

# Batf-mediated Epigenetic Control of Effector CD8+ T Cell Differentiation

Hsiao-Wei Tsao,<sup>1, 2,\*</sup> James Kaminski,<sup>3, \*</sup> Makoto Kurachi,<sup>4</sup> R. Anthony Barnitz,<sup>1, a</sup> Michael A. Dilorio,<sup>1, b</sup> Martin W. LaFleur,<sup>1, 5, 6, 7</sup> Wataru Ise,<sup>13</sup> Tomohiko Kurosaki,<sup>13, 14</sup> E. John Wherry,<sup>8</sup> W. Nicholas Haining,<sup>1, 2, 11, c, †</sup> Nir Yosef,<sup>3, 9, 10, 12, †</sup>

1. Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA
2. Broad Institute of MIT and Harvard, Cambridge, MA, USA
3. Center for Computational Biology, University of California, Berkeley, Berkeley, CA, USA
4. Department of Molecular Genetics, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Japan
5. Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, USA
6. Division of Medical Sciences, Harvard Medical School, Boston, MA, USA
7. Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA
8. Department of Microbiology and Institute for Immunology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
9. Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Boston, MA, USA
10. Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, USA
11. Division of Pediatric Hematology and Oncology, Children's Hospital, Boston, MA, USA
12. Chan Zuckerberg Biohub, San Francisco, CA, USA
13. Laboratory of Lymphocyte Differentiation, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan
14. Laboratory for Lymphocyte Differentiation, RIKEN Center for Integrative Medical Sciences, Kanagawa, Japan
- a. Present address: TCR<sup>2</sup> Therapeutics Inc. Cambridge, MA, USA
- b. Present address: Brigham and Women's Hospital, Boston, MA, USA
- c. Present address: Merck Research Laboratories, Boston, MA, USA
- \*. These authors contributed equally
- †. These authors jointly supervised this work.

Correspondence to: [niryosef@berkeley.edu](mailto:niryosef@berkeley.edu) (N.Y.) and [nick.haining@merck.com](mailto:nick.haining@merck.com) (W.N.H)  
Lead contact: N.Y.

# SUMMARY

The response of cytotoxic T cells to their cognate antigen involves rapid and broad changes in gene expression that are coupled with extensive chromatin remodeling. Here, we study the mechanisms by which the basic leucine zipper ATF-like transcription factor Batf helps regulate this process. Through genome-scale profiling, we observe critical roles for Batf in inducing transcriptional changes in stimulated naive cells, while affecting the chromatin at several levels, namely binding of key transcription factors, accessibility, and long range contacts. We identify a critical network of transcription factors that cooperate with Batf, including its binding partner Irf4, as well as Runx3 and T-bet, and demonstrate its synergistic activity in initiating aspects of the effector T cells' transcriptional and epigenetic program in ectopically-induced fibroblasts. Our results provide a comprehensive resource for studying the epigenomic and transcriptomic landscape of effector differentiation of cytotoxic T cells.

# INTRODUCTION

When CD8<sup>+</sup> T cells encounter their cognate antigens, they embark on a process of rapid clonal expansion and differentiation, generating a large population of effector cells characterized by marked changes in gene expression. The transcriptional changes that occur during effector differentiation are associated with massive shifts in chromatin accessibility, reflecting large-scale changes in both regulation of and transcription from gene loci (Kaech and Cui, 2012; Sen et al., 2016; Wong and Pamer, 2003).

The basic leucine zipper activating transcription factor (ATF)-like transcription factor (Batf) is a well-characterized transcription factor (TF) involved in the development, differentiation and function of many immune cell subsets, including dendritic cells (DC), B cells, T helper 2 (Th2), Th9, Th17, follicular T helper (Tfh) and CD8<sup>+</sup> T cells (Betz et al., 2010; Ciofani et al., 2012; Ise et al., 2011; Jabeen et al., 2013; Kurachi et al., 2014; Kuwahara et al., 2016; Sahoo et al., 2015; Schraml et al., 2009). Batf is reported to be induced by various stimuli in T cells including TCR stimulation and cytokine signaling, and contributes to T cell function through divergent mechanisms (Kuroda et al., 2011; Man et al., 2017; Pham et al., 2019; Xin et al., 2015). In Tfh, Batf is able to recruit Ctcf in an Ets1 dependent manner to establish the chromatin landscape for initiating the earliest phase of differentiation. This mechanism is independent of interferon regulatory factor 4 (Irf4) (Pham et al., 2019). However, by cooperating with Irf4, Batf is critical for initiating the differentiation of Th17 cells, acting as so-called pioneer factors (Ciofani et al., 2012; Li et al., 2012; Murphy et al., 2013), which regulate chromatin remodeling and facilitate the binding of other factors (Zaret and Carroll, 2011). Batf and Irf4 are capable of binding a composite AP-1-IRF DNA element (AICE) binding sites, which served to explain their interdependence (Glasmacher et al., 2012; Li et al., 2012). Like Batf knockout (KO) mice, mice deficient for Irf4 exhibit defects in Tfh cell differentiation, Th17 differentiation, CD8<sup>+</sup> T cell response to LCMV infection, and B cell class switch recombination (Bollig et al., 2012; Brüstle et al., 2007; Grusdat et al., 2014; Klein et al., 2006; Mudter et al., 2011; Sciammas et al., 2006).

Previously, we demonstrated that Batf is also necessary for the earliest differentiation and expansion of CD8<sup>+</sup> T cells when they encounter an antigen *in vivo* (Kurachi et al., 2014). Knockout of *Batf* in CD8<sup>+</sup> T cells results in proliferation defects, changes in mRNA expression of other transcription factors and abnormal cytokine production. Furthermore, chromatin immunoprecipitation followed by sequencing (ChIP-seq) of Batf in *in vitro* differentiated effector

CD8<sup>+</sup> T cells suggested that it may colocalize with other transcription factors and possibly facilitate their binding. Although Batf has been shown to be able to initiate the CD4<sup>+</sup> T cell chromatin reorganization and differentiation in the Tfh-prone environment (Pham et al., 2019), how Batf establishes the cell identity and effector function in CD8<sup>+</sup> background has not been described in a systematic way. Furthermore, the mechanisms by which Batf coordinates the epigenomic and transcriptomic events during CD8<sup>+</sup> T cell activation and differentiation, how Batf synchronizes with other transcription factors, and the extent to which Irf4 is required for these aspects of BATF function, are still unclear.

In order to explore these questions, we utilized multiple types of genome-scale assays to map the effects of Batf deficiency on the transcriptome and epigenome of differentiating and effector CD8<sup>+</sup> T cells. We first evaluated the relevant transcriptional changes using RNA-seq, comparing Batf deficient to normal cells. We then investigated the respective chromatin-level changes using assay for transposase accessible chromatin sequencing (ATAC-seq) and ChIP-seq of a set of critical TFs to produce a dataset of 96,490 genomic regions with accessible chromatin or bound TFs, and identified long-range contacts among these regions using Hi-C chromatin immunoprecipitation (HiChIP). Through systematic analysis of these transcriptional and epigenetic datasets, we identified a Batf-dependent transcriptional program that establishes effector differentiation. This program consisted of major changes in chromatin accessibility and three dimensional structure, binding of key TFs, and corresponding changes in gene expression. We identified a subset of key TFs that are co-localized and affected by Batf in T cells (Irf4, Runx3, and T-bet). We then demonstrated that ectopic expression of these TFs along with Batf in fibroblasts, but not Batf alone, is sufficient to reproduce aspects of the epigenetic (chromatin accessibility and loops) and transcriptional characteristics of CD8<sup>+</sup> effector T cells.

Taken together, our results not only map the binding of critical TFs but also provide a comprehensive resource for studying the epigenome of differentiating CD8<sup>+</sup> T cells. They also yield an in-depth view of the manner by which Batf plays its role in directing effector CD8<sup>+</sup> T cell differentiation after engagement with antigen, ultimately reaching proper cytotoxic T cell response against pathogens.



## RESULTS

### Loss of Batf has broad effects on gene expression and chromatin accessibility during CD8+ T cell differentiation

To understand the role of Batf in regulating the differentiation of effector CD8+ T cells, we generated inducible Batf conditional knockout (cKO) CD8+ T cells derived from the P14 TCR transgenic mouse, in which the CD8+ T cells recognize the GP<sub>33-41</sub> lymphocytic choriomeningitis virus (LCMV) epitope (Pircher et al., 1987). The Batf cKO P14 CD8+ T cells generated in donor mice were adoptively transferred to recipient mice one day before infection with the Armstrong (Arm) strain of mouse LCMV, which causes acute infection (Figures 1A and S1A). Wild type (WT) P14 CD8+ T cells were also sorted and co-transferred to recipient mice as control.

Our previous study demonstrated that Batf KO P14 TCR transgenic CD8+ T cells cannot control acute infection of LCMV, and instead exhibited cell death four days post infection (Kurachi et al., 2014). To investigate viable CD8+ T cells in this model, we therefore harvested the Batf cKO cells at 3.5 days post infection, shortly before the anticipated population collapse, and performed RNA-seq and ATAC-seq in order to compare the Batf cKO to the WT control. RNA-seq showed that deletion of Batf resulted in significant transcriptional changes, with 2,552 genes affected (1,292 increased and 1,260 decreased; adjusted p-value < 0.05; see Methods and Figure 1B, Table 1). Batf cKO CD8+ T cells had a marked decrease in mRNA levels of TFs and signaling molecules that are important for T cell activation and function, including *Fos*, *Ctla4*, *Maf*, *Il12rb2*, *Hif1a* and *Cd7*. Genes with increased expression in the cKO samples included ones that are characteristic of naive T cells such as *Sell*, *Tcf7*, and *Ccr7* as well as other genes involved in T cell function such as *Il10ra* and *Cxcr5*.

To determine if loss-of-function of Batf also affected the chromatin landscape in those antigen-exposed CD8+ T cells, we investigated the changes in chromatin accessibility observed between the WT and the cKO conditions using ATAC-seq. We first used peak calling to map all loci that were sufficiently accessible in at least one condition (Figure S1A and Methods), identifying a total of 38,055 chromatin accessible regions (ChARs). We then searched for ChARs that were differentially accessible (DA) (see Methods, Table 2). Overall, 6,011 ChARs were more accessible in the Batf cKO condition, while 6,605 ChARs were less accessible (Figure 1C). These results showed that Batf is required to establish both the chromatin landscape and transcriptional

program of activated CD8<sup>+</sup> T cells. Unlike CD4<sup>+</sup> T cells (Pham et al., 2019), the changes in chromatin accessibility and transcriptome in Batf-deficient CD8<sup>+</sup> T cells were bidirectional which indicated dual roles of Batf as a possibly indirect positive or negative regulator.

We further confirmed the relationship between chromatin changes and gene expression by associating each DA ChAR with nearby genes using GREAT ([Table 3](#) and **Methods**). We found that a significant proportion (58.7%) of the differentially expressed (DE) genes were associated with at least one DA ChAR (48.5% if we only associate a ChAR with its nearest gene). This level of overlap was higher than expected by chance (fold enrichment = 1.24,  $p < 0.001$ ) ([Figure S1B](#)). We also observed a limited quantitative correspondence between the amount of change in mRNA expression and change in accessibility around each gene locus (Pearson correlation of 0.187 ( $p < 1e-4$ ) if we consider all loci, and of 0.622 ( $p < 1e-7$ ) if we only consider ChARs that resided in the promoter of their putative target gene [Figure S1C](#)). The set of genes that were downregulated in Batf cKO and associated with DA ChARs with reduced accessibility included several key TFs and signaling molecules in effector T cells, such as *Maf*, *Ctla4*, *Ctla2a*, *Ctla2b*, *Hif1a*, *Icos* and *Cd7*. Conversely, ChARs within and around the promoter of the *Sell* gene, which is highly expressed by naive cells and encodes for the L-selectin CD62L tended to be more accessible in Batf cKO. Consistently, we observed a three-fold increase of mRNA expression of that gene in the cKO condition ([Figure S1D](#)).

Taken together, these results point to a coordinated transcriptome- and epigenome-wide dependence of CD8<sup>+</sup> T cells differentiating *in vivo* on Batf. This is consistent with previous findings that proper function of this factor is critical for mounting an effective anti-viral or other adaptive immune responses (Grusdat et al., 2014; Xin et al., 2015).

### **The effects of Irf4 on the transcriptome and epigenome overlap with those of Batf**

The TF Irf4 is a well characterized binding partner of Batf (Li et al., 2012), and the importance of the interaction between these factors has been demonstrated in several processes, including ones in T helper cells (Ciofani et al., 2012; Li et al., 2012), tissue-resident regulatory T cells (Vasanthakumar et al., 2015) and cytotoxic T cells (Grusdat et al., 2014). To study the role of Irf4 in shaping the transcriptome and chromatin landscape of activated CD8<sup>+</sup> T cells, and to investigate whether it affects a similar set of genes and loci as Batf, we generated Irf4 cKO P14 T cells *in vivo*, using the same mouse LCMV model and experiment timing described for Batf

(Figures **1A** and **S1A**), and conducted RNA-seq and ATAC-seq (Figures **S1E** and **S1F**, Tables **1** and **2**).

We found that the transcriptional effects of *Irf4* loss were largely consistent with that of *Batf* (Pearson correlation of 0.506 between the induced fold changes;  $p < 1e-4$ ; Figure **1D**). For example, *Sell*, *Socs1*, *Nfkb1* and *Tcf7* were similarly up-regulated in both *Batf* cKO and *Irf4* cKO whereas *Ctla4* and *Maf* were down-regulated in both cKO cells. The effects on chromatin accessibility were largely consistent as well ( $r = 0.494$ ;  $p < 1e-4$ ; Figure **1E**). The combined effects in both modalities was observed in several key loci, such as *Ctla4* and *Maf*, which exhibited marked decrease in accessibility and mRNA levels in both perturbations, while *Sell*, *Socs1* and *Nfkb1* showed a marked increase in both modalities and both perturbations. It should be noted, however, that generally the loss of *Batf* led to an effect of stronger magnitude on chromatin accessibility compared to *Irf4* cKO while they had comparable affected gene numbers (Figures **S1G**, **S1H** and **S1I**). This, along with the partial overlap between the effects (e.g., on the expression of the TFs Jun and Fos, Figure **1D**), suggests that although *Batf* and *Irf4* shared many aspects of chromatin and transcriptional changes when knocking out, they may also have distinct contributions to the epigenetic and transcriptional rewiring required for T cell activation.

### **The *Batf* dependent transcriptome and epigenome are associated with effector T cell differentiation**

To better connect between the observed effects of *Batf* to the process of T cell response to antigens, we produced, as a reference, a time course gene expression dataset from *in vitro* CD8+ culture, from naive to effector. To this end, naive CD8+ T cells were isolated from WT mice, stimulated with anti-CD3/CD28, and assayed up to 6 days post stimulation with an additional restimulated sample at day 6 (using PMA/ionomycin; see **Methods**). We first identified genes whose expression changed over time and divided these genes into five prototypical clusters of response (Figure **1F** and **Methods**). Using GSEA (Table **3**), we then found that the set of genes that were up-regulated in *Batf* cKO significantly overlapped with clusters of genes that had reduced expression after stimulation (Clusters 4 and 5, adjusted  $p < 0.01$ ), including *Fas* (apoptosis-inducing gene (Strasser et al., 2009)), *Scos3* (cytokine signaling regulator (Brender et al., 2007)) and *Sell* (naive T cell marker (Arbonés et al., 1994)). Conversely, genes that were down-regulated in the absence of *Batf* were associated with clusters that become induced post

primary stimulation (**Figure 1G**, Clusters 1 and 2, adjusted  $p < 0.01$ ), including the costimulatory receptor *Icos* (Dong et al., 2001), CD8<sup>+</sup> effector response genes such as *Hif1a* (Phan and Goldrath, 2015), cytokine receptors that are related to effector generation such as *Il12rb2* (Danilo et al., 2018) and the cell proliferation marker *Mki67* (Gerdes et al., 1984). From the epigenome perspective, we found that regions which were more accessible in the Batf cKO were most strongly associated with clusters of reduced expression after stimulation (using GREAT; Clusters 4 and 5,  $q$  value  $< 1e-5$ ; **Table 4**). Regions that were more accessible in the WT had a less clear trend, with significant overlap with cluster 1 (induced expression post stimulation) and cluster 4 (decreased;  $q$  value  $< 0.05$ ). Results for Irf4 cKO were similar (**Figure S1J** and **Table 4**), suggesting that deficiency of Batf or Irf4 impairs the expression of TCR-inducible genes and the subsequent transition from naive to effector state. These results accord with our previous work whereby loss of Batf led to defects in the differentiation and expansion of stimulated CD8<sup>+</sup> T cells *in vivo* (Kurachi et al., 2014).

To further explore this hypothesis, we utilized a compendium of published data of transcriptional changes in CD8<sup>+</sup> T cells under different conditions. We tested whether these previously observed changes overlap with the effects we observed of Batf on the transcriptome and epigenome (using GSAE and GREAT, see **Method**; **Figures 1H**, **1I**, **S1J** and **S1K**, **Tables 3** and **4**). Both data modalities clearly associated the effects of Batf in our data with previously observed differences between naive and effector CD8<sup>+</sup> T cells. For instance, genes that were down-regulated in the absence of Batf in our data tend to be up-regulated in the effector state. Similarly, genes that were proximal to ChARs that lost accessibility in Batf cKO tended to be over-expressed by the effector and activated populations. Results from Irf4 cKO also showed a similar trend towards the naive CD8<sup>+</sup> state when comparing to the effector state (**Figure S1J** and **S1K**). Using a similar GSEA procedure we also found that the sets of Batf-affected genes and loci significantly overlap with genes and loci that have been associated with interferon signaling and response. These results are concordant with our previous work showing abnormal interferon gamma (IFN $\gamma$ ) production in Batf KO cells *in vivo* (Kurachi et al., 2014). Interestingly, Irf4 cKO did not display a strong impact on interferon signalling pathways at transcriptomic level (**Figure S1J**), suggesting a potential mechanism of Batf that is independent of Irf4. Finally, analysis of DNA binding motifs in Batf-dependent ChARs revealed a significant over-representation of motifs that are associated with transcription factors or complexes that are important for (but not necessarily specific to) T cell function such as the tryptophan cluster (IRF, ETS and etc.), high mobility group (HMG), and

homeodomain. We also observed strong enrichment of basic leucine zipper (bZIP) motifs, which can be potentially bound by Batf itself (Figure S1L, Table 4).

Overall, these results serve to place the effects of Batf and Irf4 in the context of the process of CD8<sup>+</sup> T cell effector differentiation. We observed, in both cases, that the transcriptional and epigenetic changes induced by the loss of Batf or Irf4 had a significant level of similarity to the ones observed prior to the antigen exposure and activation of naive T cells. These results therefore suggest that in the cKO conditions, the cells were prone to stall in a naive state compared to WT.

### ***In vitro* profiling of transcription factors reveals genome-wide dependence on Batf**

While our motif analysis suggested plausible dependence of other transcription factors on Batf function, it affords a limited view of the actual binding landscape (Yue et al., 2014). For a more direct analysis we used ChIP-seq of WT and Batf KO P14 CD8<sup>+</sup> T cells (Kurachi et al., 2014) cultured *in vitro* to identify key binding locations and evaluate their dependence on Batf. To this end, naive cells were stimulated with anti-CD3/CD28, cultured for 6 days and restimulated with PMA/ionomycin (similarly to Figure 1F; see Methods). We assayed a list of transcription factors whose presence is important for T cell function and development, including the CD8<sup>+</sup> lineage determinant, T-bet (Intlekofer et al., 2005; Sullivan et al., 2003), the factor for memory generation, Eomes (Intlekofer et al., 2005; Pearce et al., 2003), factors critical for CD8<sup>+</sup> development such as Runx3 and (Cruz-Guilloty et al., 2009; Shan et al., 2017), Ets1 (Grenningloh et al., 2011; Zamisch et al., 2009), a factor required for maintaining effector function, Fosl2 (Lingel et al., 2017; Lund et al., 2007; Stelekati et al., 2018), the Batf binding partner and transactivator Jund (Li et al., 2012), and the cytokine signaling transducers such as Stat3 (Cui et al., 2011; Kujawski et al., 2010) and Stat5 (Tripathi et al., 2010), in addition to Batf and Irf4. The set of putative binding locations (peaks) in each sample was identified using MACS2 (Zhang et al., 2008) (Table 2, Figures 2A and S2A).

Starting with the WT samples, we found a broad range in the extent of binding, from widespread (tens of thousands of loci for Jund and Irf4), to more specific (few thousand loci or less for Runx3 and T-bet; Figure 2B), with most sites residing in non-coding regions (Figure 2C, Table 2). We also observed a large extent of colocalization between the TFs in our dataset (Figure 2D). For instance, and as expected, Batf bound loci significantly overlapped with those of its well-

characterized partners, *Irf4* and *Jund* (Figure 2D middle). We also observed that most of the binding regions of *Runx3* and *T-bet* were proximal to a *Batf* binding site (Figure 2D right). Furthermore, most TFs exhibited a clear preference for localizing near ChARs that lost accessibility in *Batf* KO (more than expected by chance,  $p < 0.001$ ; Figure 2E and Methods). This effect was most pronounced in regions bound by *Runx3* and *T-bet* (Figure 2F) or in “hot” regions occupied by multiple TFs (Figure 2G, Table 2). A similar trend was observed with DA ChARs in the *Irf4* cKO (Figure S2B). Taken together, these results demonstrate that a group of TFs that are critical for CD8+ differentiation and function tend to operate in overlapping regions of the genome, and that *Batf* and *Irf4* have a broad effect on the accessibility of these regions. We next investigated whether the affinity of binding in those regions indeed changed as a consequence of *Batf* deficiency.

To explore this we plotted, for each TF, the changes in its binding signal versus the changes in accessibility in all loci where it bound next to *Batf* and where at least one of the two signals (evaluated by ATAC-seq and ChIP-seq) changed significantly (*Batf* KO vs. WT; adjusted  $p$ -value  $< 0.05$ ; Figure 2H, Table 2, and Methods). The majority of TFs had a significant positive association between changes in accessibility and changes in binding, suggesting that a nearby *Batf* binding event may help facilitate, directly or indirectly, their own binding by means of chromatin organization. The most pronounced loss of binding was observed with *Jund*, which is capable of dimerizing and co-binding with *Batf* (Murphy et al., 2013), and with *Stat5*, which mediates IL-2 signaling and consequently differentiation and homeostasis (Ross and Cantrell, 2018). Interestingly, and in contrast to *Stat5*, the binding of *Stat3* has primarily increased upon loss of *Batf* function, where a substantial part (48.3%) of the gained sites overlapped with a loss in *Stat5* binding (using csaw for differential binding with FDR  $< 5\%$ ; See Table 3 and Methods). An example for such *Stat5* to *Stat3* switch was found proximal to the BCL-2 family member *Bim* (*Bcl2l11*), which is regulated in T cells by JAK/STAT signaling (Bouillet et al., 2002) and whose expression increased in *Batf* cKO (adjusted  $p$ -value  $< 0.001$ ; Figure S6A). These results accord with previous reports on competitive binding and opposing regulatory effects of these two TFs (Walker et al., 2013) and suggest a role for *Batf* in maintaining their opposing effects in activated CD8+ T cells. Unlike most other TFs we profiled, the changes in the binding of *Stat3* in the absence of *Batf* did not have a strong association with changes to accessibility. This may suggest other means in which *Batf* exerts its effects in *Stat3* sites, such as changes in the activity of binding competitors (Sadreev et al., 2018; Walker et al., 2013; Wingelhofer et al., 2018) and/or in post translational modifications at nearby histone tails (Chan and Maze, 2020). Notably, we observed



a similar trend of low correlation between changes in binding versus changes in accessibility in the case of Fosl2. This result accords with previous reports on antagonistic relationship between Fosl2 and Batf in Th0 and Th17 conditions, whereby Fosl2 binding sites are overlapped with those bound by Batf, thus competing for the same AP-1 sites (Ciofani et al., 2012).

Taken together, our results suggest that Batf may, at least partially, exert its effects on CD8+ effector differentiation by modulating the binding landscape of critical TFs and may be dependent or independent of chromatin accessibility. Of note, out of the ten TFs we studied, Runx3 and T-bet had the largest overlap with Batf binding sites and with Batf-dependent DA ChARs, while showing clear positive association between these two signals.

### **Batf-mediated chromatin loops are associated with Batf-dependent accessible regions and transcription factor binding**

While we investigated the effects of Batf on the chromatin by mapping local changes in accessibility and TF binding (Figure 1 and Figure 2), these effects may also be associated with higher-order alterations, namely long-range contacts between otherwise distant genomic regions (Pham et al., 2019). We therefore used HiChIP (Mumbach et al., 2016) for the *in vitro* differentiated effector CD8+ T cells (naive CD8+ T cells stimulated with anti-CD3/28, cultured for 6 days and restimulated with PMA/ionomycin, same condition as in figure 1F, figure 2, see Methods) to identify chromatin loops that are likely dependent on the presence of Batf. For the purpose of comparison, we used the CCCTC-binding factor (Ctcf), which plays a housekeeping role in defining genomic boundaries in many types of mammalian cells, including T cells (Rowley and Corces, 2018). Starting with a global perspective, we observed a significant level of overlap between regions nearby Batf-dependent loops and binding locations of TFs in our study. We also found a significant overlap between Batf-dependent loops and ChARs that lost accessibility in the absence of Batf or Irf4. Notably, the observed overlaps were generally more pronounced than the ones observed with Ctcf-anchored loops (Figure 3A, Table 5).

Taking a closer look, we found that regions that were bound by certain TFs tended to form larger than expected numbers of long-range interactions (Figure 3B; Wilcoxon test adjusted p-value < 0.05 on each test comparing regions with a given feature to those without). This tendency was especially pronounced in regions bound by Runx3 or T-bet, and in “hot” regions that are occupied by many TFs (resonating with the notion of super enhancers (Huang et al., 2018)). Finally, through

pairwise analysis of our data, we found that many TF binding sites came into contact with one another or with Batf-dependent ChARs through Batf-mediated loops (Figure 3C). These trends may be reflective of the effects of Batf on transcription, as we observed many cases by which key regions (TF binding sites, or Batf-dependent ChARs) come into contact with the promoters of genes with a Batf-dependent expression level (Figure 3D). As an example, we observed a dense set of Batf-anchored loops around the *Ctla4* and *Icos* loci, which were both downregulated in Batf cKO (Figure 3E). These loops brought in close proximity a group of regulatory regions that required Batf for accessibility and were bound by most of the TFs we examined (Figure S3A).

Taken together, these results associate many of the observed effects of Batf on the transcriptome and epigenome with higher-order organization of the chromatin. Specifically, they suggest that Batf-dependent chromatin loops may serve to bring in proximity regions that are bound by key TFs, and that Batf may be critical for both the accessibility of these regions and for the long range contacts they make. Among the TFs we studied, the extent of overlap was most pronounced for Runx3 and T-bet, with many binding sites coinciding with Batf mediated loops (Figure 3A), forming larger numbers of loops per-site (Figure 3B). Furthermore, Runx3 and T-bet exhibited the largest proportion of Batf-mediated interactions with regions bound by other TFs and with regions with Batf-dependent accessibility (Figure 3C). These results thus suggest a strong link between Batf-mediated organization of the chromatin and the activity of Runx3 and T-bet.

### **Overexpression of Batf and other key transcription factors recapitulates aspects of CD8+ T cell chromatin landscape and gene expression in fibroblasts**

Having observed that Batf and Irf4 affect critical chromatin features that are required for CD8+ T cell differentiation and function, and that T-bet and Runx3 preferentially bind regions that are regulated by Batf and Irf4, we sought to determine whether these four transcription factors are sufficient for recapitulating CD8+ T cell chromatin and transcriptional properties in a different cell type. Thus, we set up an ectopic gene expression system in NIH/3T3 fibroblasts, which allowed us to express all possible combinations of these four TFs, and measure the resultant effects on chromatin accessibility and the transcriptome.

We first used western blot and RNA-seq to confirm that the endogenous expression of Batf, Irf4, Runx3 and T-bet in NIH/3T3 fibroblasts was undetectable (Figure S4A and 4A). We then transduced the fibroblasts with a doxycycline-inducible lentiviral vector consisting of any



combination of Batf, Irf4, Runx3 and T-bet (16 conditions altogether). Successfully transduced cells were selected with antibiotics or fluorescence (Figure S4B). Ectopic expression of the respective TFs in each sample was then induced simultaneously with doxycycline for 72 hours, as confirmed by western blots and RNA-seq (Figure S4A and Figure 4A). Transduced fibroblasts were subjected to profiling with RNA-seq (Table 6) and ATAC-seq (Table 7), providing an estimate for the changes in the transcriptome and chromatin landscape for each combination of TFs. As a negative control we also assayed the same set of transcription factor combinations with no doxycycline induction (16 control conditions altogether).

Focusing on the gene expression data, we found that different combinations of the four TFs led to increased expression of several genes involved in the regulation of T cell homeostasis, activation, proliferation, differentiation and function (Figure 4A), such as IRF family members (*Irf1*, *Irf2*, *Irf5*, *Irf7* and *Irf9* (Hammami et al., 2015; Hida et al., 2000; Martinet et al., 2015; Simon et al., 1997; Zhou et al., 2012)), JAK-STAT signaling pathway members (*Jak1*, *Jak2*, *Stat1*, *Stat2* and *Stat5a* (Seif et al., 2017)), SOCS family members (*Socs1*, *Socs2* and *Socs4* (Palmer and Restifo, 2009)), cytokine signaling (*Tgfb1*, *Il2rb*, *Ifngr1* and *Il17ra* (Curtsinger et al., 2012; Oh and Li, 2013; Ross and Cantrell, 2018; Tosello Boari et al., 2018)) and other factors (*Atf3*, *Batf2*, *Ets2*, *Gadd45a*, *Klf4*, etc. (Filén et al., 2010; Kayama et al., 2019; Salvador et al., 2005; Yamada et al., 2009; Zaldumbide et al., 2002)). To more systematically evaluate these transcriptional effects, we created a gene signature consisting of genes that are the most differentially expressed (up or down) between untreated fibroblasts and CD8<sup>+</sup> T cells (using DESeq2; see Methods). We then used a method akin to Vision (DeTomaso et al., 2019) to calculate a signature score for each condition, estimating its position in the spectrum between CD8<sup>+</sup> and fibroblast transcriptional states (Figure 4B). We observed that the ectopic expression of Batf and Irf4 is generally associated with a high T cell score (closer to T cell state than others). Conversely, ectopic expression of Irf4 only or Batf only resulted in less pronounced effects, supporting the model of synergism between these two TFs.

To assess each TF's individual contribution to changes in our T cell score, we used the Shapley Value - a game theoretic measure that summarizes the contribution of individuals in cooperative scenarios (Keinan et al., 2004). Batf and Irf4 scored the highest, indicating that these two TFs played the primary role in reshaping the transcriptome of fibroblasts toward a T cell-like profile (Figure 4C). The score for T-bet was substantially lower, and Runx3 had a slightly negative score.

This indicates that these two TFs had smaller impacts from the genome-wide perspective, but not excluding a more nuanced contribution in regulating specific genes.

Indeed, while over expression of *Batf* and *Irf4* resulted in marked increase in the expression of critical T cell factors (Figure 4A), we found parts of the T cell program that were most strongly upregulated when *Runx3* and/or *T-bet* were ectopically expressed alongside *Batf* and *Irf4*, including chemokine receptors such as *Ccr7*, cytokines and cytokine receptors such as *Il2rb*, *Il17ra* and *Tgfb1*, and T cell transcription factors *Stat5a*, *Egr2* and *Elf4* (Figure 4A, Table 6). To confirm this, we selected genes that play important roles in CD8+ T cell homeostasis, survival and function (Acharya et al., 2017; Gaffen, 2009; Jung et al., 2016; Tinoco et al., 2009; Tosello Boari et al., 2018) and examined their expression with flow cytometry (*Il17ra* and *Ccr7*, Figure 4D left) or western blot (*Tgfb1*, Figure 4D right). Consistent with our RNA-seq analysis, we observed that these genes become induced in the treated fibroblasts and that their expression is most pronounced when all four TFs are present. To gain further insight into these transcriptional changes, we applied GSEA using as reference the literature-curated MSigDB C7 collection of transcriptional dissimilarities between subsets of immune cells (Godec et al., 2016) (Figure 4E, Table 8). Overall, we found that ectopic expression of *Batf* and *Irf4* resulted in transcriptional differences akin to those between effector CD8+ T cells and other CD8+ T cell states (e.g. naive and memory; Figure 4E, left). In agreement with most of the immune signatures we listed, the addition of *Runx3*, *T-bet* or both further strengthen the differences (compared to fibroblasts treated with only *Batf* and *Irf4*) between effector CD8+ versus other CD8+ populations. *Runx3*, *T-bet* or both TFs were also observed to differentiate the CD8+ identity when comparing CD8+ signatures with other cell types (e.g. CD4+, B cell, monocyte, NK and NKT cells; Figure 4E, right). Taken together, this analysis indicates that over expression of the four TFs resulted in a shift toward the effector identity within the CD8+ T cell populations.

Given the role of *Batf* in modulating the epigenome of CD8+ T cells, we reasoned that the transcriptional changes we observed are accompanied by epigenetic reprogramming of the treated fibroblasts. Indeed, signature analysis of global changes in chromatin accessibility with ATAC-seq (using as signature regions that are differentially accessible between fibroblasts and T cells, in a manner akin to Figure 4B) points to a similar phenomenon, with gradual modulation of the chromatin accessibility landscape in fibroblasts toward T cell like profile (Figure S4C). As an example, we considered loci around *Tgfb1* (Oh and Li, 2013), which is not normally expressed by fibroblasts but became induced in our ectopic system (Figure 4F). We observed a specific locus

downstream of *Tgfb1* that became accessible only when the four TFs were expressed. This locus is also associated with loss of accessibility in Batf cKO and with direct binding by Batf and Irf4 in T cells (Table 2, Figure S5A) and may therefore act as an enhancer of the nearby *Tgfb1* locus, whose activity in T cells depends on all four TFs.

Taken together, our results suggest a major role for Batf and Irf4 in shaping the transcriptome and epigenome of CD8<sup>+</sup> T cells, due to their effects on a subset of key genes and loci that are otherwise inactive in fibroblasts. The results also suggest that Runx3 and T-bet may be important for fine-tuning or enhancing these global effects, toward a more specific effector CD8<sup>+</sup> T cell like phenotype.

### **Establishment of long range chromatin loops characteristic of T cells requires overexpression of Runx3 and T-bet in addition to Batf and Irf4**

Having found that Batf bound loci engage in long-range genomic contacts in T cells, we sought to determine if its ectopic expression (alone or with its associated factors) can reconstitute similar chromatin structures in fibroblasts. To this end, we performed Batf HiChIP experiments in fibroblasts expressing only Batf or a combination of all four transcription factors.

Examples of Batf-anchored chromatin loops in these manipulated fibroblasts around the *Nfatc1* and *Tgfb1* loci are shown in figure 5A and S5A. Ectopic expression of Batf alone generated mostly short range loops near these loci, which only partially recapitulated the Batf loops in WT effector CD8<sup>+</sup> T cells. However, when Batf/Irf4/Runx3/T-bet were co-expressed, the resulting Batf-anchored loops recapitulated many of the loops found in T cells, especially long-range contacts. Noticeably, the 5' proximal promoter region of *Nfatc1* has a Batf-dependent accessibility and is co-bound by Batf and Irf4 in T cells. It is also linked by Batf-mediated loops to the 3' end of that gene, which (in T cells) includes binding sites for Batf, Irf4, Runx3 and T-bet. Finally, we note that Batf alone did not alter the expression of *Nfatc1* in fibroblasts, which required the addition of other factors (Figure 4A). These findings therefore point to a composite epigenetic structure, (TF binding, chromatin accessibility and chromatin loops) that may contribute to the expression of a nearby gene, while depending (directly or indirectly) on the joint activity of several TFs, in addition to Batf.

A more global comparison of Batf-dependent genomic contacts in T cells vs. manipulated fibroblasts indeed reveals a higher level of concordance in cells co-expressing Batf/Irf4/Runx3/T-bet vs. Batf alone, and with differences most pronounced in the case of long-range loops (**Figure 5B and S5B**). Indeed, overall we observed that approximately 40% of the Batf-dependent loops in T cells were recapitulated in fibroblasts when overexpressing all four TFs versus 20% when over expressing Batf alone. The increase in concordance was observed more specifically next to regions that were functionally important in CD8<sup>+</sup> T cells, namely Batf-dependent ChARs (especially ones that lose accessibility in Batf cKO), Irf4-dependent ChARs, and regions bound by key TFs (**Figure 5B and 5C, Table 10**). This trend was especially strong in regions bound by Runx3, T-bet, Stat3, and Ets1 in T cells, or regions occupied by many TFs (**Figure 5B, 5C and S5B**).

Taken together, these results demonstrate that overexpressing the four TFs, Batf, Irf4, Runx3 and T-bet in fibroblasts, but not Batf alone, was sufficient to reconstitute a chromatin architecture that was closer to that observed in CD8<sup>+</sup> T cells. It therefore supports the notion that Batf exerts its regulatory effects, at least in part, by modulating the long range architecture of the chromatin, and that these changes required the presence of other cofactors, including Irf4, Runx3 and T-bet.

## DISCUSSION

The differentiation of naive CD8<sup>+</sup> T cells into an effector state is a critical step in adaptive immunity. Although many of the TFs that govern T cell differentiation or response to stimulation have been characterized, the regulation of the earliest epigenetic and transcriptional events that initiate the transition of cytotoxic T cells into an effector state remains partially understood. Previous work by us and others (Godec et al., 2015; Grusdat et al., 2014; Kurachi et al., 2014; Kuroda et al., 2011) identified the transcription factor Batf to be essential for establishing commitment into the effector state. Here, we further explored how Batf, along with other key TFs, may achieve its influence on this process.

Our study provides a detailed account of the transcriptome and epigenome of differentiating CD8<sup>+</sup> T cells in mice, and their dependence on the proper functioning of Batf. Our data set includes genome-wide characterization of the transcriptome of CD8<sup>+</sup> T cells (RNA-seq) in an *in vivo* infection model and in a time course of naive CD8<sup>+</sup> T stimulated *in vitro*. It also includes a map of the accessible chromatin landscape (ATAC-seq), binding of ten key TFs (ChIP-seq) and mapping of long-range chromatin loops that are anchored by Batf (Hi-ChIP). To probe the causal effects that Batf may have on the differentiation and function of CD8<sup>+</sup> T cells, we repeated our primary measurements using samples from Batf KO animals. Joint analysis of these data resulted in a map of the CD8<sup>+</sup> T cell epigenome, consisting of 96,490 genomic regions along with the TFs that bind them, their accessibility, their contacts with other regions, the expression from nearby genes, and the dependence of all these modalities on the presence of Batf. Through this joint analysis, we observed that the absence of Batf had a substantial effect on both the transcriptome and the structure of the chromatin. The transcriptional effects, which were observed in thousands of genes, largely indicated that Batf contributes to a shift into an effector state, whereby KO cells tend to retain a naive-like gene expression program with markers such as *Sell* and *Ccr7*. The respective epigenetic changes were often observed in proximity to the affected genes and had a similar level of abundance, with thousands of affected loci. Intriguingly, we observed these effects on several levels: the increase or decrease in accessibility often coincided with changes to binding of key TFs, which in turn were observed close to Batf-anchored chromatin loops.

The availability of genome wide estimates of both chromatin accessibility and protein-DNA interactions can lend insight into the manner by which Batf affects different TFs. In most cases, we observed a positive correlation between changes in binding and changes in accessibility

(comparing KO to WT cells). While it may suggest a simple global mechanism by which Batf-dependent accessibility is required for binding, we cannot rule out other alternatives. Indeed, it is interesting to note that for some TFs (most prominently T-bet, Stat5 and Jun), changes to binding were restricted to decrease, which became more severe as accessibility decreased. Conversely, in other TFs (e.g., Runx3, Irf4) we also see an increase in binding at loci that became more accessible in the absence of Batf. While Batf-dependent changes to the mRNA expression of these TFs might in principle serve to explain such differences, we have not observed such trends in our data (Figure S6D). It is therefore conceivable that our observations may result from additional (direct or indirect) consequences of Batf function, such as changing the expression of other TFs outside our panel, modifications of TF proteins or nearby histone tails, and local alterations in binding affinity through protein-protein contacts. The latter option is further illustrated in the case of Fos12 and Stat3. Unlike the other transcription factors we assayed, Fos12 and Stat3 exhibited marked increase in binding in Batf KO, but this increase had only weak association with accessibility. While there may again be several alternative explanations, it is interesting to note that in both cases the increase in binding may result from reduced activity of a binding competitor (Batf and Stat5, respectively).

The observation that Stat5 lost its binding while Stat3 gained binding in the absence of Batf further suggests a general regulatory mechanism that controls the activity of Stat family proteins. Stat3 and Stat5 are known to be activated by the IL-2 signaling pathway, however while they have been demonstrated to be non-redundant and function differently in T cells (Laurence et al., 2007; Moriggi et al., 1999), the mechanisms that balance their usage are not well characterized. Our data therefore suggests an additional function of Batf that may serve to regulate the activity of these two proteins. Within our system, this balancing raises an hypothesis as to what may contribute to the enhanced T cell death and subsequent lack of functional effector cells in infected naive mice transferred with Batf KO cells (Kurachi et al., 2014). Although Batf-deficient cell counts diminished on day 4 after LCMV challenge, the KO cells can be easily cultured and expanded *in vitro* in the presence of recombinant IL-2. Given that the expression of IL-2 receptors (*Il2ra*, *Il2rb* and *Il2rg*) were intact in Batf cKO (Figure S6A), the defects of Stat5 binding and increased Stat3 binding in Batf cKO (Figure S6B) may contribute to the IL-2 unresponsiveness and cell death observed in Batf KO since Stat5 is the immediate downstream signal transducer of IL-2 signaling pathway (Lin et al., 1995). Our data points to additional (possibly downstream) potential mediators of the observed cell death in Batf KO *in vivo*, which may rely on some, but not all, of our measurement modalities. For instance, the apoptotic factor Bcl211 (*Bim*) (Bouillet et al., 2002)

has elevated expression in the absence of Batf. We further find that its promoter comes in Batf-mediated interaction with a distal site (> 300 kb), and that both promoters and distal sites are bound by Batf, Irf4, Runx3, and T-bet (Figure S6C). However, the accessibility of these sites and the majority of their binding events are not dependent on Batf, raising the option of involvement by additional factors that are not included in our measurements.

In agreement with previous observations of similar phenotypes in mice deficient for either Batf or Irf4 (Grusdat et al., 2014), we found that perturbing these two proteins resulted in similar trends in terms of the transcriptome and epigenome. These effects, however, were not completely overlapping. For instance, while the absence of Batf had a significant effect on the expression of interferon signaling genes, this effect was not as strongly observed with Irf4. More globally, the number of regions differentially accessible in Batf cKO was almost twice as much as in Irf4 cKO, suggesting that it plays a broader or more critical role in reshaping the chromatin post stimulation. Either Batf or Irf4 alone, however, were insufficient for inducing broad changes in the epigenome of fibroblasts, while their joint effect was sufficient to engage some aspects of a CD8<sup>+</sup> core program, thus adding to a growing body of evidence on cooperativity between these two molecules (Grusdat et al., 2014; Man et al., 2017; Yao et al., 2013).

In addition to Irf4, cumulative evidence from all of our genomic assays pointed to two additional TFs, Runx3 and T-bet, as key regulators which may work in concert with Batf and Irf4. Specifically, we found that Runx3 and T-bet tend to bind in regions whose accessibility depends on Batf and that the strength of binding changed proportionally to Batf-dependent changes in accessibility. They also had the strongest tendency to bind next to other transcription factors (particularly Batf) through direct genomic proximity or through long-range Batf-mediated loops. Finally, both Runx3 and T-bet tended to be connected, by long-range chromatin loops, to the promoters of genes with a Batf-dependent expression. Indeed, ectopic expression in fibroblasts provided evidence that these factors contribute to strengthening the induction of a T-cell like transcriptional program, when co-expressed with Batf and Irf4.

These results accord with reports that T-box proteins and Runx3 contribute cooperatively to the effector program of CD8<sup>+</sup> T cells (Cruz-Guilloty et al., 2009). They further agree with more recent publications (van der Veecken et al., 2019; Wang et al., 2018), which showed that T-box factors and RUNX family proteins are required to drive and maintain the changes in the chromatin landscape during effector differentiation and memory generation of CD8<sup>+</sup> T cells. Adding to that,



our over-expression analysis suggests that T-box and Runx3 may act synergistically with Batf and Irf4 to induce a T-cell like program, in terms of gene expression, chromatin accessibility and long range chromatin loops. The mechanism behind these effects is yet to be determined though. While it has been observed that Runx3 and T-bet can physically interact with each other (Djuretic et al., 2007), currently there is no evidence showing physical associations with Batf or Irf4. One study showed the interaction between AP-1 (Fos and Jun) and Runx family proteins (D'Alonzo et al., 2002) which raised the possibility of Batf interacting with AP-1, Runx3 and T-bet in the same complex, since Batf interacts with AP-1 constantly. Interestingly, however, crystal structure study showed that T-bet was capable of binding to two independent strands of DNA and may function as a chromatin linker (Liu et al., 2016). Thus, Batf may connect the co-regulated regions through the binding of T-bet to form a regulatory complex.

While our focus was on the cooperation between Irf4, Runx3, T-bet and Batf in effector T cells, our data may be further explored for insight on the activity of other TFs, and possibly on the regulation of other CD8<sup>+</sup> T cell states. Indeed, Batf has been shown to also be expressed by naive and exhausted T cells, and may interact with different proteins in different cell contents (Chang et al., 2014; Man et al., 2017). For instance, while the expression of T-bet was shown to characterize naive and effector T cells, the activity of Eomes, a T-box TF with high homology to T-bet which we profiled here as well, has been more prominently associated with memory and exhausted T cells (Buggert et al., 2014; Sen et al., 2016). The extensive overlap we observed between Batf and Eomes binding sites, their proximity to genes and their dependence on Batf may therefore provide further hypotheses for regulation in these subsets.



## FIGURE LEGENDS

### Figure 1. Batf and Irf4 remodels chromatin accessibility during effector CD8<sup>+</sup> T differentiation.

**(A)** Schematic of Cre-lox conditional KO system and LCMV acute infection mouse model. **(B)** Volcano plot of log<sub>2</sub> fold change in gene expression (Batf cKO/WT) and adjusted p-values for differential expression tests in Batf cKO CD8<sup>+</sup> T cells. A subset of genes with important T cell functions are labeled. **(C)** Volcano plot of fold change in chromatin accessibility in a chromatin ChAR, measured by Tn5 cuts in locus (Batf cKO/WT), and adjusted p-values for tests of differential accessibility. ChARs were obtained by calling ATAC-seq peaks in the WT or Batf cKO conditions with MACS2. ChARs were mapped to putative target genes with GREAT, and a subset of ChARs mapped to genes with important T cell functions are labeled. **(D)** Scatterplot showing correlation between Batf cKO DE gene fold-change and Irf4 cKO DE gene fold-change. The universe of genes is limited to those with a significant change in expression (adjusted p-value < 0.05) in one of the two experiments. **(E)** Scatterplot showing the correlation between fold-change of Batf cKO and Irf4 cKO DA ChARs. The universe of genes is limited to those with a significant change in expression (adjusted p-value < 0.05) in one of the two experiments. A subset of ChARs mapped to putative target genes with important T cell functions are highlighted. **(F)** Transcriptomic profiling of *in vitro* cultured CD8<sup>+</sup> T cells. Gene expression was measured with RNA-seq daily, for six days. Row normalized expression counts are shown pre-stimulation (day 0), after primary anti-CD3/CD28 stimulation (days 1 to 6), and after PMA/ionomycin restimulation (day 6 restimulated). Genes were clustered into 5 groups using *hierarchical* clustering. Average expression trajectory for each cluster is shown on the left. **(G)** Normalized enrichment score (NES) for GSEA in the clusters shown in **(F)** in Batf cKO vs WT conditions. **(H)** NES for selected immunologic signatures of Batf cKO vs WT DE gene sets. Higher NES indicates enrichment in Batf cKO. Only results with adjusted p-value < 0.05 are shown. **(I)** Fold enrichment for selected immunologic signatures of WT vs Batf cKO DA ChARs analyzed using GSEA C7 immunological signature gene sets and GREAT. WT DA ChAR-associated gene sets are shown in red and Batf cKO DA ChAR-associated gene sets are shown in blue.

### Figure 2. Batf is required for binding of key transcription factors.

**(A)** Representative pileup of ChIP-seq fragments for main TFs in WT CD8<sup>+</sup> T cells around selected gene loci via IGV. Scale bars represent 5 kb. **(B)** The total counts of peaks for each TF in *in vitro* WT CD8<sup>+</sup> T cell samples. **(C)** Percentage of TF peaks across genomic features for WT

samples. **(D)** (Left heatmap) Percentage of each row TF's peaks that are co-bound by the column TF. (Center barplot) Percentage of each TF's peaks that are co-bound by BATF. (Right barplot) Percentage of Batf peaks that are co-bound by each TF. **(E)** TF ChIP-seq peaks stratified by change in accessibility in Batf cKO ATAC-seq: lost, gained, or unchanged. **(F)** Fold enrichment of TF binding over the DA ChARs lost in Batf cKO samples. Off-diagonal entries in the heatmap show the enrichment of regions co-bound by the two respective TFs. Diagonal entries in the heatmap and the barplot to the right show the enrichment for individual TFs. See **Method** for details of fold enrichment calculation. **(G)** Percentage of TF-bound regions, categorized by number of TFs bound, each stratified by change in accessibility in Batf cKO ATAC-seq: lost, gained, or unchanged. **(H)** Plot of relationship between change in TF binding (x-axis) and change in DNA accessibility (y-axis) for main TFs at loci bound by the TFs in WT or Batf KO conditions. The  $\log_2(\text{WT/Batf KO})$  values for binding and accessibility were estimated by applying csaw and DESeq2 to ChIP-seq and ATAC-seq data, respectively, and these points are binned into hexes colored by density of loci falling within them. Only loci exhibiting significant change in either binding or accessibility are shown (adjusted p-value < 0.05). TFs are arranged in order of decreasing Pearson correlation, which is shown on the lower-right of each plot.

**Figure 3. Differentially accessible chromatin regions and loci bound by key TFs are linked by Batf loops.**

**(A)** Fold enrichment of area spanned by HiChIP loop anchors for Batf (red) and Ctcf (blue-green) in area spanned by key CD8+ TFs and ChARs gained and lost in Batf cKO and Irf4 cKO CD8+ T cells. The enrichment score is defined as “percent of genome spanned by feature / percent of HiChIP loop anchor area spanned by feature”. **(B)** Boxplot of number of Batf-associated contacts to other regions, categorized by epigenetic feature. This is the “degree” of each locus in a network where each TF bound locus or ChAR is a node, and an edge between nodes exists if linked by Batf HiChIP loops. **(C)** Percent of regions with row feature (loci bound by TF or ChAR gained and lost in Batf cKO and Irf4 cKO) with a long-range HiChIP loop connected to column feature. Loop length threshold = 1,000,000 bp. Enrichment > 0.4 are highlighted with black edges. **(D)** Percent of regions that interact with DE genes linked by Batf long-range loops. A loop was considered to contact a DE gene if one of its loop anchors spanned the promoter of the gene. **(E)** Representative long-range contacts/loops in WT CD8+ T cells around *Ctla4* and *Icos* loci identified by HiChIP for Batf (red arc) and Ctcf (blue-green arc). Coordinates of ChIP-seq peaks identified by MACS2 for Batf are displayed as black bars.

#### Figure 4. Overexpression of Batf and other key TFs recapitulates aspects of CD8+ T cell chromatin landscape and gene expression

(A) Row normalized gene expression of selected T cell and fibroblast genes co-expressing four TFs combinations with or without doxycycline induction. (B) Relative T cell score of normalized gene expression for all experimental conditions in fibroblast. See **Methods**. (C) Box and whisker plot of the Shapley value for each TF, based on the relative T cell score. (D) Validation of selected surface markers and protein expression by flow cytometry and western blot quantified with densitometry for doxycycline-treated fibroblasts expressing no transcription factors, Batf only or Batf/Irf4/Runx3/T-bet. (E) Normalized Enrichment Score (NES) for pathways found in GSEA based on differential expression results in fibroblasts. The conditions tested are indicated by color and include fibroblasts expressing Batf/Irf4, Batf/Irf4/Runx3/T-bet, or no TFs controls. Gene sets and immunologic signatures related to CD8+ T cells are shown. (F) Representative tracks showing the chromatin accessibility changes in fibroblasts around *Tgfb1* loci, when expressing Batf/Irf4/Runx3/T-bet, Batf/Irf4 and no TFs. Batf, Irf4, Runx3, T-bet ChIP-seq peaks in CD8+ T cells and DA ChARs from Batf cKO are also shown as references for TF binding and chromatin accessibility. Shaded areas indicated putative regulatory regions for *Tgfb1* in fibroblasts. Bar plot of the relative mRNA expression of *Tgfb1* from control, Batf/Irf4 and Batf/Irf4/Runx3/T-bet expressing fibroblasts is shown on the right.

#### Figure 5. Overexpression of Batf, Irf4, Runx3 and T-bet can partially reconstruct chromatin architectures of fibroblasts

(A) Representative long-range genomic contacts/loops around *Nfatc1* locus identified by HiChIP for Batf in WT CD8+ T cells (red arcs), Batf in fibroblasts expressing Batf (blue arcs), and Batf in fibroblasts co-expressing Batf/Irf4/Runx3/T-bet (orange arcs). DA ChARs identified in WT vs Batf cKO CD8+ T cells are shown as red and blue bars (lost and gained in Batf cKO, respectively). ChIP-seq peaks identified for Batf, Irf4, Runx3 and T-bet in WT CD8+ T cells are displayed in black bars. (B) Coverage (percent of Batf HiChIP loops where both loops anchors are overlapped by both anchors of at least one loop in fibroblasts expressing Batf, or fibroblasts expressing Batf/Irf4/Runx3/T-bet) for different lengths of loops overlapped with defined features, including all loops, TF binding sites and DA ChARs in Batf cKO. (C) Fold enrichment of area spanned by HiChIP loop anchors by HiChIP for Batf in fibroblasts expressing Batf (blue), Batf in fibroblasts co-expressing Batf/Irf4/Runx3/T-bet (orange) and Ctcf (brown) in WT fibroblasts in areas spanned by key CD8+ TFs and ChARs gained/lost in Batf cKO and Irf4 cKO CD8+ T cells. Fold enrichment

is defined as “percent of genome spanned by feature / percent of HiChIP loop anchor area spanned by feature”.

## SUPPLEMENTARY FIGURE LEGENDS

### Supplementary Figure 1.

**(A)** Schematic of Cre-lox conditional knockout system and lymphocytic choriomeningitis virus (LCMV) acute infection mouse model. **(B)** Venn diagram showing the overlap between RNA-seq DE genes (WT vs Batf cKO, both up-regulated and down-regulated) and the genes associated with DA ChARs in ATAC-seq (WT vs Batf cKO, both up-regulated and down-regulated, consider only the nearest gene defined by GREAT). Hypergeometric p-value for the overlap is shown. **(C)** Scatter plot showing the relationship between chromatin accessibility and transcriptome changes in Batf cKO vs WT CD8<sup>+</sup> T cells. Only genes/loci which are significant at adjusted p-value <0.05 in either test displayed. **(D)** Representative Tn5 pileup from ATAC-seq at differential loci (*Sell* and *Maf*) identified by DESeq2. Relative mRNA expression levels of *Sell* and *Maf* from RNA-seq in Batf cKO and WT CD8<sup>+</sup> T cells are shown as barplots. **(E)** Volcano plot of log2 fold change (*Irf4* cKO/WT) in gene expression. **(F)** Volcano plot of log2 fold change (*Irf4* cKO/WT) in chromatin accessibility (Tn5 cuts) at loci identified as accessible (MACS2 peak) in either *Irf4* cKO or WT conditions. Name of the nearest gene mapped by GREAT is displayed for a subset of loci. **(G)** PCA plot of cuts in ATAC-seq peaks of naive, Batf cKO, *Irf4* cKO and control CD8<sup>+</sup> T cells in LCMV infection model. **(H)** DE gene counts of WT vs Batf cKO or *Irf4* cKO. **(I)** DA ChAR counts of WT vs Batf cKO or *Irf4* cKO. **(J)** GSEA of DE genes from Batf cKO and *Irf4* cKO CD8<sup>+</sup> T cells. Selected gene signatures and clusters of gene expression from in vitro culture are shown. Only results with adjusted p-value < 0.05 are displayed. **(K)** Fold enrichment for selected immunologic signatures of WT vs *Irf4* cKO DA ChARs GREAT pathways. WT DA ChAR-associated gene sets are shown in red and *Irf4* cKO DA ChAR-associated gene sets are shown in blue. **(L)** Top 50 enriched transcription factor motifs in DA ChARs identified in Batf cKO CD8<sup>+</sup> T cells with JASPAR database (Khan et al., 2017), by fold-enrichment in peaks differentially open in Batf cKO or WT conditions relative to union of peaks across the two conditions. Red circles represent DA ChARs enriched in WT (lost in cKO) and blue circles represent DA ChARs enriched in Batf cKO (lost in WT). Filled circles are datasets with an adjusted p-value < 0.05 whereas open circles are datasets with > 0.05.

### Supplementary Figure 2.

**(A)** Representative pileup of ChIP-seq fragments from Batf KO and WT CD8<sup>+</sup> T cells (normalized by control sample in MACS2) for key transcription factors (transcription factors) around selected genes. **(B)** Fold enrichment of transcription factor binding over the DA ChARs lost in *Irf4* cKO

samples. Off-diagonal entries in the heatmap show the enrichment of regions co-bound by the two respective transcription factors. Diagonal entries in the heatmap and the barplot to the right show the enrichment for individual transcription factors. See methods section for details of fold enrichment calculation.

### Supplementary Figure 3.

**(A)** Long-range contacts/loops in WT CD8<sup>+</sup> T cells around *Ctla4* and *Icos* loci identified by HiChIP for Batf (red arc) and Ctcf (blue-green arc). Coordinates of ChIP-seq peaks identified by MACS2 for indicated transcription factors from WT CD8<sup>+</sup> T cells are displayed as blue bars. The DA ChARs identified in ATAC-seq datasets for regions that are lost and gained in Batf cKO cells are shown as black and red bars, respectively.

### Supplementary Figure 4.

**(A)** Western blot validating protein expression of the four ectopically expressed CD8<sup>+</sup> TFs across all experimental conditions.. Heat shock protein 90 (Hsp90) was used as internal control. **(B)** Top: constructs of the doxycycline inducible TF expression system. HPH: Hygromycin B phosphotransferase, PAC: Puromycin N-acetyltransferase, BSD: Blasticidin-S deaminase. Bottom: schematic of lentiviral transduction and doxycycline inducible system for expressing four key CD8<sup>+</sup> T cell TFs in fibroblasts. Chart was created with BioRender.com **(C)** Percentage of T cell score of normalized DNA accessibility for all experimental conditions. T cell accessibility signatures obtained from a DESeq2 comparison of ATAC-seq of effector and naive CD8<sup>+</sup> T cells versus fibroblasts. Signature is defined by giving a weight of +1 to the top 1000 peaks differentially upregulated in T cells and weight of -1 to the top 1000 peaks differentially upregulated in fibroblasts. These weights were then applied to the vector of normalized counts for each condition.

### Supplementary Figure 5.

**(A)** Representative long-range genomic contacts/loops around *Tgfb1* locus identified by HiChIP for Batf in WT CD8<sup>+</sup> T cells (red arcs), Batf in fibroblasts expressing Batf (blue arcs), and Batf in fibroblasts co-expressing Batf/Irf4/Runx3/T-bet (orange arcs). DA ChARs lost in WT vs Batf cKO CD8<sup>+</sup> T cells are shown as blue bars. ChIP-seq peaks identified for Batf, Irf4, Runx3 and T-bet in WT CD8<sup>+</sup> T cells are displayed in black bars. **(B)** Jaccard statistics assessing recovery of T cell loops in fibroblasts expressing Batf or Batf/Irf4/Runx3/T-bet. Statistics are stratified by epigenetic features found on either anchor of the loop. Minimum loop length was set to 0.

## Supplementary Figure 6

**(A)** Relative mRNA levels of *Il2ra*, *Il2rb* and *Il2rg* of WT, Batf cKO and Irf4 cKO CD8<sup>+</sup> T cells sorted from day 3.5 LCMV challenged mice. **(B)** Representative IGV pileup of Stat3 and Stat5 ChIP-seq fragments from CD8<sup>+</sup> T cells culture *in vitro* around *Bim* (*Bcl2l11*) locus, both WT and Batf KO. **(C)** Representative plot (adapted from WashU Epigenome Browser) to show the ChIP-seq peaks (Batf, Irf4, Runx3 and T-bet from WT CD8<sup>+</sup> T cells), DA ChARs identified with MACS2 (WT vs Batf cKO), merged regulatory regions and HiChIP (Batf, WT CD8<sup>+</sup> T cell) loops around *Bim* (*Bcl2l11*) locus. **(D)** Normalized RNA counts of transcription factors from WT and Batf cKO CD8<sup>+</sup> T cells sorted from LCMV mouse model.

# SUPPLEMENTARY TABLE LEGENDS

## Table 1 - T cell RNA-seq

This table lists our RNA-seq experiments in T cells (in vitro and in vivo), and reports the normalized counts for each experiment as well as the differential expression results.

## Table 2 - T cell ATAC-seq and ChIP-seq

Each row in this table corresponds to a locus in the genome. Each column indicates an in vitro ChIP-Seq experiment or in vivo ATAC-Seq (listed in the first row), and indicates whether a peak was found with MACS2 and whether it was differentially accessible (ATAC-Seq) or differentially bound (ChIP-Seq) between WT and KO conditions.

## Table 3 - T cell GSEA

Gene set enrichment analysis (GSEA) results for differentially expressed (DE) genes, comparing knockouts (Irf4 or Batf) to WT CD8<sup>+</sup> T cells. The underlying DE genes can be found in Table 1.

## Table 4 - T cell chromatin enrichment analysis

This table reports whether the T cell chromatin features in Table 2 were found to be associated with transcription factor binding motifs, disease-associated SNPs, and other genomic annotations, and includes analysis from GREAT. See section “**Locus- level enrichment analysis**” in methods for details.

## Table 5 - T cell HiChIP

This table reports the summary statistics for the T Cell HiChIP experiments.

## Table 6 - Fibroblast RNA-Seq

This table reports differential expression results and normalized counts for the in vitro fibroblast experiments.

## Table 7 - Fibroblast ATAC-seq

Each row in this table corresponds to a locus in the genome. Each column indicates an in vitro ATAC-Seq (listed in the first row) experiment, and indicates whether a peak was found with MACS2 and whether it was differentially accessible under different experimental conditions.

## Table 8 - Fibroblast GSEA

Gene set enrichment analysis (GSEA) results for fibroblast experiments. The underlying DE results can be found in Table 6.

## Table 9 - Fibroblast chromatin enrichment analysis

This table reports whether the fibroblast ATAC-seq peaks from different experimental conditions in Table 7 were found to be associated with transcription factor binding motifs, disease-associated



SNPs, and other genomic annotations, and includes analysis from GREAT. See section “**Locus-level enrichment analysis**” in methods for details.

#### **Table 10 - Fibroblast HiChIP**

This table reports the summary statistics for the fibroblast HiChIP experiments.

## SUPPLEMENTARY RESOURCES

Our epigenetic datasets (ATAC-Seq peaks, ChIP-Seq peaks, and HiChIP loops) are available for browsing on the WashU Epigenome Browser (Li et al., 2019) at [https://epigenomegateway.wustl.edu/browser/?genome=mm10&sessionFile=https://yosef-lab-public-resources.s3-us-west-1.amazonaws.com/BATF/WashU\\_Browser/Datahub/BATF\\_Session\\_All\\_Tracks.json](https://epigenomegateway.wustl.edu/browser/?genome=mm10&sessionFile=https://yosef-lab-public-resources.s3-us-west-1.amazonaws.com/BATF/WashU_Browser/Datahub/BATF_Session_All_Tracks.json) .

# MATERIALS AND METHODS

## Mice and LCMV infection model

Batf KO mice in B6 background were purchased from the Jackson Laboratory (Schraml et al., 2009) and crossed with P14 TCR transgenic mice to generate P14 Batf KO mice (Kurachi et al., 2014). Five to 10 weeks of age female mice were used in this study. All animal work was performed in accordance with the Institute Animal Care and Use Guidelines of the University of Pennsylvania.

For the generation and adoptive transfer of Batf cKO P14 CD8<sup>+</sup> T cells, CD45.2<sup>+</sup> Batf-floxed and Irf4-floxed ERT2-Cre P14 donor female mice was treated with 2 mg 4-hydroxytamoxifen (4-OHT) per day per mouse intraperitoneally for 5 consecutive days. After the Cre-mediated recombination of the loxP flanked sequence, Batf was removed along with GFP. For Irf4 conditional KO constructs, Irf4 followed by a stop signal was excised by the induced Cre activity, thus inducing expression of YFP. Five days after the final dosing, mice spleens were then harvested and Batf cKO and Irf4 cKO CD8<sup>+</sup> P14 T cells were sorted with cell sorter (BD, FACS Aria) according to the GFP and YFP fluorescence expression, respectively. The sorted Batf cKO and Irf4 cKO CD8<sup>+</sup> T cells were mixed with WT P14 CD8<sup>+</sup> T cells ( $2 \times 10^4$ ) and adoptively transferred into WT mice expressing CD45.1 congenic marker.

LCMV Armstrong strain was produced and viral titers were quantified as described previously (Kao et al., 2011). Mice were challenged by intraperitoneal injection (i.p.) of LCMV with  $2 \times 10^5$  plaque forming units (PFU) one day after the adoptive transfer.

## In vitro CD8<sup>+</sup> T cell culture

Spleens from WT and Batf KO P14 mice were harvested, homogenized and passed through 70  $\mu$ m cell strainers. Splenocytes were spun down at 500 g for 10 min at 4 °C in 50 ml centrifuge tubes and the red blood cells were lysed with 1 ml ACK lysis buffer per spleen for 1 min at room temperature. Cells were washed once with MACS buffer (1X PBS, 0.5% BSA, 2 mM EDTA) and cell numbers were counted. Mouse naive CD8a<sup>+</sup> T cells isolation kit (Miltenyi Biotec, 130-096-543) was used for CD8<sup>+</sup> T cells isolation according to manufacturer's instructions. The isolated CD8<sup>+</sup> T cells were plated on anti-CD3 (BD Pharmingen, 553057, 2  $\mu$ g/ml in PBS, coating at 4°C overnight) pre-coated 24-well plates with the concentration of  $1 \times 10^6$  cells/well in 1 ml complete RPMI medium (RPMI with 1X NEAA, 10 mM HEPES, 1 mM sodium pyruvate, 10% FBS, 55 mM 2-Mercaptoethanol, 50 U/ml Penicillin and 50  $\mu$ g/ml Streptomycin) supplemented with 100 U/ml hIL-2 and 2  $\mu$ g/ml anti-CD28 (BD Pharmingen, 553294). Two days after the primary stimulation, 1 ml fresh complete RPMI was added to the same well with 100 U/ml hIL-2. The cultured CD8<sup>+</sup> T cells were split and expanded into larger culture wells or flasks every day from day 3 with fresh RPMI medium with 100 U/ml hIL-2 until day 6. On culture day 6, cells were re-stimulated with 50 ng/ml Phorbol 12-myristate 13-acetate (PMA, Sigma P1585) and 1  $\mu$ M ionomycin (Sigma, I0634) for additional 3 hours and harvested for further experiments.

## RNA sequencing

Total RNA was extracted with QIAGEN RNeasy Plus Mini Kit (QIAGEN, 74134). mRNA was isolated with NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) and the RNA-seq

libraries were generated with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, E7760) according to manufacturer's instructions.

## ChIP sequencing

The detailed ChIP-seq protocol was described (Kurachi et al., 2014). *Cross-linking the cells.* The in vitro cultured WT and Batf KO P14 CD8<sup>+</sup> T cells were collected, and the cell numbers were counted. Cells were suspended in cell culture media with 1% formaldehyde and incubate the cells for 10 min at 37°C. Formaldehyde was quenched with glycine in a final concentration of 125 mM at room temperature for 5 min. The fixed cells were washed 2 times with ice cold PBS and the fixed cell pellets were frozen at -80 °C for storage. *Lyse the cells and shear the chromatin.* The frozen, fixed cells were thawed on ice for 10 min. Add 120 ul of fresh ChIP SDS Lysis Buffer (0.5% SDS, 50 mM Tris, pH 8, 10 mM EDTA, 1 x protease inhibitor cocktail (Roche)) per 5 million cells. Incubated the lysate on ice for 10 min and transferred to Covaris microTUBEs (Covaris, cat# 520045). The Covaris microTUBEs were transferred to the tube rack, placed into the Covaris E220 sonicator and sonicated the samples for 6 treatments (Peak power: 175.0, Duty factor: 10.0, Cycle/Burst: 200, Ave. Power:17.5) of 60 sec for each tube. *Prepare the IP beads.* Prepare the final ChIP buffer by adding 1-part ChIP SDS Lysis Buffer with 4 parts ChIP Dilution Buffer (1.25% Triton X-100, 12.5 mM Tris, pH 8, 187.5 mM NaCl) with protease inhibitors. 30-50 ul of Protein G Dynabeads (Invitrogen, cat# 10004D) were aliquoted into each ChIP tube, one for each ChIP up to 10 million cells. The beads were washed once in 300 ul of ChIP buffer and resuspended the beads in 500 ul of ChIP buffer. 2 to 10 ug of ChIP antibodies were added to the tube and the ChIP tubes were placed on a rotator at 4°C for 3 to 4 hours to bind the antibody to the beads. *Chromatin Immunoprecipitation.* After the samples have been sheared, pooled the lysates together in a microcentrifuge tube. Spin the pooled lysates at 13000 rpm, 4°C for 10 min. The cleared lysate was transferred to a new tube, added 4 parts of ChIP Dilution Buffer (with protease inhibitors) and mixed by pipetting up and down. 100 ul of lysate was aliquoted as an input fraction into another tube and stored in fridge at 4°C overnight. Once the antibodies were bound to the beads, quickly spun down the ChIP bead tubes (5-10 sec spin) and placed the tubes on a magnet rack. Aspirated the entire buffer and aliquoted an appropriate amount of ChIP lysate into the ChIP bead tubes. (Each 600 ul = 5 million cells). Placed the ChIP tubes on a rotator at 4°C overnight. *Prepare Elution Buffer.* Make fresh ChIP Elution Buffer: 1% SDS, 0.1 M sodium bicarbonate. *Washing and Elution of ChIP samples.* Removed the ChIP tubes from the rotator, quickly spun down the tubes, and placed the tubes on a magnet rack. Aspirated the lysate without disturbing the beads and then performed 5 washes for the ChIP beads with 1 ml of the following wash buffers for 3-5 min. Wash 1: ChIP Low Salt Buffer: 0.1% SDS, 1% Triton X-100, 20 mM Tris, pH 8, 2 mM EDTA, 150 mM NaCl. Wash 2: ChIP High Salt Buffer: 0.1% SDS, 1% Triton X-100, 20m M Tris, pH 8, 2 mM EDTA, 500 mM NaCl. Wash 3: ChIP LiCl Buffer: 0.7% sodium deoxycholate, 1% NP-40, 20m M Tris, pH 8, 1 mM EDTA, 500 mM LiCl. Wash 4: ChIP LiCl Buffer. Wash 5: ChIP TE Buffer: 10m M Tris, pH 8, 1 mM EDTA. After the wash steps, eluted the DNA twice with 100 ul ChIP Elution Buffer on a vortex mixer for 30 min. Tubes were removed from the vortex mixer, quickly spun down the tubes, and placed the tubes on a magnet rack. The samples were transferred to new tubes. Retrieved the input tubes from 4°C. Added 5 M NaCl to a final concentration of 0.2 M for each tube (input and ChIP samples). All tubes were placed (input and ChIP samples) on the heat block at 65°C overnight to reverse the formaldehyde cross-linking of the DNA and protein.

**RNA/Protein Degradation.** Removed samples from the heat block, added 2 ul of RNase A (Qiagen), and 2 ul of Proteinase K (RNA grade, Invitrogen/Life Technologies) to each tube. Tubes were placed on the heat block at 37°C for ~2 hrs. **Purify the ChIP DNA fragments.** Removed samples from the heat block, and added 10 ul of 3M sodium acetate, pH 5.5 to each input and ChIP sample. Qiagen MinElute Reaction Cleanup Kit was used to purify the DNA from the input and ChIP samples and eluted the DNA in 23 ul of EB buffer. Determine the concentration of ChIP DNA for library preparation using a Nanodrop. **ChIP-seq library construction.** NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina (NEB cat# E6240L) was used for ChIP-seq library preparation. **End Repair of ChIP DNA.** Added up to 50 ng of ChIP DNA in a maximum of 20 ul to a PCR plate. Added 30 ul of the master mix (5 ul End Repair Reaction Buffer (10X), 24 ul nuclease free water, 1 ul End Repair Enzyme Mix) to each sample of ChIP DNA. Incubated the samples in a thermal cycler for 30 min at 20°C. Added 10 ul of 3M sodium acetate to each sample, purified the samples with Qiagen QIAquick PCR Purification Kit and eluted end-repaired DNA in 34 ul of EB. **dA-Tailing of End-Repaired DNA.** Added 16 ul of the master mix (5 ul dA-Tailing Reaction Buffer (10X), 10 ul nuclease free water, 1 ul Klenow Fragment) to each sample of 34 ul of end-repaired DNA. Incubated the samples in a thermal cycler for 30 min at 37°C. Added 10 ul of 3M sodium acetate to each sample. Use Qiagen MinElute Reaction Cleanup Kit to isolate and elute dA-Tailed DNA in 10 ul of EB. **Adaptor Ligation of dA-Tailed DNA and size-selection.** Diluted the NEBNext Adaptor (15 mM stock) 10X to a 1.5 mM working stock in nuclease free water. Added the 10 ul of dA-Tailed DNA to a PCR plate and added 20 ul of the master mix (6 ul Quick Ligation Reaction Buffer (5X), 1 ul Diluted Adaptor (1.5 mM), 9 ul nuclease free water, 4 ul End Repair Enzyme Mix) to each sample of ChIP DNA. Incubated the samples in a thermal cycler for 15 min at 20°C. Added 3 ul of USER enzyme to each sample and incubated the samples in a thermal cycler for 15 min at 37°C. After the incubation, used AMPure XP beads (Beckman Coulter A63881) for size selection of ChIP-seq DNA. Performed Double Size selection with beads to sample ratio: 0.5 for right side selection and 0.9 for left side selection. Eluted the DNA 20 ul of EB and the DNA is ready for amplification. **PCR Enrichment of Size Selected ChIP-Seq DNA.** NEBNext Multiplex Oligos for Illumina Index Primers Sets were used for indexing ChIP-seq libraries. Added the 2.5 ul of Universal PCR Primer and 2.5 ul index primers to samples. Added 25 ul of NEBNext HotStart Q5 2x PCR Master Mix to each sample and performed PCR with the following conditions: (1) 98°C 30 sec, (2) 98°C 10 sec, (3) 98°C 10 sec, (4) 65°C 30 sec, (5) 72°C 30 sec (6) repeat (3)-(5) 12X, (7) 72°C 5 min, (8) hold at 4 °C. **SPRI Bead Clean Up of ChIP-Seq DNA.** Performed double size selection again with 0.55 for the right-side selection and 1.0 for the left-side selection. Eluted the samples with 40 ul Elution Buffer and examined the size range of your library using TapeStation. Antibodies used in ChIP-seq: Batf (Brookwood Biomedical, PAB4003), JunD (Santa Cruz, 329), Fosl2 (Santa Cruz, Q-20), Irf4 (Santa Cruz M-17), Stat3 (Santa Cruz, C-20), ETS-1 (Santa Cruz, C-20), T-bet (Santa Cruz, H-210), Stat5 (Santa Cruz A-9), Runx3 (Abcam, ab11905) and Eomes (Abcam, ab23345)

## ATAC-seq

50,000 cells were harvested and spun down at 500 g for 5 min, 4°C. The cell pellet was washed once with 50 µl of cold 1X PBS buffer and resuspend in 50 µl cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL CA-630). Spun down immediately at 500 g for

10 min, 4°C. Supernatants were discarded, and proceeded to the transposition reaction. To make the transposition reaction mix, combined the following: 25 µl 2x TD Buffer (Illumina Cat #FC-121-1030), 2.5 µl Tn5 Transposase (Illumina Cat. #FC-121-1030), 22.5 µl nuclease free water to 50 µl of total volume. Resuspended nuclei in the transposition reaction mix and incubate at 37°C for 30 min. Immediately following incubation, purified the DNA using Qiagen MinElute Kit according to the manufacturer's protocol and eluted transposed DNA in 20 µl Elution Buffer (10mM Tris buffer, pH 8). Purified DNA can be stored at -20°C. To amplify transposed DNA fragments and barcoded the libraries, Illumina Nextera Index Kit (FC-121-1011) was used. Combined the following in a PCR tube: 20 µl Transposed DNA, 2.5 µl Nextera PCR Primer with barcode 1, 2.5 µl Nextera PCR Primer with barcode 2, 25 µl NEBNext High-Fidelity 2x PCR Master Mix (New England Labs Cat. #M0541) to 50 µl total volume. PCR cycle as follows: (1) 72°C for 5 min, (2) 98°C 30 sec, (3) 98°C 10 sec, (4) 63°C 30 sec, (5) 72°C, 1 min, (6) Repeat steps 3-5, 12x (7) Hold at 4°C. Used AMPure XP Beads to purify the ATAC-seq libraries with a 1:1 PCR product to Beads ratio and followed manufacturer's instructions. Eluted the purified ATAC-seq libraries in 40 µl Elution Buffer.

### Library quantification and sequencing

KAPA Library Quantification Kit for Illumina Platforms (KAPA Biosystems, KK4824) was used for sequencing library quantification according to the manufacturer's protocol with ABI ViiA 7 Real-Time PCR System. NextSeq 500 V2 High Output Kit (Illumina, FC-404-2005, 75 cycles) and Illumina NextSeq 550 system were used for library sequencing.

### HiChIP

The detailed HiChIP protocol was described (Mumbach et al., 2016).

*Cell crosslinking.* Resuspend every 1 million target cells in 1 ml 1% formaldehyde and incubate cells at room temperature for 10 minutes with rotation. Add glycine (125 mM final) to quench formaldehyde, and then incubate at room temperature for 5 minutes with rotation. Wash the cells once in PBS and then store in -80°C or proceed into the HiChIP protocol. *Lysis and restriction digest.* Resuspend 20 million crosslinked cells in 2 tubes of 500 µl of ice-cold Hi-C Lysis Buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.2% NP-40 and protease inhibitors) and rotate at 4 °C for 30 minutes. Spin down at 2500 g for 5 minutes at 4 °C and discard the supernatant. Wash pelleted nuclei once with 500 µl of ice-cold Hi-C Lysis Buffer. Remove the supernatant and resuspend the pellet in 100 µl of 0.5% SDS. Incubate at 62 °C for 10 minutes and add 285 µl of H<sub>2</sub>O and 50 µl of 10% Triton X-100 to quench the SDS. Mix well and incubate at 37 °C for 15 minutes with rotation. Add 50 µl of 10X NEB Buffer 2 and 375 U (15 µl of 25 U/ µl) of MboI restriction enzyme (NEB, R0147), and digest chromatin for 2 hours at 37°C with rotation. Heat inactivated MboI at 62 °C for 20 minutes. *Incorporation and proximity ligation.* To fill in the restriction fragment overhangs and mark the DNA ends with biotin, add 52 µl of fill-in master mix: 0.4 mM biotin-dATP (Thermo 19524016) 37.5 µl, 10 mM dCTP 1.5 µl, 10 mM dGTP 1.5 µl, 10 mM dTTP 1.5 µl, 5U/ µl DNA Polymerase I, Large (Klenow) Fragment (NEB, M0210) 10 µl. Mix and incubate at 37 °C for 1 hour with rotation. Add 948 µl of ligation master mix: 150 µl 10X NEB T4 DNA ligase buffer (NEB, B0202), 125 µl 10% Triton X-100, 3 µl 50 mg/mL BSA (Thermo Fisher, AM2616), 10 µl 400 U/ µl T4 DNA Ligase (NEB, M0202), 660 µl nuclease free water. Incubate at room temperature for 4 hours with rotation. Pellet nuclei at 2500 g for 5 minutes and remove supernatant. *Sonication.*



For Batf and Ctcf HiChIP, pool the two sample tubes into 880 µl Nuclear Lysis Buffer and proceed to the sonication step. Transfer to a Covaris milliTUBE (Covaris #520135) and shear with Covaris E220 with the following program: Fill Level 5, Duty Cycle 5, PIP 140, Cycles/Burst 200, Time 4 minutes. Preclearing, immunoprecipitation, IP bead capture, and washes. Centrifuge sample for 15 minutes at 16000 g at 4°C. Add 3320 µl Dilution Buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 7.5, 167 mM NaCl) to the 880 µl sonicated sample and split into 4 tubes, 1050 µl per tube (~5 million cells). Wash 30 µl of Protein G Dynabeads (Invitrogen, 10004D) for every 5 million cells in 50 µl ChIP Dilution Buffer. Resuspend Protein G Dynabeads in 50 µl of Dilution Buffer per tube (100 µl per HiChIP), add to sample and rotate at 4 °C for 1 hour. Put samples on magnet and transfer supernatant into new tubes. Add 2.5 µg of antibody for every 5 million cells and incubate at 4°C overnight with rotation. Wash 30 µl of Protein G Dynabeads for every 5 million cells in ChIP Dilution Buffer. Resuspend Protein G Dynabeads in 50 µl of Dilution Buffer (100 µl per HiChIP), add to sample and rotate at 4 for 2 hours. Combine all the split samples and wash beads four times each with Low Salt Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 7.5, 150 mM NaCl), High Salt Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 7.5, 500 mM NaCl), and twice with the LiCl Wash Buffer (10 mM Tris-HCl pH 7.5, 250 mM LiCl, 1% NP-40, 1% Sodium deoxycholate, 1 mM EDTA). Washing was performed at room temperature on a magnet by adding 500 µl of a wash buffer, swishing the beads back and forth twice by moving the sample relative to the magnet, and then removing the supernatant. *ChIP DNA elution.* Resuspend ChIP sample beads in 100 µl of fresh DNA Elution Buffer (50 mM NaHCO<sub>3</sub>, 1% SDS). Incubate at room temperature for 10 minutes with rotation, followed by 3 minutes at 37 °C shaking. For ChIP samples, place on magnet and remove supernatant to a fresh tube. Add another 100 µl of DNA Elution Buffer to ChIP samples and repeat incubations. Remove ChIP samples supernatant again to the new tube. There should now be 200 µl of ChIP sample. Add 10 µl of Proteinase K and 8 µl of 5M NaCl to each sample and incubate at 55 °C for 45 minutes. Increase temperature to 67 and incubate for at least 1.5 hours with shaking. Use QIAGEN minElute Kit to purify the samples and elute in 12 µl of water. Quantify post-ChIP DNA with Quant-iT dsDNA HS Kit (ThermoFisher, Q33120). For libraries with greater than 150 ng of post-ChIP DNA, set aside material and take a maximum of 150 ng into the biotin capture step. *Biotin pull-down and preparation for Illumina sequencing.* Prepare for biotin pull-down by washing 5 µl of Streptavidin C-1 (Invitrogen, 65001) beads with Tween Wash Buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20). Resuspend the beads in 10 µl of 2X Biotin Binding Buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl) and add to the samples. Incubate at room temperature for 15 minutes with rotation. Separate on a magnet and discard the supernatant. Wash the beads twice by adding 500 µl of Tween Wash Buffer and incubating at 55 °C for 2 minutes with shaking. Wash the beads in 100 µl of 1X TD Buffer (10 mM Tris-HCl pH 7.5, 5 mM MgCl<sub>2</sub>, 10% Dimethylformamide). Resuspend beads in 25 µl of 2X TD Buffer (20 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 20% dimethylformamide). For Batf and Ctcf HiChIP, use 2.5 µl Tn5 transposase for 50 ng post-ChIP DNA and add water to 50 µl. Incubate at 55 °C with interval shaking for 10 minutes. Place samples on magnet and remove supernatant. Add 50 mM EDTA to samples and incubate at 50 °C for 30 minutes, then quickly place on magnet and remove supernatant. Wash samples twice with 100 µl 50 mM EDTA at 50 °C for 3 minutes and wash samples twice in 100 µl Tween Wash Buffer at 55 °C for 2 minutes. And then wash samples in 100 µl 10 mM Tris. Resuspend beads in 20 µl Elution Buffer. PCR and Post-PCR Size

Selection. Add 30  $\mu$ l PCR master mix (2.5  $\mu$ l Illumina dual indexes primer 1, 2.5  $\mu$ l Illumina dual index primer 2, 25  $\mu$ l NEB Q5 HotStart Master Mix) to the 20  $\mu$ l beads bound with samples. For 50 ng post-ChIP DNA, Run PCR program: (1) 72°C 5 minutes, (2) 98°C 1 minute, (3) 98°C 15 seconds, (4) 63°C 30 seconds, (5) 72°C 1 minute, (6) repeat (3)-(5) 5X, (7) hold at 4°C. *Purify and size-select the PCR product with Ampure XP beads.* After PCR, place libraries on a magnet and elute into new tubes (~50  $\mu$ l). Then add 25  $\mu$ l of Ampure XP beads and keep the supernatant to capture fragments less than 700 bp. Transfer supernatant to a new tube and add 15  $\mu$ l of fresh beads to capture fragments greater than 300 bp. Eluted in 22  $\mu$ l Elution Buffer and the samples were ready to be quantified and sequencing. Antibodies used in HiChIP experiments: Batf (Brookwood Biomedical, PAB4003) and Ctcf (Millipore, 07729MI).

### **Lentiviral vector construction and lentivirus production**

The coding domain sequence (CDS) of mouse Batf, Irf4, Runx3 and T-bet were PCR amplified and cloned into pDONR221 vector to generate entry clones pENTR221-Batf, pENTR221-Irf4, pENTR221-Runx3 and pENTR221-T-bet respectively. The CDS of Batf, Irf4, Runx3 and T-bet from the entry clones were then cloned into the pLoxp404 lentiviral expression vectors (Broad Institute) pLoxp404-HygroR, pLoxp404-PuroR, pLoxp404-BlastR and pLoxp404-GFP by Gateway Cloning (Thermo Fisher Scientific) to generate final lentiviral constructs pLoxp404-HygroR-Batf, pLoxp404-PuroR-Irf4, pLoxp404-BlastR-Runx3 and pLoxp404-GFP-T-bet. The lentivirus was packaged in HEK 293T cells by co-transfecting the lentiviral vectors with pCMV-VSV-G (envelope) and pCMV-dR8.91 (gag-pol) plasmids. The supernatants containing lentivirus were harvested and used for NIH/3T3 cell line transduction. The GFP-T-bet lentivirus was concentrated with ultracentrifugation and resuspended in medium as 50X concentrated virus for transduction.

### **Lentiviral transduction, antibiotics selection and doxycycline induction of NIH/3T3 cells**

24 hours before transduction, seed NIH/3T3 cells in 6-well plates,  $2 \times 10^5$  cells/well in 2 ml DMEM (10% FBS) medium. On the day of transduction, replace with fresh medium and add 300  $\mu$ l lentiviral supernatant with 5  $\mu$ g/ml polybrene. After overnight incubation, replace the virus containing medium with fresh DMEM and incubated for an additional 48 hours before adding antibiotics. 200  $\mu$ g/ml Hygromycin B, 10  $\mu$ g/ml Blasticidin, 7.5  $\mu$ g/ml Puromycin were added to corresponding transduced 3T3 cells and selection for 1 to 2 weeks for generating stable clones. The transduction efficiency of GFP-T-bet lentivirus transduced 3T3 cells was determined by flow cytometry for GFP expression and GFP expressing cells were sorted if necessary. 4  $\mu$ g/ml doxycycline was used to induce the gene expressions in transduced 3T3 fibroblasts. The cells were treated for 72 hours and the culture medium was changed every 24 hours with fresh doxycycline.

### **Western blot**

Fibroblasts were trypsinized, washed with cold PBS and lysed in RIPA lysis buffer (Thermo Scientific) on ice for 30 min for extracting protein lysates. Samples were centrifuged >13,000 rcf at 4 °C for 10 min, supernatants were collected and protein concentrations were quantified with Bio-Rad protein assay (Bio-Rad). 20-40  $\mu$ g of lysates were mixed with LDS sample buffer



(ThermoFisher) and heated at 70 degrees for protein denaturation according to the manufacturer's protocol. Samples were run in Bolt 4-12% Bis-Tris Plus Gels (ThermoFisher) and transferred to PVDF membrane (Millipore). The following primary antibodies were used to detect designated proteins: Batf (Brookwood Biomedical, PAB4003), Irf4 (Santa Cruz, M-17), Runx3 (Abcam, ab11905), T-bet (Santa Cruz, H-210), TGFb1 (Abcam, EPR18163) and Hsp90 (Santa Cruz, F-8). HRP conjugated secondary antibodies (Goat anti-rabbit IgG and Goat anti-mouse IgG light chain) were purchased from Jackson ImmunoResearch Laboratories. HRP substrates (Immobilon western chemiluminescent HRP substrate, Millipore) were applied to the membrane according to manufacturer's protocol and imaged using Fujifilm LAS4000 imaging system.

### **Flow cytometry**

Fibroblasts were trypsinized (Trypsin-EDTA 0.25%, Gibco), washed with cold PBS and resuspended in flow cytometry staining buffer (eBioscience). The cells were stained at 4 degrees for 30 mins with the following fluorochrome-conjugated antibodies: APC-IL17Ra (eBioscience, clone PAJ-17R) and APC-CD197(CCR7) (BioLegend, clone 4B12). Samples were washed with flow cytometry staining buffer and analyzed on LSRFortessa (BD). The data were analyzed with FlowJo software (FlowJo).

### **Alignment of RNA-seq and quantification of transcript abundance**

All RNA-Seq reads were trimmed using Trimmomatic (Bolger et al., 2014) to remove primer and low-quality bases. Reads were then passed to FastQC [<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>] to check the quality of the trimmed reads. The single-end reads were then aligned to the Mus\_musculus.GRCm38.82 transcriptome from Ensembl using RSEM (Li and Dewey, 2011) with the parameters “--num-threads 4 --bowtie2 --sampling-for-bam --output-genome-bam --sort-bam-by-coordinate --sort-bam-memory-per-thread 1G --estimate-rspd --fragment-length-mean 200 --fragment-length-sd 80’.

### **Analysis of differential gene expression**

Differential expressions were assessed using DESeq2 (Love et al., 2014). Genes were considered differentially expressed if they had an adjusted p-value < 0.05. ‘betaPrior’ was set equal to FALSE.

### **Analysis of time course in vitro RNA-Seq**

Using DESeq2 (Love et al., 2014), we conducted a likelihood ratio test to identify any genes that changed over time (adjusted p value <0.05). We then clustered the genes using the pheatmap package in R with correlation as the row distance, and hierarchical complete clustering (default settings), cutting the tree at the point at which there were five clusters.

### **Alignment of ATAC-seq and peak calling**

All ATAC-seq reads were trimmed using Trimmomatic (Bolger et al., 2014) to remove primer and low-quality bases. Reads were then passed to FastQC (Andrews, 2010) to check the quality of the trimmed reads. The paired-end reads were then aligned to the mm10 reference genome using bowtie2 (Langmead and Salzberg, 2012), allowing maximum insert sizes of 2000 bp, with the “--no-mixed” and “--no-discordant” parameters added. Reads with a mapping quality (MAPQ) below

30 were removed. Duplicates were removed with PicardTools , and the reads mapping to the blacklist regions and mitochondrial DNA were also removed. Reads mapping to the positive strand were moved +4 bp, and reads mapping to the negative strand were moved -5bp following the procedure outlined by Buenrostro et al. (Buenrostro et al., 2013) to account for the binding of the *Tn5* transposase.

Peaks were called using macs2 on the aligned fragments (Zhang et al., 2008) with a q-value cutoff of 0.05 and overlapping peaks among replicates were merged. A region was considered a valid peak if it had a q-value below 0.05 in at least two replicates, and a qvalue below 0.001 in at least one of the replicates.

### **Tests of differential accessibility in ChARs**

Differential accessibility was assessed by applying DESeq2 (Love et al., 2014) to a matrix with rows as peaks, columns as samples, and counts of Tn5 cuts in each cell. Peaks were considered differentially accessible if they had an adjusted p-value < 0.05. 'betaPrior' was set equal to FALSE.

### **Alignment of ChIP-seq and peak calling**

All ChIP-seq reads were trimmed using Trimmomatic (Bolger et al., 2014) to remove primer and low-quality bases. Reads were then passed to FastQC (Andrews, 2010) to check the quality of the trimmed reads. These single-end reads were then aligned to the mm10 reference genome using bowtie2 (Langmead and Salzberg, 2012), allowing maximum insert size of 2000 bp, with the "--no-mixed" and "--no-discordant" parameters added. Reads with a mapping quality (MAPQ) below 30 were removed. Duplicates were removed with Picard Tools , and the reads mapping to the blacklist regions and mitochondrial DNA were also removed.

Each combination of transcription factor and condition had two replicates, with the exception of Runx3-WT, which we limited to one high-quality replicate. ChIP-Seq peaks were called in each replicate, versus a control sample, using macs2 (Zhang et al., 2008) with a q-value cutoff of 0.05 and overlapping peaks among replicates were merged. A region was considered a valid peak for a transcription factor if it had a q-value below 0.05 in at least two replicates, and a qvalue below 0.001 in at least one of the replicates.

### **Tests of differential binding**

Differential binding was assessed using the csaw package (Lun and Smyth, 2016) in regions identified as being bound in either the WT or *Batf*<sup>-/-</sup> for the given transcription factor. Only log2 fold changes greater than one with an adjusted p-value below 0.05 were considered differentially bound.

### **Universe of transcription factor bound regions and accessible peaks**

The universe of transcription factor bound regions and accessible peaks was created by merging the overlapping ChIP-seq and ATAC-seq peaks previously described. ATAC-seq peaks from the experiments in vivo flox experiments were used, and All WT and Batf KO ChIP-seq peaks previously described were used. This dataset is available as [Table 2](#).

## **Test for Association of transcription factor binding with reduced accessibility of ChARs in Batf cKO**

We used the 51,259 regions from Table 2 where an ATAC-seq peak was present in the WT (floxed) or Batf cKO condition. We used the results from the test of differential accessibility (described above) to code each peak as having reduced accessibility in Batf cKO (1) or not (0). We then ran a separate logistic regression for each of our TF's (coded 1 if bound, 0 if not bound) to test for its association with reduced accessibility in Batf cKO. All TF's had a Bonferroni corrected p-value <0.001.

## **Fold enrichment of transcription factor binding**

Fold-enrichment of individual transcription factors is shown on the diagonal (barplot) of Figure 2F, and fold-enrichment for co-binding of transcription factors is shown on off-diagonal cells. X = locus bound by column transcription factor, Y = locus bound by row transcription factor, and L = locus less accessible in Batf KO. Fold-enrichment is calculated as:

$$[(\text{loci } X \& Y \& L) / (\text{loci } X \& Y)] / \max [(\text{loci } X \& L) / (\text{loci } X), (\text{loci } Y \& L) / (\text{loci } Y)]$$

## **Alignment of HiChIP data and peak calling**

HiChIP data were aligned using the HiCPro pipeline (Servant et al., 2015). Normalized counts were plotted using the HiCPlotter python package (Akdemir and Chin, 2015).

The Hichipper package (Lareau and Aryee, 2018) was used to identify loops in the HiChIP data. The Hichipper package uses the Mango package (Phanstiel et al., 2015) to test the statistical significance of loops. We retained loops with an adjusted p-value <0.001 and at least 5 paired-end tags linking the two anchors.

## **Gene-level enrichment analysis**

Gene Set Enrichment Analysis was conducted with clusterProfiler (Yu et al., 2012) which utilizes fgsea (Sergushichev, 2016). The gsea, gseGO, and gseKEGG functions were run with 1000 permutations and an adjusted p-value cutoff of 0.10. Results are provided for the ImmSigDB C2 and C7 collections (min set size = 10, max set size = 1000), GO Biological Processes and Molecular Functions (min set size = 100, max set size = 500), and KEGG (min set size = 100, max set size = 500). The "simplify" function was run on the GO results to condense them.

The Wald statistic from DESeq2 was used to order the genes. Mouse ensembl gene id's were mapped to human orthologs using Biomart (Durinck et al., 2005, 2009) for the C2 and C7 collections. For instances in which multiple mouse genes mapped to one human ortholog, the average of the Wald statistics was used. Genes which were knocked out or ectopically expressed were removed prior to GSEA to avoid biasing the results.

## **T cell score calculation and Shapley value**

T cell gene expression signatures were obtained from DESeq2 comparison of RNA-seq of effector and naive CD8+ T cells vs fibroblasts. Signature was defined by giving a weight of +1 to the top 1000 genes differentially upregulated in T cells and weight of -1 to the top 1000 genes differentially upregulated in fibroblasts. Weights were then applied to the vector of normalized counts for each condition. Shapley value for each TF was calculated based on the relative T cell score. Bootstrap

estimates were calculated by computing the score using one randomly chosen replicate from each condition, with 2,500 iterations.

### **Locus- level enrichment analysis**

Subsets of loci deemed “foregrounds” were analyzed for fold enrichment of features relative to a background set of loci. The features analyzed were motifs and other genomic annotations, SNPs associated with immune diseases, proximity to gene groups of interest, and pathways in GREAT. q-values were estimated using the q-value package (Storey JD, Bass AJ, Dabney A, Robinson D, 2019).

### **Motifs / annotation tracks**

PWM’s for motifs were downloaded from the 2018 release of JASPAR (Khan et al., 2017). We used fimo (Grant et al., 2011) to identify motifs in mm10, and applied the default threshold of 1e-4. We also included regulatory features from the ORegAnno database (Lesurf et al., 2016), (iii) conserved regions annotated by the multiz30way (Blanchette et al., 2004) algorithm, and repeat regions annotated by RepeatMasker (Smit et al., 1996).

### **SNPs**

SNPs associated with different immune diseases were retrieved from the experimental factor ontology (Malone et al., 2010) by downloading all SNPs in the GWAS Catalog (Buniello et al., 2019) corresponding to a given node and its children. The orthologous regions in mice were identified by using liftOver (Hinrichs et al., 2006).

### **GREAT pathways / genes**

Loci were associated with pathways using GREAT (McLean et al., 2010), submitted with the rGREAT package (Gu, 2015). We retrieved pathways found in the MSigDB Immunologic Signatures, MSigDB Pathway, and GO Biological Process databases. Loci were mapped to genes using GREAT.

# **ACKNOWLEDGMENTS**

We thank Dr. Robert Manguso and Kathleen Yates of the Haining laboratory and members of Yosef laboratory for comments and suggestions, the Birgit Knoechel laboratory of Dana-Farber Cancer Institute for sequencing assistance, and the Hematologic Neoplasia and Jimmy Fund Flow Cytometry Cores of Dana-Farber Cancer Institute for technical support. This work was supported by National Institutes of Health (R01AI115712 to H.W.N. and W.E.J.), and the “Postdoctoral Research Abroad Program” of the Ministry of Science and Technology, Taiwan (MOST 104-2917-I-564-018 to H.W.T.)

# **Author Contributions**

H.W.T., J.K., W.N.H. and N.Y. designed the research and wrote the manuscript.

H.W.T. performed most of the experiments.

J.K. performed most of the computational analyses.

M.K. conducted the in vivo LCMV challenge and sorted the cells for this study.

R.A.B., M.A.D. and M.W.L. assisted the ChIP-seq and ectopic TF expression experiments.

W.I. and T.K. provided the critical mouse model for this study.

N.Y., W.N.H. and E.J.W. supervised the study.

# REFERENCES

- Acharya, D., Wang, P., Paul, A.M., Dai, J., Gate, D., Lowery, J.E., Stokic, D.S., Leis, A.A., Flavell, R.A., Town, T., et al. (2017). Interleukin-17A Promotes CD8+ T Cell Cytotoxicity To Facilitate West Nile Virus Clearance. *J. Virol.* 91.
- Akdemir, K.C., and Chin, L. (2015). HiCPlotter integrates genomic data with interaction matrices. *Genome Biol.* 16, 198.
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.
- Arbonés, M.L., Ord, D.C., Ley, K., Ratech, H., Maynard-Curry, C., Otten, G., Capon, D.J., and Tedder, T.F. (1994). Lymphocyte homing and leukocyte rolling and migration are impaired in L-selectin-deficient mice. *Immunity* 1, 247–260.
- Betz, B.C., Jordan-Williams, K.L., Wang, C., Kang, S.G., Liao, J., Logan, M.R., Kim, C.H., and Taparowsky, E.J. (2010). Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. *J. Exp. Med.* 207, 933–942.
- Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F.A., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14, 708–715.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bolig, N., Brüstle, A., Kellner, K., Ackermann, W., Abass, E., Raifer, H., Camara, B., Brendel, C., Giel, G., Bothur, E., et al. (2012). Transcription factor IRF4 determines germinal center formation through follicular T-helper cell differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8664–8669.
- Bouillet, P., Purton, J.F., Godfrey, D.I., Zhang, L.-C., Coultas, L., Puthalakath, H., Pellegrini, M., Cory, S., Adams, J.M., and Strasser, A. (2002). BH3-only Bcl-2 family member Bim is required for apoptosis of autoreactive thymocytes. *Nature* 415, 922–926.
- Brender, C., Tannahill, G.M., Jenkins, B.J., Fletcher, J., Columbus, R., Saris, C.J.M., Ernst, M., Nicola, N.A., Hilton, D.J., Alexander, W.S., et al. (2007). Suppressor of cytokine signaling 3 regulates CD8 T-cell proliferation by inhibition of interleukins 6 and 27. *Blood* 110, 2528–2536.
- Brüstle, A., Heink, S., Huber, M., Rosenplänter, C., Stadelmann, C., Yu, P., Arpaia, E., Mak, T.W., Kamradt, T., and Lohoff, M. (2007). The development of inflammatory T(H)-17 cells requires interferon-regulatory factor 4. *Nat. Immunol.* 8, 958–966.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218.
- Buggert, M., Tauriainen, J., Yamamoto, T., Frederiksen, J., Ivarsson, M.A., Michaëlsson, J., Lund, O., Hejdeman, B., Jansson, M., Sönnernborg, A., et al. (2014). T-bet and Eomes are differentially linked to the exhausted phenotype of CD8+ T cells in HIV infection. *PLoS Pathog.* 10, e1004251.



Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 47, D1005–D1012.

Chan, J.C., and Maze, I. (2020). Nothing Is yet Set in (Hi)stone: Novel Post-Translational Modifications Regulating Chromatin Function. *Trends in Biochemical Sciences*.

Chang, J.T., Wherry, E.J., and Goldrath, A.W. (2014). Molecular regulation of effector and memory T cell differentiation. *Nat. Immunol.* 15, 1104–1115.

Ciofani, M., Madar, A., Galan, C., Sellars, M., Mace, K., Pauli, F., Agarwal, A., Huang, W., Parkhurst, C.N., Muratet, M., et al. (2012). A validated regulatory network for Th17 cell specification. *Cell* 151, 289–303.

Cruz-Guilloty, F., Pipkin, M.E., Djuretic, I.M., Levanon, D., Lotem, J., Lichtenheld, M.G., Groner, Y., and Rao, A. (2009). Runx3 and T-box proteins cooperate to establish the transcriptional program of effector CTLs. *J. Exp. Med.* 206, 51–59.

Cui, W., Liu, Y., Weinstein, J.S., Craft, J., and Kaech, S.M. (2011). An interleukin-21-interleukin-10-STAT3 pathway is critical for functional maturation of memory CD8<sup>+</sup> T cells. *Immunity* 35, 792–805.

Curtsinger, J.M., Agarwal, P., Lins, D.C., and Mescher, M.F. (2012). Autocrine IFN- $\gamma$  promotes naive CD8 T cell differentiation and synergizes with IFN- $\alpha$  to stimulate strong function. *J. Immunol.* 189, 659–668.

D'Alonzo, R.C., Selvamurugan, N., Karsenty, G., and Partridge, N.C. (2002). Physical interaction of the activator protein-1 factors c-Fos and c-Jun with Cbfa1 for collagenase-3 promoter activation. *J. Biol. Chem.* 277, 816–822.

Danilo, M., Chennupati, V., Silva, J.G., Siegert, S., and Held, W. (2018). Suppression of Tcf1 by Inflammatory Cytokines Facilitates Effector CD8 T Cell Differentiation. *Cell Rep.* 22, 2107–2117.

DeTomaso, D., Jones, M.G., Subramaniam, M., Ashuach, T., Ye, C.J., and Yosef, N. (2019). Functional interpretation of single cell similarity maps. *Nat. Commun.* 10, 4376.

Djuretic, I.M., Levanon, D., Negreanu, V., Groner, Y., Rao, A., and Ansel, K.M. (2007). Transcription factors T-bet and Runx3 cooperate to activate Ifng and silence Il4 in T helper type 1 cells. *Nat. Immunol.* 8, 145–153.

Dong, C., Juedes, A.E., Temann, U.A., Shresta, S., Allison, J.P., Ruddle, N.H., and Flavell, R.A. (2001). ICOS co-stimulatory receptor is essential for T-cell activation and function. *Nature* 409, 97–101.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 21, 3439–3440.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* 4,



1184–1191.

Filén, S., Ylikoski, E., Tripathi, S., West, A., Björkman, M., Nyström, J., Ahlfors, H., Coffey, E., Rao, K.V.S., Rasool, O., et al. (2010). Activating transcription factor 3 is a positive regulator of human IFNG gene expression. *J. Immunol.* *184*, 4990–4999.

Gaffen, S.L. (2009). Structure and signalling in the IL-17 receptor family. *Nat. Rev. Immunol.* *9*, 556–567.

Gerdes, J., Lemke, H., Baisch, H., Wacker, H.H., Schwab, U., and Stein, H. (1984). Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J. Immunol.* *133*, 1710–1715.

Glasmacher, E., Agrawal, S., Chang, A.B., Murphy, T.L., Zeng, W., Vander Lugt, B., Khan, A.A., Ciofani, M., Spooner, C.J., Rutz, S., et al. (2012). A genomic regulatory element that directs assembly and function of immune-specific AP-1-IRF complexes. *Science* *338*, 975–980.

Godec, J., Cowley, G.S., Barnitz, R.A., Alkan, O., Root, D.E., Sharpe, A.H., and Haining, W.N. (2015). Inducible RNAi in vivo reveals that the transcription factor BATF is required to initiate but not maintain CD8<sup>+</sup> T-cell effector differentiation. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 512–517.

Godec, J., Tan, Y., Liberzon, A., Tamayo, P., Bhattacharya, S., Butte, A.J., Mesirov, J.P., and Haining, W.N. (2016). Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation. *Immunity* *44*, 194–206.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* *27*, 1017–1018.

Grenningloh, R., Tai, T.-S., Frahm, N., Hongo, T.C., Chicoine, A.T., Brander, C., Kaufmann, D.E., and Ho, I.-C. (2011). Ets-1 maintains IL-7 receptor expression in peripheral T cells. *J. Immunol.* *186*, 969–976.

Grusdat, M., McIlwain, D.R., Xu, H.C., Pozdeev, V.I., Knievel, J., Crome, S.Q., Robert-Tissot, C., Dress, R.J., Pandya, A.A., Speiser, D.E., et al. (2014). IRF4 and BATF are critical for CD8<sup>+</sup> T-cell function following infection with LCMV. *Cell Death Differ.* *21*, 1050–1060.

Gu, Z. (2015). rGREAT: client for GREAT analysis.

Hammami, A., Charpentier, T., Smans, M., and Stäger, S. (2015). IRF-5-Mediated Inflammation Limits CD8<sup>+</sup> T Cell Expansion by Inducing HIF-1 $\alpha$  and Impairing Dendritic Cell Functions during Leishmania Infection. *PLoS Pathog.* *11*, e1004938.

Hida, S., Ogasawara, K., Sato, K., Abe, M., Takayanagi, H., Yokochi, T., Sato, T., Hirose, S., Shirai, T., Taki, S., et al. (2000). CD8(+) T cell-mediated skin disease in mice lacking IRF-2, the transcriptional attenuator of interferon-alpha/beta signaling. *Immunity* *13*, 643–655.

Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F., et al. (2006). The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* *34*, D590–D598.

Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S.H., Xu, J., and Yuan, G.-C. (2018).

Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat. Commun.* 9, 943.

Intlekofer, A.M., Takemoto, N., Wherry, E.J., Longworth, S.A., Northrup, J.T., Palanivel, V.R., Mullen, A.C., Gasink, C.R., Kaech, S.M., Miller, J.D., et al. (2005). Effector and memory CD8<sup>+</sup> T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* 6, 1236–1244.

Ise, W., Kohyama, M., Schraml, B.U., Zhang, T., Schwer, B., Basu, U., Alt, F.W., Tang, J., Oltz, E.M., Murphy, T.L., et al. (2011). The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat. Immunol.* 12, 536–543.

Jabeen, R., Goswami, R., Awe, O., Kulkarni, A., Nguyen, E.T., Attenasio, A., Walsh, D., Olson, M.R., Kim, M.H., Tepper, R.S., et al. (2013). Th9 cell development requires a BATF-regulated transcriptional network. *J. Clin. Invest.* 123, 4641–4653.

Jung, Y.W., Kim, H.G., Perry, C.J., and Kaech, S.M. (2016). CCR7 expression alters memory CD8 T-cell homeostasis by regulating occupancy in IL-7- and IL-15-dependent niches. *Proc. Natl. Acad. Sci. U. S. A.* 113, 8278–8283.

Kaech, S.M., and Cui, W. (2012). Transcriptional control of effector and memory CD8<sup>+</sup> T cell differentiation. *Nat. Rev. Immunol.* 12, 749–761.

Kao, C., Oestreich, K.J., Paley, M.A., Crawford, A., Angelosanto, J.M., Ali, M.-A.A., Intlekofer, A.M., Boss, J.M., Reiner, S.L., Weinmann, A.S., et al. (2011). Transcription factor T-bet represses expression of the inhibitory receptor PD-1 and sustains virus-specific CD8<sup>+</sup> T cell responses during chronic infection. *Nat. Immunol.* 12, 663–671.

Kayama, H., Tani, H., Kitada, S., Opasawatchai, A., Okumura, R., Motooka, D., Nakamura, S., and Takeda, K. (2019). BATF2 prevents T-cell-mediated intestinal inflammation through regulation of the IL-23/IL-17 pathway. *Int. Immunol.* 31, 371–383.

Keinan, A., Sandbank, B., Hilgetag, C.C., Meilijson, I., and Ruppin, E. (2004). Fair attribution of functional contribution in artificial and biological networks. *Neural Comput.* 16, 1887–1915.

Khan, A., Fomes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., et al. (2017). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.*

Klein, U., Casola, S., Cattoretti, G., Shen, Q., Lia, M., Mo, T., Ludwig, T., Rajewsky, K., and Dalla-Favera, R. (2006). Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat. Immunol.* 7, 773–782.

Kujawski, M., Zhang, C., Herrmann, A., Reckamp, K., Scuto, A., Jensen, M., Deng, J., Forman, S., Figlin, R., and Yu, H. (2010). Targeting STAT3 in adoptively transferred T cells promotes their in vivo expansion and antitumor effects. *Cancer Res.* 70, 9599–9610.

Kurachi, M., Barnitz, R.A., Yosef, N., Odorizzi, P.M., Dilorio, M.A., Lemieux, M.E., Yates, K., Godec, J., Klatt, M.G., Regev, A., et al. (2014). The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8<sup>+</sup> T cells. *Nat. Immunol.* 15, 373–383.

Kuroda, S., Yamazaki, M., Abe, M., Sakimura, K., Takayanagi, H., and Iwai, Y. (2011). Basic

leucine zipper transcription factor, ATF-like (BATF) regulates epigenetically and energetically effector CD8 T-cell differentiation via Sirt1 expression. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14885–14889.

Kuwahara, M., Ise, W., Ochi, M., Suzuki, J., Kometani, K., Maruyama, S., Izumoto, M., Matsumoto, A., Takemori, N., Takemori, A., et al. (2016). Bach2-Batf interactions control Th2-type immune response by regulating the IL-4 amplification loop. *Nat. Commun.* **7**, 12596.

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

Lareau, C.A., and Aryee, M.J. (2018). hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. *Nat. Methods* **15**, 155–156.

Laurence, A., Tato, C.M., Davidson, T.S., Kanno, Y., Chen, Z., Yao, Z., Blank, R.B., Meylan, F., Siegel, R., Hennighausen, L., et al. (2007). Interleukin-2 signaling via STAT5 constrains T helper 17 cell generation. *Immunity* **26**, 371–381.

Lesurf, R., Cotto, K.C., Wang, G., Griffith, M., Kasaian, K., Jones, S.J.M., Montgomery, S.B., Griffith, O.L., and Open Regulatory Annotation Consortium (2016). ORegAnno 3.0: a community-driven resource for curated regulatory annotation. *Nucleic Acids Res.* **44**, D126–D132.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323.

Li, D., Hsu, S., Purushotham, D., Sears, R.L., and Wang, T. (2019). WashU Epigenome Browser update 2019. *Nucleic Acids Res.* **47**, W158–W165.

Li, P., Spolski, R., Liao, W., Wang, L., Murphy, T.L., Murphy, K.M., and Leonard, W.J. (2012). BATF-JUN is critical for IRF4-mediated transcription in T cells. *Nature* **490**, 543–546.

Lin, J.X., Migone, T.S., Tsang, M., Friedmann, M., Weatherbee, J.A., Zhou, L., Yamauchi, A., Bloom, E.T., Mietz, J., and John, S. (1995). The role of shared receptor motifs and common Stat proteins in the generation of cytokine pleiotropy and redundancy by IL-2, IL-4, IL-7, IL-13, and IL-15. *Immunity* **2**, 331–339.

Lingel, H., Wissing, J., Arra, A., Schanze, D., Lienenklaus, S., Klawonn, F., Pierau, M., Zenker, M., Jänsch, L., and Brunner-Weinzierl, M.C. (2017). CTLA-4-mediated posttranslational modifications direct cytotoxic T-lymphocyte differentiation. *Cell Death Differ.* **24**, 1739–1749.

Liu, C.F., Brandt, G.S., Hoang, Q.Q., Naumova, N., Lazarevic, V., Hwang, E.S., Dekker, J., Glimcher, L.H., Ringe, D., and Petsko, G.A. (2016). Crystal structure of the DNA binding domain of the transcription factor T-bet suggests simultaneous recognition of distant genome sites. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6572–E6581.

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

Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* **44**, e45.

- Lund, R.J., Löytömäki, M., Naumanen, T., Dixon, C., Chen, Z., Ahlfors, H., Tuomela, S., Tahvanainen, J., Scheinin, J., Henttinen, T., et al. (2007). Genome-wide identification of novel genes involved in early Th1 and Th2 cell differentiation. *J. Immunol.* 178, 3648–3660.
- Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., Zhukova, A., Brazma, A., and Parkinson, H. (2010). Modeling sample variables with an Experimental Factor Ontology. *Bioinformatics* 26, 1112–1118.
- Man, K., Gabriel, S.S., Liao, Y., Gloury, R., Preston, S., Henstridge, D.C., Pellegrini, M., Zehn, D., Berberich-Siebelt, F., Febbraio, M.A., et al. (2017). Transcription Factor IRF4 Promotes CD8 T Cell Exhaustion and Limits the Development of Memory-like T Cells during Chronic Infection. *Immunity* 47, 1129–1141.e5.
- Martinet, V., Tonon, S., Torres, D., Azouz, A., Nguyen, M., Kohler, A., Flamand, V., Mao, C.-A., Klein, W.H., Leo, O., et al. (2015). Type I interferons regulate eomesodermin expression and the development of unconventional memory CD8<sup>+</sup> T cells. *Nat. Commun.* 6, 7089.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28, 495–501.
- Moriggl, R., Topham, D.J., Teglund, S., Sexl, V., McKay, C., Wang, D., Hoffmeyer, A., van Deursen, J., Sangster, M.Y., Bunting, K.D., et al. (1999). Stat5 is required for IL-2-induced cell cycle progression of peripheral T cells. *Immunity* 10, 249–259.
- Mudter, J., Yu, J., Zufferey, C., Brüstle, A., Wirtz, S., Weigmann, B., Hoffman, A., Schenk, M., Galle, P.R., Lehr, H.A., et al. (2011). IRF4 regulates IL-17A promoter activity and controls ROR $\gamma$ t-dependent Th17 colitis in vivo. *Inflamm. Bowel Dis.* 17, 1343–1358.
- Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* 13, 919–922.
- Murphy, T.L., Tussiwand, R., and Murphy, K.M. (2013). Specificity through cooperation: BATF-IRF interactions control immune-regulatory networks. *Nat. Rev. Immunol.* 13, 499–509.
- Oh, S.A., and Li, M.O. (2013). TGF- $\beta$ : guardian of T cell function. *J. Immunol.* 191, 3973–3979.
- Palmer, D.C., and Restifo, N.P. (2009). Suppressors of cytokine signaling (SOCS) in T cell differentiation, maturation, and function. *Trends Immunol.* 30, 592–602.
- Pearce, E.L., Mullen, A.C., Martins, G.A., Krawczyk, C.M., Hutchins, A.S., Zediak, V.P., Banica, M., DiCioccio, C.B., Gross, D.A., Mao, C.-A., et al. (2003). Control of effector CD8<sup>+</sup> T cell function by the transcription factor Eomesodermin. *Science* 302, 1041–1043.
- Pham, D., Moseley, C.E., Gao, M., Savic, D., Winstead, C.J., Sun, M., Kee, B.L., Myers, R.M., Weaver, C.T., and Hatton, R.D. (2019). Batf Pioneer the Reorganization of Chromatin in Developing Effector T Cells via Ets1-Dependent Recruitment of Ctcf. *Cell Rep.* 29, 1203–1220.e7.
- Phan, A.T., and Goldrath, A.W. (2015). Hypoxia-inducible factors regulate T cell metabolism

and function. *Mol. Immunol.* 68, 527–535.

Phanstiel, D.H., Boyle, A.P., Heidari, N., and Snyder, M.P. (2015). Mango: a bias-correcting ChIA-PET analysis pipeline. *Bioinformatics* 31, 3092–3098.

Pircher, H., Michalopoulos, E.E., Iwamoto, A., Ohashi, P.S., Baenziger, J., Hengartner, H., Zinkernagel, R.M., and Mak, T.W. (1987). Molecular analysis of the antigen receptor of virus-specific cytotoxic T cells and identification of a new V alpha family. *Eur. J. Immunol.* 17, 1843–1846.

Ross, S.H., and Cantrell, D.A. (2018). Signaling and Function of Interleukin-2 in T Lymphocytes. *Annu. Rev. Immunol.* 36, 411–433.

Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19, 789–800.

Sadreev, I.I., Chen, M.Z.Q., Umezawa, Y., Biktashev, V.N., Kemper, C., Salakhieva, D.V., Welsh, G.I., and Kotov, N.V. (2018). The competitive nature of signal transducer and activator of transcription complex formation drives phenotype switching of T cells. *Immunology* 153, 488–501.

Sahoo, A., Alekseev, A., Tanaka, K., Obertas, L., Lerman, B., Haymaker, C., Clise-Dwyer, K., McMurray, J.S., and Nurieva, R. (2015). Batf is important for IL-4 expression in T follicular helper cells. *Nat. Commun.* 6, 7997.

Salvador, J.M., Mittelstadt, P.R., Belova, G.I., Fornace, A.J., and Ashwell, J.D. (2005). The autoimmune suppressor Gadd45 $\alpha$  inhibits the T cell alternative p38 activation pathway. *Nat. Immunol.* 6, 396–402.

Schraml, B.U., Hildner, K., Ise, W., Lee, W.-L., Smith, W.A.-E., Solomon, B., Sahota, G., Sim, J., Mukasa, R., Cemerski, S., et al. (2009). The AP-1 transcription factor Batf controls T(H)17 differentiation. *Nature* 460, 405–409.

Sciammas, R., Shaffer, A.L., Schatz, J.H., Zhao, H., Staudt, L.M., and Singh, H. (2006). Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* 25, 225–236.

Seif, F., Khoshmirsafa, M., Aazami, H., Mohsenzadegan, M., Sedighi, G., and Bahar, M. (2017). The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun. Signal.* 15, 23.

Sen, D.R., Kaminski, J., Barnitz, R.A., Kurachi, M., Gerdemann, U., Yates, K.B., Tsao, H.-W., Godec, J., LaFleur, M.W., Brown, F.D., et al. (2016). The epigenetic landscape of T cell exhaustion. *Science* 354, 1165–1169.

Sergushichev, A.A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 16, 259.



Shan, Q., Zeng, Z., Xing, S., Li, F., Hartwig, S.M., Gullicksrud, J.A., Kurup, S.P., Van Braeckel-Budimir, N., Su, Y., Martin, M.D., et al. (2017). The transcription factor Runx3 guards cytotoxic CD8<sup>+</sup> effector T cells against deviation towards follicular helper T cell lineage. *Nat. Immunol.* 18, 931–939.

Simon, A.K., Desrois, M., and Schmitt-Verhulst, A.M. (1997). Interferon-regulatory factors during development of CD4 and CD8 thymocytes. *Immunology* 91, 340–345.

Smit, A., Hubley, R., Green, P., and Smit, H.A. (1996). RepeatMasker Open-3.0.

Stelekati, E., Chen, Z., Manne, S., Kurachi, M., Ali, M.-A., Lewy, K., Cai, Z., Nzingha, K., McLane, L.M., Hope, J.L., et al. (2018). Long-Term Persistence of Exhausted CD8 T Cells in Chronic Infection Is Regulated by MicroRNA-155. *Cell Rep.* 23, 2142–2156.

Storey JD, Bass AJ, Dabney A, Robinson D (2019). qvalue: Q-value estimation for false discovery rate control.

Strasser, A., Jost, P.J., and Nagata, S. (2009). The many roles of FAS receptor signaling in the immune system. *Immunity* 30, 180–192.

Sullivan, B.M., Juedes, A., Szabo, S.J., von Herrath, M., and Glimcher, L.H. (2003). Antigen-driven effector CD8 T cell function regulated by T-bet. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15818–15823.

Tinoco, R., Alcalde, V., Yang, Y., Sauer, K., and Zuniga, E.I. (2009). Cell-intrinsic transforming growth factor-beta signaling mediates virus-specific CD8<sup>+</sup> T cell deletion and viral persistence in vivo. *Immunity* 31, 145–157.

Tosello Boari, J., Araujo Furlan, C.L., Fiocca Vernengo, F., Rodriguez, C., Ramello, M.C., Amezcua Vesely, M.C., Gorosito Serrán, M., Nuñez, N.G., Richer, W., Piaggio, E., et al. (2018). IL-17RA-Signaling Modulates CD8<sup>+</sup> T Cell Survival and Exhaustion During Trypanosoma cruzi Infection. *Front. Immunol.* 9, 2347.

Tripathi, P., Kurtulus, S., Wojciechowski, S., Sholl, A., Hoebe, K., Morris, S.C., Finkelman, F.D., Grimes, H.L., and Hildeman, D.A. (2010). STAT5 is critical to maintain effector CD8<sup>+</sup> T cell responses. *J. Immunol.* 185, 2116–2124.

Vasanthakumar, A., Moro, K., Xin, A., Liao, Y., Gloury, R., Kawamoto, S., Fagarasan, S., Mielke, L.A., Afshar-Sterle, S., Masters, S.L., et al. (2015). The transcriptional regulators IRF4, BATF and IL-33 orchestrate development and maintenance of adipose tissue-resident regulatory T cells. *Nat. Immunol.* 16, 276–285.

van der Veeken, J., Zhong, Y., Sharma, R., Mazutis, L., Dao, P., Pe'er, D., Leslie, C.S., and Rudensky, A.Y. (2019). Natural Genetic Variation Reveals Key Features of Epigenetic and Transcriptional Memory in Virus-Specific CD8 T Cells. *Immunity* 50, 1202–1217.e7.

Walker, S.R., Nelson, E.A., Yeh, J.E., Pinello, L., Yuan, G.-C., and Frank, D.A. (2013). STAT5 outcompetes STAT3 to regulate the expression of the oncogenic transcriptional modulator BCL6. *Mol. Cell. Biol.* 33, 2879–2890.

Wang, D., Diao, H., Getzler, A.J., Rogal, W., Frederick, M.A., Milner, J., Yu, B., Crotty, S.,

Goldrath, A.W., and Pipkin, M.E. (2018). The Transcription Factor Runx3 Establishes Chromatin Accessibility of cis-Regulatory Landscapes that Drive Memory Cytotoxic T Lymphocyte Formation. *Immunity* 48, 659–674.e6.

Wingelhofer, B., Neubauer, H.A., Valent, P., Han, X., Constantinescu, S.N., Gunning, P.T., Müller, M., and Moriggl, R. (2018). Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia* 32, 1713–1726.

Wong, P., and Pamer, E.G. (2003). CD8 T cell responses to infectious pathogens. *Annu. Rev. Immunol.* 21, 29–70.

Xin, G., Schauder, D.M., Lainez, B., Weinstein, J.S., Dai, Z., Chen, Y., Esplugues, E., Wen, R., Wang, D., Parish, I.A., et al. (2015). A Critical Role of IL-21-Induced BATF in Sustaining CD8-T-Cell-Mediated Chronic Viral Control. *Cell Rep.* 13, 1118–1124.

Yamada, T., Park, C.S., Mamonkin, M., and Lacorazza, H.D. (2009). Transcription factor ELF4 controls the proliferation and homing of CD8+ T cells via the Krüppel-like factors KLF4 and KLF2. *Nat. Immunol.* 10, 618–626.

Yao, S., Buzo, B.F., Pham, D., Jiang, L., Taparowsky, E.J., Kaplan, M.H., and Sun, J. (2013). Interferon regulatory factor 4 sustains CD8(+) T cell expansion and effector differentiation. *Immunity* 39, 833–845.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515, 355–364.

Zaldumbide, A., Carlotti, F., Pognonec, P., and Boulukos, K.E. (2002). The role of the Ets2 transcription factor in the proliferation, maturation, and survival of mouse thymocytes. *J. Immunol.* 169, 4873–4881.

Zamisch, M., Tian, L., Grenningloh, R., Xiong, Y., Wildt, K.F., Ehlers, M., Ho, I.-C., and Bosselut, R. (2009). The transcription factor Ets1 is important for CD4 repression and Runx3 up-regulation during CD8 T cell differentiation in the thymus. *J. Exp. Med.* 206, 2685–2699.

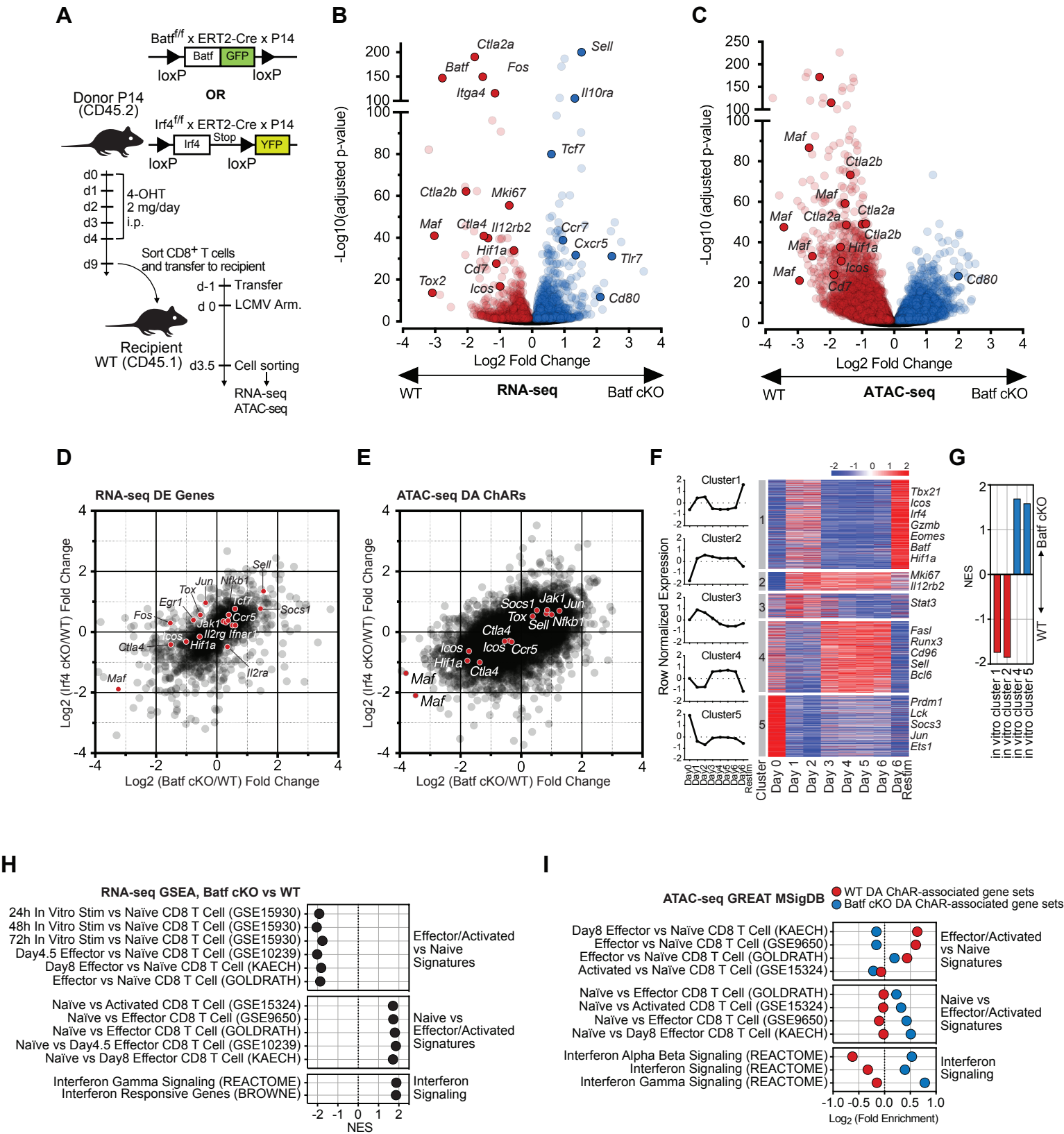
Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137.

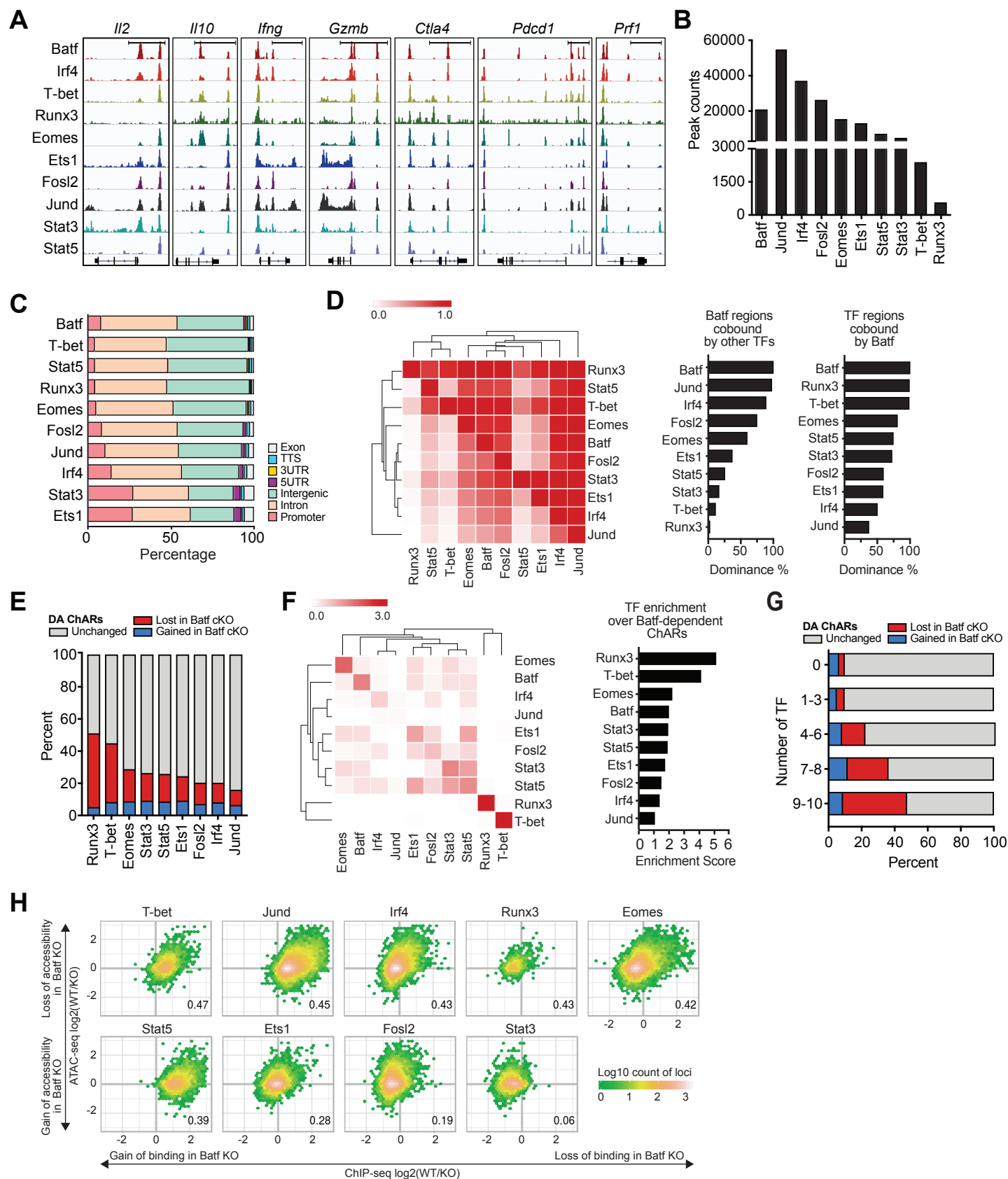
Zhou, S., Cerny, A.M., Fitzgerald, K.A., Kurt-Jones, E.A., and Finberg, R.W. (2012). Role of interferon regulatory factor 7 in T cell responses during acute lymphocytic choriomeningitis virus infection. *J. Virol.* 86, 11254–11265.



**Figure 1. Batf and Irf4 remodel chromatin accessibility at key regulatory regions during effector CD8+ T differentiation**

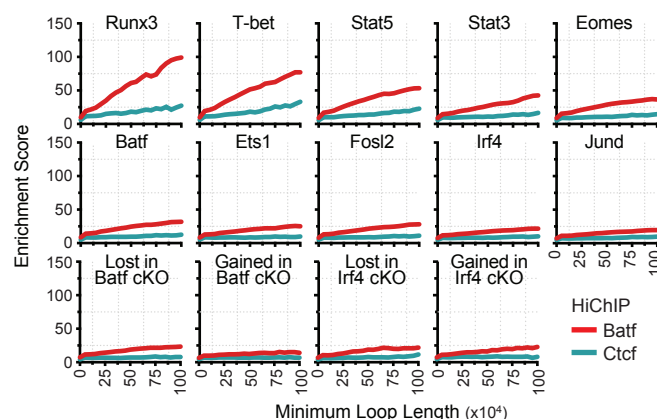


# Figure 2: Batf is required for the binding of key transcription factors in CD8 T cells and contributes to CD8 T cell phenotypes by cooperating with Runx3 and T-bet

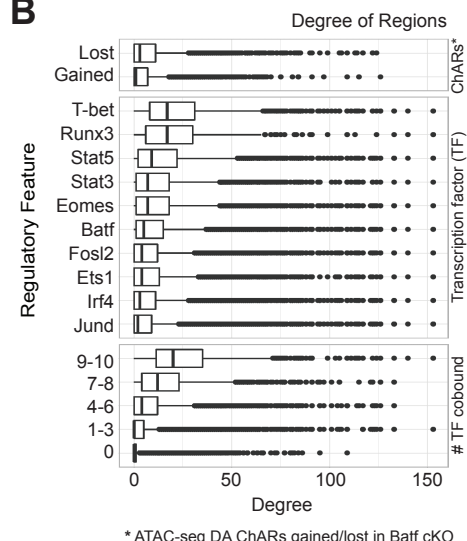


# Figure 3. Differential accessible ChARs and loci bound by key transcription factors are linked by Batf loops in CD8 T cells.

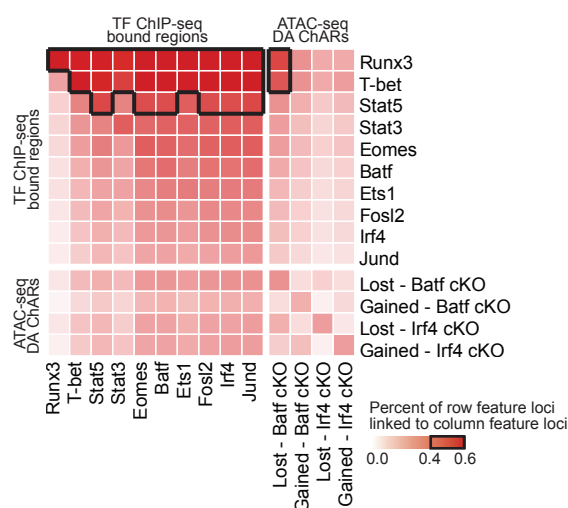
**A**



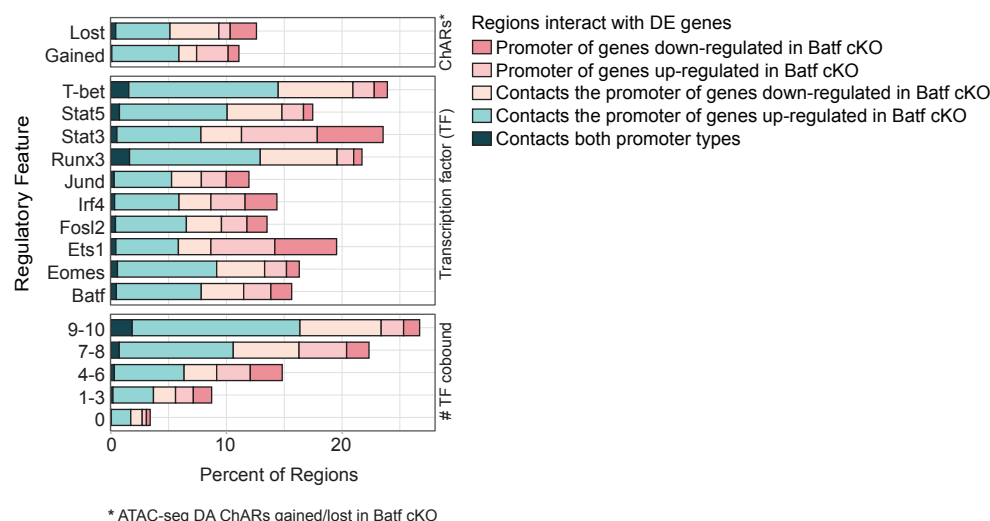
**B**



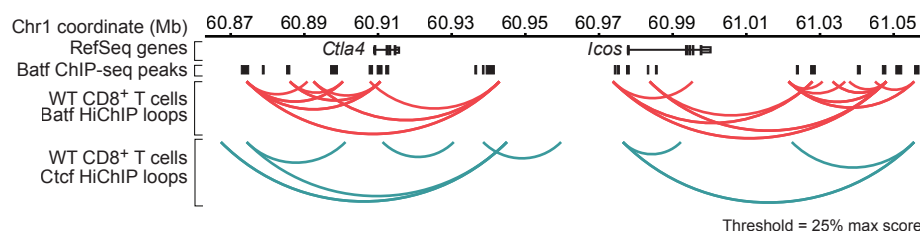
**C**

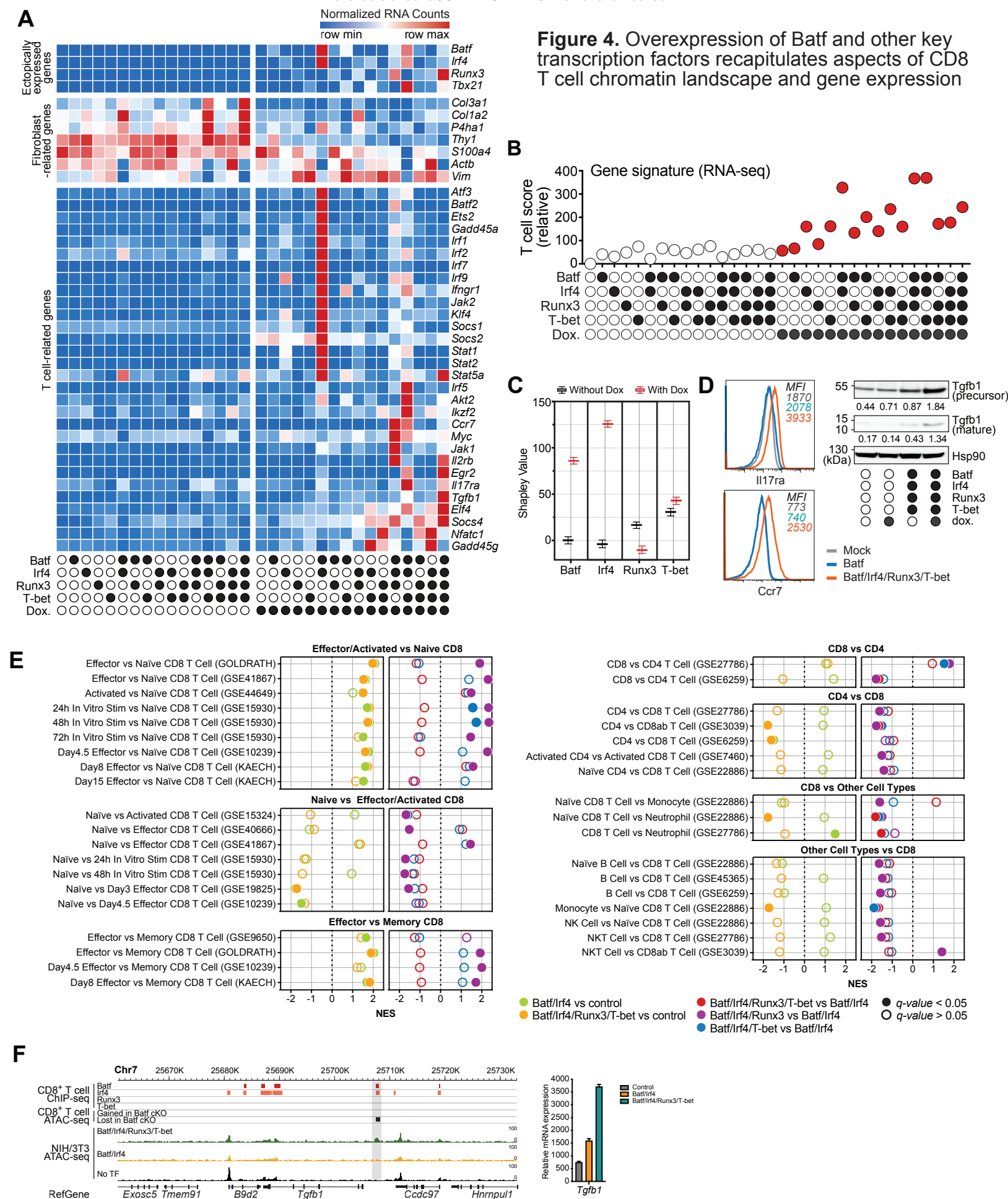


**D**



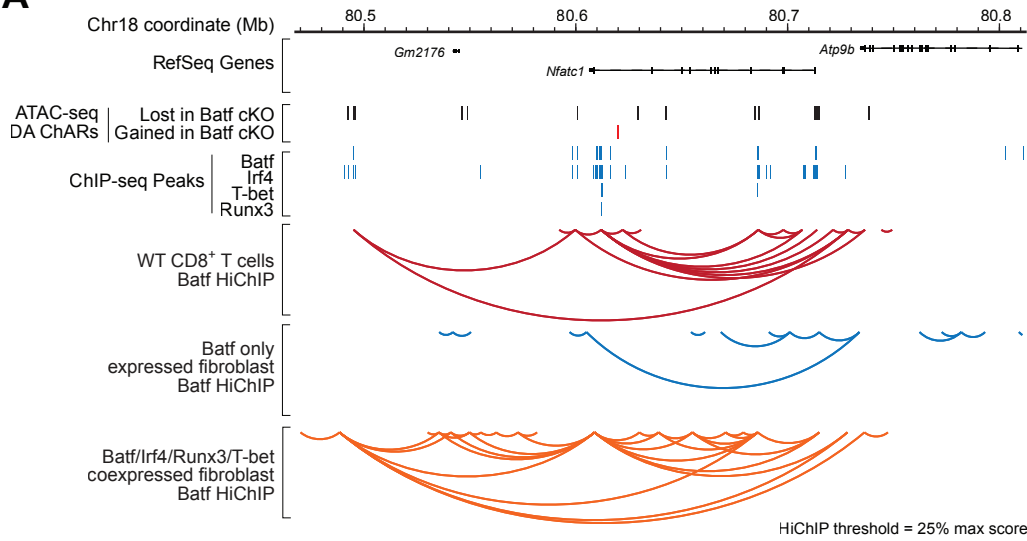
**E**



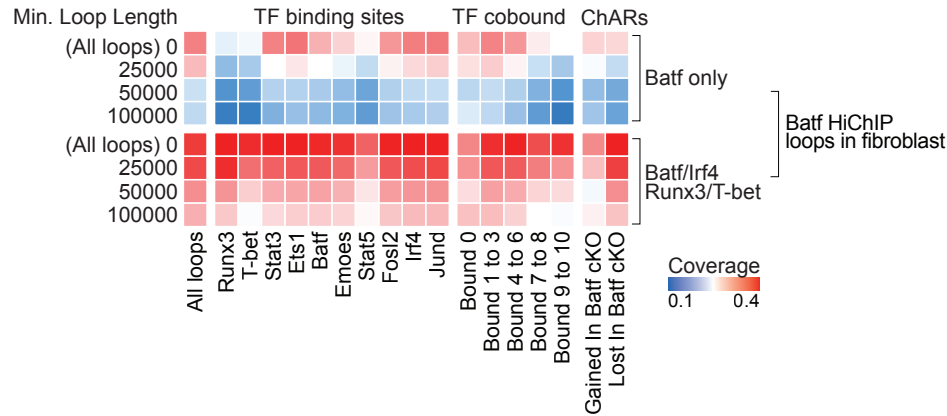


# Figure 5: Overexpression of Batf, Irf4, Runx3 and T-bet are able to reconstruct chromatin architectures of NIH/3T3 fibroblasts

**A**



**B**



**C**

