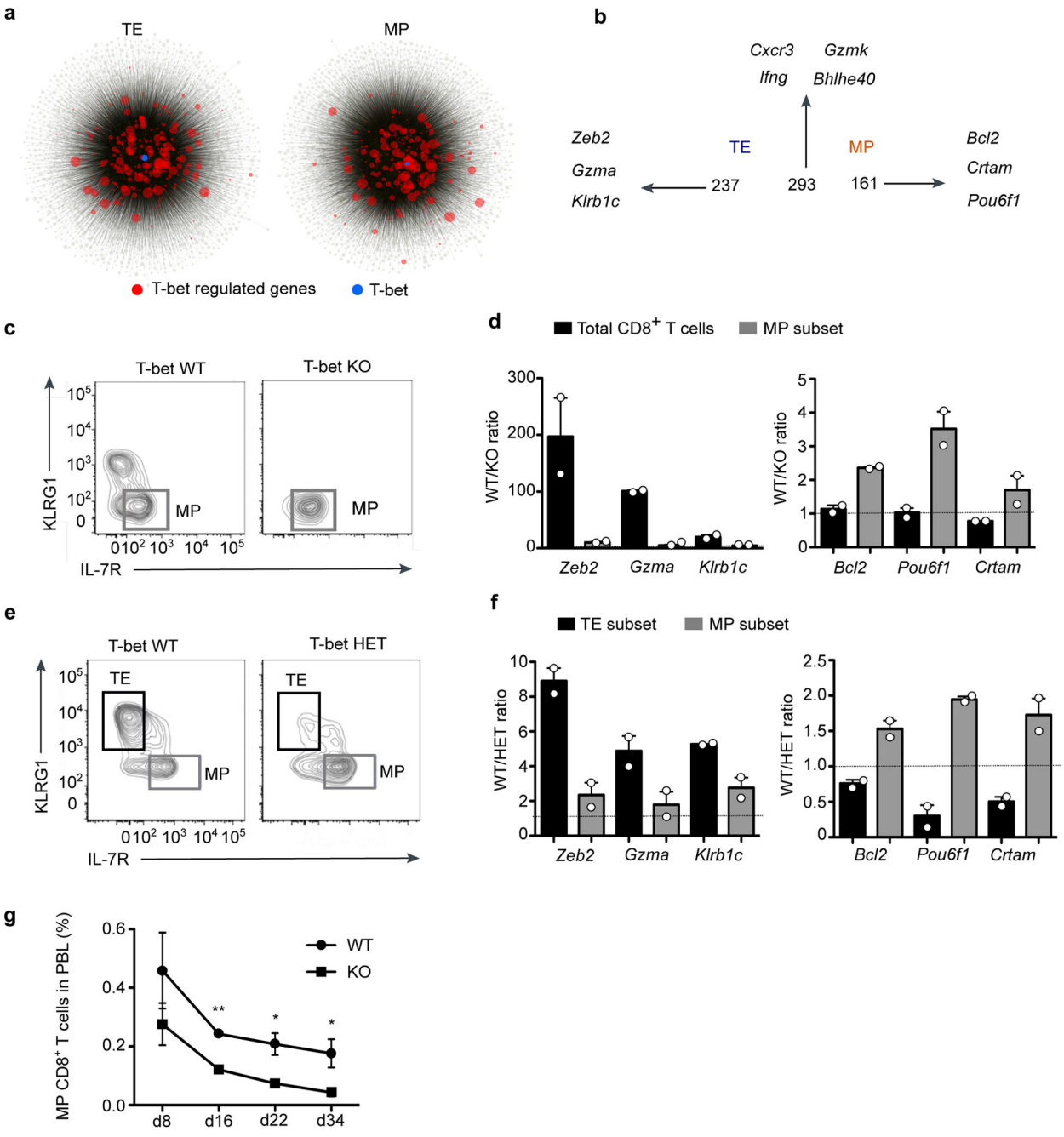

Figure 4.

Network analysis reveals subset-specific T-bet regulatory circuits. (**a**) Global regulatory network in the TE and MP subsets. T-bet regulated genes are highlighted in red; T-bet is labeled in blue. (**b**) Comparison of T-bet regulated genes between the TE and MP subsets. (**c–d**) *Tbx21*$^{+/+}$ and *Tbx21*$^{-/-}$ OT-I cells were co-transferred into recipient mice followed by Lm-OVA infection. At day 9 of infection, total CD8^{+} T cells or the MP subset of *Tbx21*$^{+/+}$ and *Tbx21*$^{-/-}$ populations were sorted to and mRNA expression levels of subset-specific T-bet regulated gene targets determined by qPCR (d). The dashed line indicates that *Tbx21*$^{+/+}$/

Tbx21$^{-/-}$ ratio=1. (**e–f**) *Tbx21*$^{+/+}$ OT-I and *Tbx21*$^{+/-}$ OT-I cells were co-transferred into recipient mice followed by Lm-OVA infection. At day 8 of infection, TE and MP subset of *Tbx21*$^{+/+}$ and *Tbx21*$^{+/-}$ were sorted to measure RNA expression of subset-specific T-bet regulated gene targets (f). The dashed line indicates that *Tbx21*$^{+/+}$/*Tbx21*$^{+/-}$ ratio=1. (**g**) Kinetic analysis of the percentage of MP of *Tbx21*$^{+/+}$ and *Tbx21*$^{-/-}$ cells during Lm-OVA infection. Data in (c–g) are representative of two independent experiments (n=3 in (c–f), n=4 in (g), mean ± s.e.m.). *P* value was calculated by paired two tailed Student's *t*-test: n.s. *: p<0.05; **: p<0.01

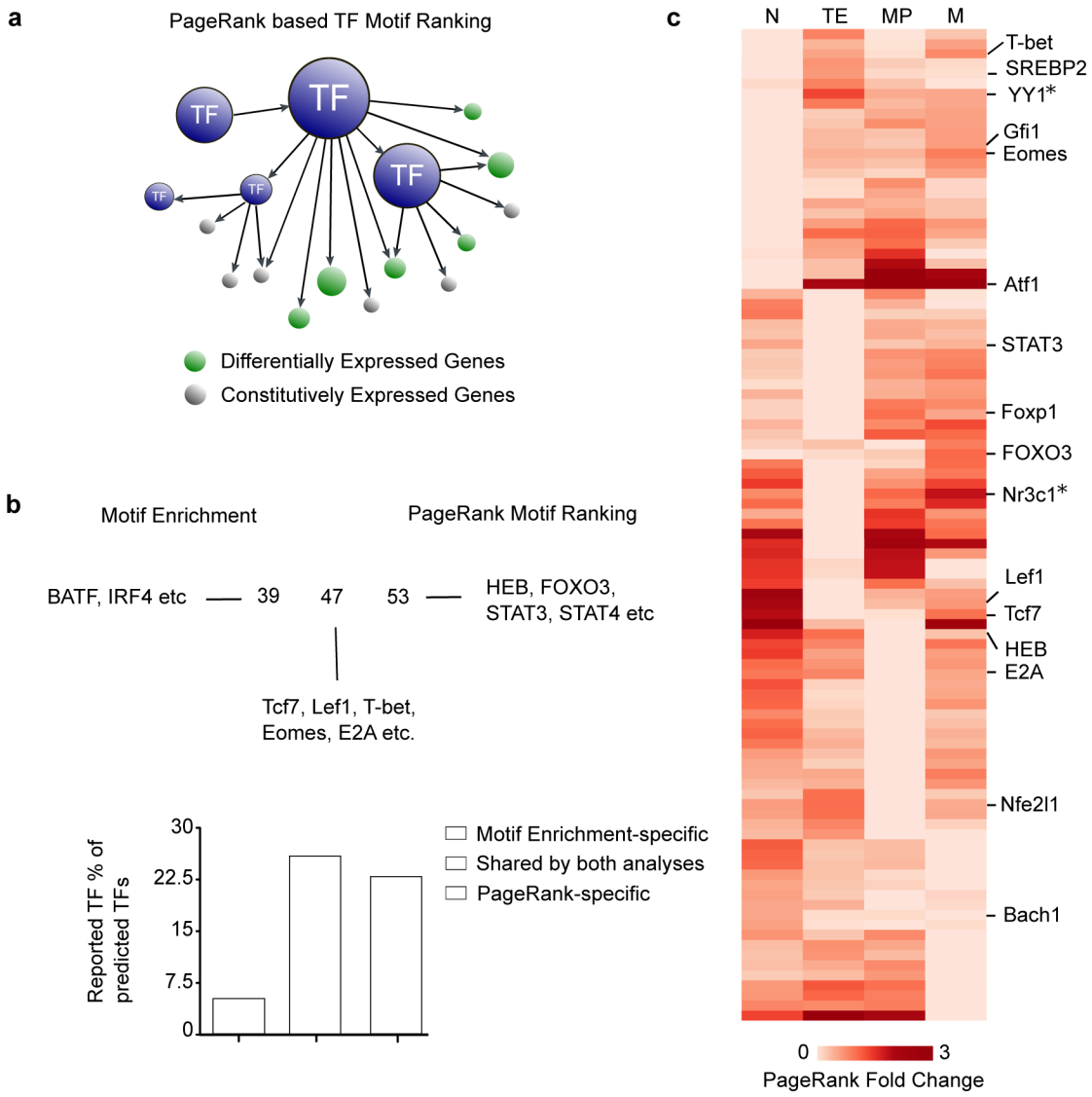

Figure 5.

PageRank-based TF ranking highlights key TF candidates. (**a**) Schematic view of PageRank-based TF motif ranking. The size of circles in the network represents the importance of gene targets which are assessed by relative expression across different cell types generated by microarray data. (**b**) Comparison of PageRank analysis with motif enrichment analysis in Figure 3. Bar graph showing the percentage of known TFs reported previously recovered from predicted TF candidates for each analysis. (**c**) Heatmap of PageRank fold enrichment of TFs across $CD8^+$ T cell subsets. Data in (c) is based on representative network generated from two independent ATAC-seq and ChIP-seq experiments (n=5 for ATAC-seq and n=10 for ChIP-seq).

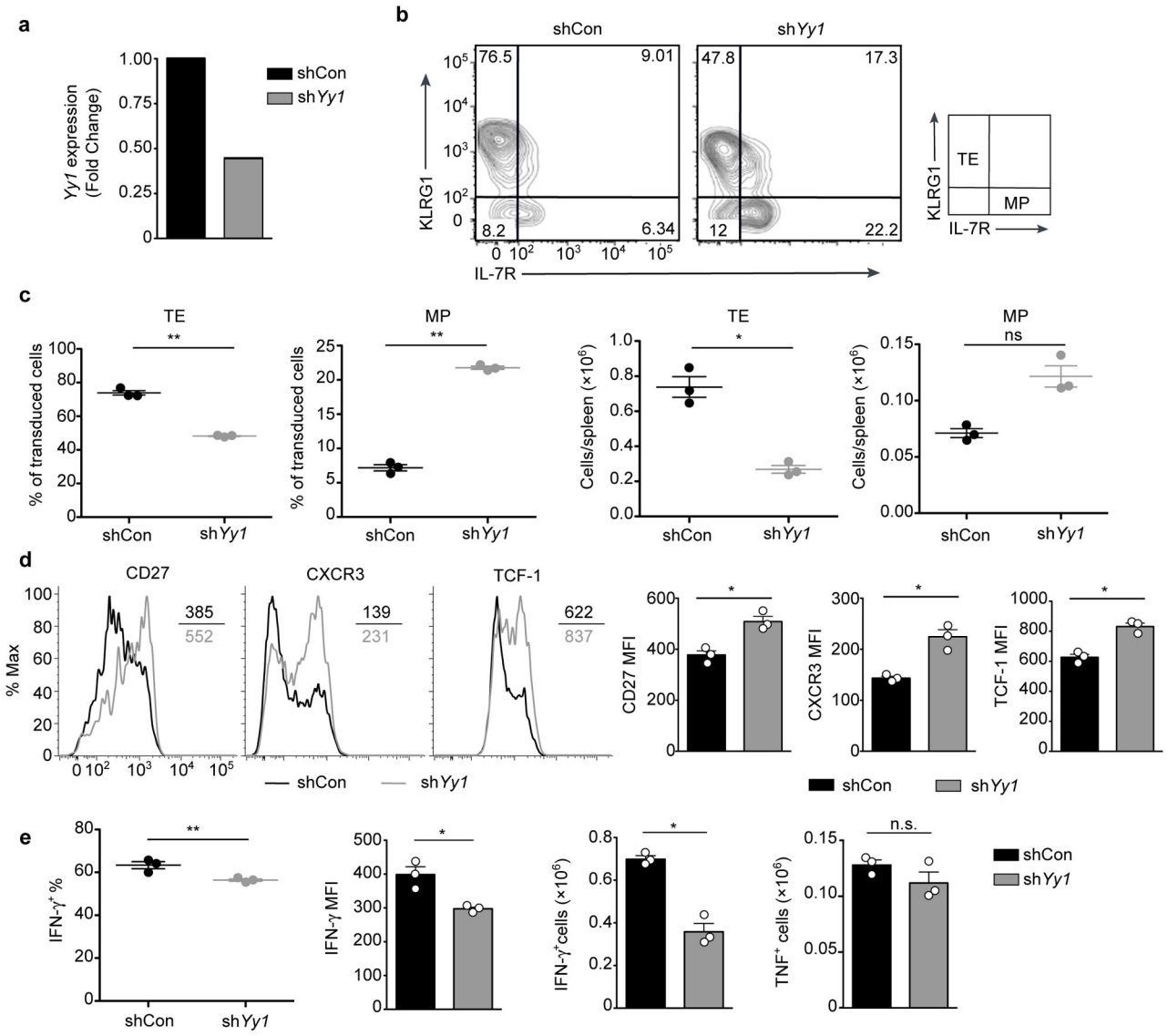

Figure 6.

YY1 is a transcriptional regulator of TE CD8$^+$ T cell differentiation. (**a**) Expression of *Yy1* mRNA quantified by RT-qPCR after *Yy1* shRNA knockdown in CD8$^+$ T cells after 72 h *in vitro* activation. (**b**) Flow cytometric analysis of KLRG1 and IL-7R expression for cells transduced with control shRNA and sh*Yy1*. (**c**) The percentage and the number of TE and MP CD8$^+$ T cells after knockdown of *Yy1*. (**d**) Histogram (left) and MFI expression (right) of CD27, CXCR3 and TCF-1 after knockdown of *Yy1*. (**e**) Flow cytometric analysis of the frequency and number of IFN-γ producing cells and IFN-γ MFI and the number of TNF producing cells using intracellular cytokine staining of splenocytes restimulated by OVA peptide for 4 h. Data are representative of two (a, d, e) or three (b, c) independent experiments (n=3, mean ± s.e.m.) For comparison of two groups, two tailed paired Student's *t*-test was performed. n.s. *: p<0.05; **: p<0.01; ***: p<0.001

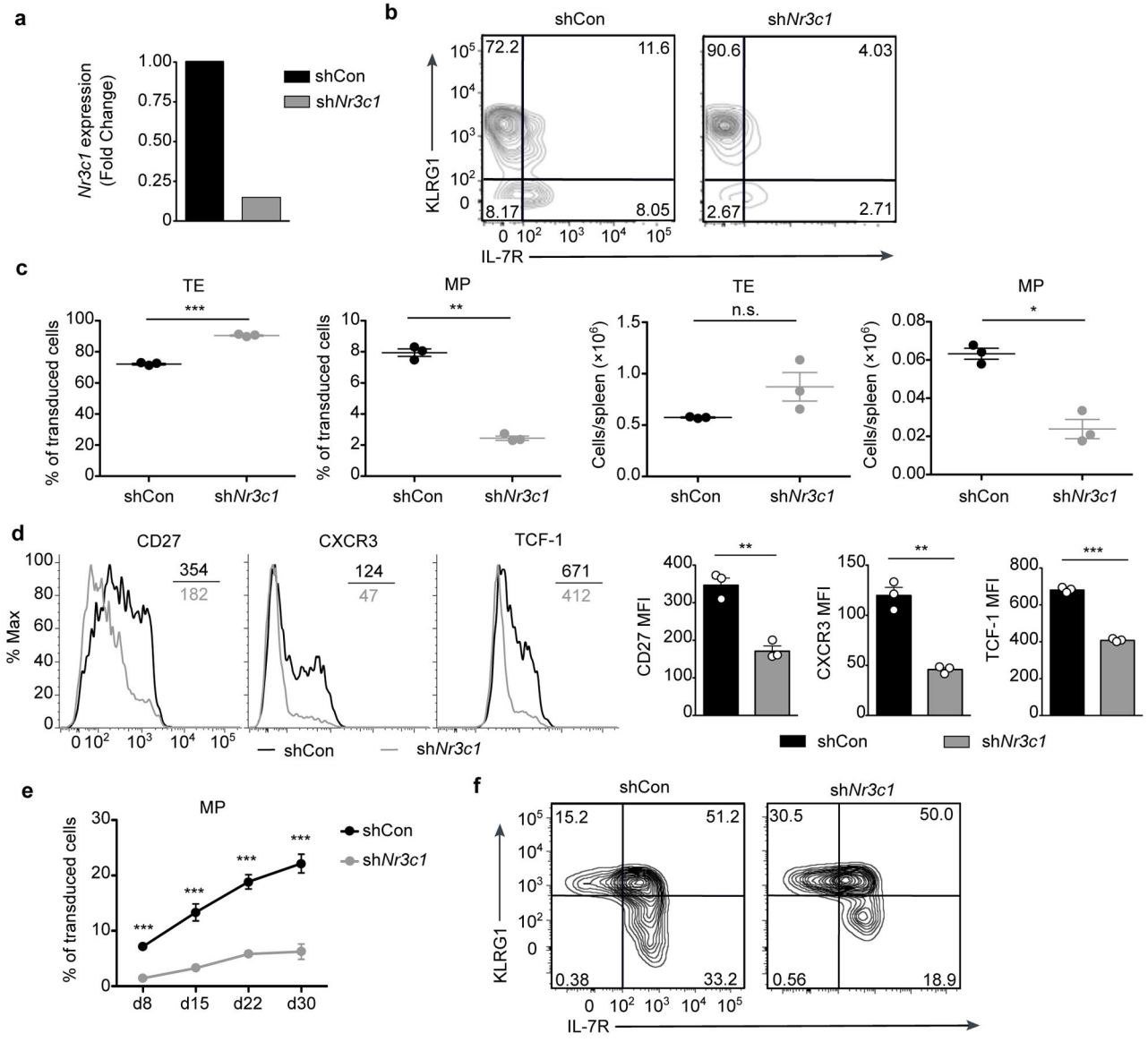

Figure 7.

Glucocorticoid receptor Nr3c1 is essential for the formation of MP CD8$^+$ T cells. (**a**) Expression of *Nr3c1* mRNA quantified by RT-qPCR after *Nr3c1* shRNA knockdown in CD8$^+$ T cells after 72 h *in vitro* activation. (**b–d**) OT-I CD8$^+$ T cells were activated *in vitro* and transduced with retrovirus containing control shRNA or shNr3c1 for 24 h, co-transfered to recipient mice followed by i.v. infection with Lm-OVA. Splenocytes were analyzed on day 7 of infection. Flow cytometric analysis of KLRG1 and IL-7R expression shown in (b) and the frequency and number of TE and MP CD8$^+$ T cells after knockdown of *Nr3c1* shown in (c). Histograms and MFI of CD27, CXCR3, and TCF-1 after knockdown of *Nr3c1* shown in (d). (**e**) Kinetic analysis of MP subset frequency after knockdown of *Nr3c1*. (**f**) Flow cytometric analysis of KLRG1 and IL-7R expression of shRNA transduced cells in the spleen on day 30 of Lm-OVA infection. Data are representative of two independent

experiments (n=3 in (a–d), n=5 in (e), mean ± s.e.m.). For comparison of two groups, two tailed paired Student's *t*-test was performed. n.s. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$