A Genome-wide Regulatory Network Identifies Key Transcription Factors for Memory CD8$^+$ T Cell Development

Guangan Hu and Jianzhu Chen

David H. Koch Institute for Integrative Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Corresponding to Jianzhu Chen (jchen@mit.edu)

Author's contacts:

Guangan Hu (gahu@mit.edu)

Jianzhu Chen (jchen@mit.edu)

Abstract

Memory $CD8^+$ T cell development is defined by the expression of a specific set of memory signature genes (MSGs). Despite recent progress, many components of the transcriptional control of memory $CD8^+$ T cell development are still unknown. To identify transcription factors (TFs) and their interactions in memory $CD8^+$ T cell development, we construct a genome-wide regulatory network and apply it to identify key TFs that regulate MSGs. Most of the known TFs in memory $CD8^+$ T cell development are rediscovered and about a dozen new TFs are also identified. *Sox4*, *Bhlhe40*, *Bach2* and *Runx2* are experimentally verified and *Bach2* is further shown to promote both development and recall proliferation of memory $CD8^+$ T cells through *Prdm1* and *Id3*. Gene perturbation study identifies the mode of interactions among the TFs with *Sox4* as a hub. The identified TFs and insights into their interactions should facilitate further dissection of molecular mechanisms underlying memory $CD8^+$ T cell development.

Immunological memory refers to faster and stronger responses to re-encountering of the same antigen. The basis for this enhanced response is the persistence of more abundant and intrinsically more reactive antigen-specific memory T and B lymphocytes that are generated following the initial antigen stimulation. Memory $CD8^+$ T cells are usually generated following antigen-stimulated T cell activation and expansion. In a typical $CD8^+$ T cell response, naïve $CD8^+$ T cells are activated to undergo clonal expansion when stimulated by appropriate antigen [1]. The resulting T cells acquire effector functions and migratory properties that allow them to clear antigens in both lymphoid and non-lymphoid organs. As antigen is cleared, most of the effector T cells die by apoptosis and only a small fraction survive and differentiate into memory $CD8^+$ T cells. Memory $CD8^+$ T cells are often divided into two subsets. Effector memory T cells (T_{EM}) are $CD62L^{lo}CCR7^{lo}$ and capable of rapid expression of effector functions following antigen stimulation to confer faster memory response. Central memory T cells (T_{CM}) are $CD62L^{hi}CCR7^{hi}$ and proliferate extensively upon antigen restimulation to confer stronger memory response.

Memory $CD8^+$ T cells are developmentally programmed as they express a specific set of memory signature genes (MSGs) [2, 3], which confer them with characteristic memory phenotype and function. Like many developmental processes, memory $CD8^+$ T cell development is ultimately controlled by transcription factors (TFs) that integrate external and internal signals to regulate the expression of the MSGs. In recent years, several studies have shed light on TFs that regulate the development of memory $CD8^+$ T cells. T-bet (encoded by *Tbx21*) and Eomesodermin (encoded by *Eomes*), both member of the T-box family, are essential for the differentiation of effector and memory $CD8^+$ T cells [4, 5, 6]. *Tcf7* is a TF downstream of the Wnt signaling. Consistent with the observation that activation of Wnt/β-catenin signaling promotes memory $CD8^+$ T cell development by suppressing terminal differentiation of effector T cells [7, 8], Tcf7-deficiency in $CD8^+$ T cells impairs T_{CM} differentiation [9]. *Klf2* has been shown to be associated with memory $CD8^+$ T cell development [10] probably by directly controlling the expression of cell surface receptors S1P1 and CD62L [11, 12]. *Id2* and *Id3*, the E-box-containing transcription suppressors, appear to regulate the development of memory $CD8^+$ T cell subsets [13]. Id2-knockout mice are deficient in memory $CD8^+$ T cells, but effector T cells generated in these mice are $CD127^{hi}CD62L^{hi}$, a phenotype similar to T_{CM}[14]. Deficiency in Id3 inhibits, whereas

overexpression of *Id3* promotes memory $CD8^+$ T cell development [15]. The B-cell transcriptional repressor Blimp-1 (encoded by *Prdm1*) promotes the terminal differentiation of effector $CD8^+$ T cells and is required for recall response of memory T cells [16, 17]. Despite these progresses, the current understanding of transcriptional regulation of memory $CD8^+$ T cell development is still limited, as additional TFs as well as their coordination are likely required to respond to external and internal signals in order to establish the MSG program for memory $CD8^+$ T cell development.

In this study, we assemble a genome-wide regulatory network associated with the development of $CD8^+$ T cells using publicly available gene expression data and a reverse-engineering algorithm. This regulatory network is applied to identify key TFs that regulate memory $CD8^+$ T cell development using the master regulator analysis (MRA) of the MSGs of $CD8^+$ T cells. The inferred TFs include most of the known TFs as well as a dozen new TFs with limited functional information in $CD8^+$ T cell differentiation. A regulatory module controlling the MSGs is constructed and the high accuracy of the regulations in the module is verified using ChIP-PCR. Gene perturbations identify multiple regulatory motifs among the key TFs, suggesting their complex regulations during the memory $CD8^+$ T cell development. Four of the newly identified key TFs (*Sox4*, *Bhlhe40*, *Bach2*, and *Runx2*) are experimentally validated to regulate memory $CD8^+$ T cell development and function. Bach2 is shown to promote memory $CD8^+$ T cell development and recall proliferation through Id3 and Prdm1. Our study represents the most comprehensive analysis of TFs and their interactions in memory $CD8^+$ T cell development to date. The identified TFs and the insights into their mode of interactions provide a foundation for further dissecting the molecular mechanisms underlying memory $CD8^+$ T cell development.

Results

Identification of TFs associated with memory $CD8^+$ T cells

We collected 386 gene expression profiles of naïve, effector and memory $CD8^+$ T cells of the mouse from 35 independent GEO datasets (Supplementary Table S1). 1,445 genes coding putative TFs [18] were manually mapped to the latest mouse genome to eliminate redundant and erroneous annotations, resulting in a total of 1,038 putative TFs. Among these putative TFs, 464 were expressed during the naïve to effector to memory $CD8^+$ T cell development (see Methods

for detail). Using a reverse-engineering algorithm CLR (context likelihood of relatedness)[19], the 386 gene expression profiles and the 1,038 putative TFs, a genome-wide regulatory network was assembled. The network consisted of 107,157 interactions among 11,032 genes. 62,272 interactions (58%) were between the 276 of the 464 expressed putative TFs and 8,572 target genes, suggesting that interactions are enriched among the expressed genes (P<0.001, binomial test). Furthermore, 3,219 of these interactions involve 154 out of 196 (79%) identified MSGs [3] (Supplementary Fig. S1).

To identify key TFs that regulate MSGs, we applied the MRA to the CLR-inferred interactions (Fig. 1a, see Methods). The MRA algorithm computes the statistical significance of overlaps of all interactions of each TF (inferred by CLR) with MSGs or a control gene set by a binomial test. From the 1,038 putative TFs, MRA identified 60 MSG-specific TFs at P<0.05 (binomial test), all of which are expressed in $CD8^+$ T cells (Supplementary Table S2). These 60 candidates were filtered by removing those whose knockout do not have any immune system phenotype as defined in MGI (Mouse Genome Informatics) [20]. The positive candidates were then analyzed for enrichment of DNA-binding motifs among MSGs or differential expression among naïve, effector and memory $CD8^+$ T cells (see Methods). This led to 21 key TFs that were ranked according to the numbers of MSG they regulate (Table 1). Text-mining of public references on these 21 TFs revealed that 8 of 12 known TFs, which have been reported to be involved in memory $CD8^+$ T cell development and function [21, 22, 23, 24], were identified by our analysis. These results show that our systematic approach is valid for identifying TFs that regulate memory $CD8^+$ T cell development.

Validation of a regulatory module for memory signature genes

To further explore the relative importance of the 21 identified TFs in regulating MSGs, we constructed a regulatory module using the top 10-ranked TFs (Fig. 1b). The resulting module contained 56% (86 out of 154) of MSGs that were present in the entire network (Supplementary Fig. S1). To verify this regulatory module, chromatin immunoprecipitation (ChIP) was performed for the top 3 TFs, *Sox4*, *Tcf7* and *Eomes*, in $CD8^+$ T cells followed by PCR amplification of promoter regions (within 1 kb upstream of the transcription-starting site) of randomly selected MSGs that were predicted to be regulated by *Sox4* or *Tcf7* or *Eomes*. As shown in Fig. 1c and 1d, promoter regions of 10 out of the 12 randomly selected *Sox4*-regulated

MSGs were amplified. Similarly, 12 out of 14 randomly selected *Tcf7*-regulated MSGs and 6 out of 9 randomly selected Eomes-regulated MSGs were amplified. On average, 80% of the tested promoter regions were immunoprecipitated with antibodies specific for each of the three TFs (Table 2), confirming the high accuracy of the constructed regulatory module. Furthermore, when the cumulative coverage of MSGs was plotted as a function of each of the 21 TFs, the top 2 TFs, *Sox4* and *Tcf7,* were shown to regulate 42% of MSGs (Supplementary Fig. S2).

Perturbation network of key TFs

Although the constructed regulatory module predicts interactions between TFs (Fig. 1b), the directions of regulation are not known. To find out these, the top 10 TFs and another two known memory-regulating TFs (*Id3* and *Tbx21*, #12 and #19 in the list, Table 1) [6, 15, 25] were perturbed in $CD8^+$ T cells *in vitro* by overexpression through retroviral transduction. The transcript level of each of the 12-selected TFs was measured by quantitative real-time PCR (Table 3). If changes in transcript level of ≥2 fold were taken as directional regulations, the perturbation results identified 41 regulations among the 12x12 matrix (31%). Notably, the top 3 TFs (*Sox4*, *Tcf7* and *Eomes*) directed 19 of the 41 regulations. To verify these regulations, ChIP-PCR was performed using antibodies specific for Sox4, Tcf7 and Eomes. As shown in Fig. 2a, 18 of the 19 regulations were confirmed. ChIP-PCR also identified 4 more regulations that were not observed in the perturbation study. Thus, compared to ChIP-PCR, perturbation studies is able to identify the directional regulations with 82% sensitivity and 91% specificity.

We then constructed a perturbation network of the 12 TFs with directional regulations (Fig. 2b). The top 3 TFs (*Sox4*, *Tcf7* and *Eomes*) and *Bach2* had more downstream targets than the number of TFs that regulate them (Supplementary Fig. S3), suggesting that they are at the upstream of a regulatory structure. TFs in the perturbation network formed multiple motifs, such as feedback and feed-forward loops (Supplementary Fig. S4). For example, in a feedback motif of *Sox4*-*Tcf7*-*Eomes*-*Tbx21* (Fig. 2c), *Sox4* and *Tcf7* regulate each other and they also regulate expression of *Eomes* and/or *Tbx21*. The latter regulations were further confirmed at the protein level as indicated by suppression of Eomes and Tbx21 by overexpression of *Sox4* or *Tcf7* (Supplementary Fig. S5). These results suggest that complex regulations involving multiple regulatory motifs among these TFs are involved in memory $CD8^+$ T cell development.

Validation of *Sox4* and *Bach2* in memory CD8$^+$ T cells

Among the top 10 TFs (Table 1), 6 are known to play important roles in memory CD8$^+$ T cell development and/or function. We then investigated whether the other 4 TFs (*Sox4*, *Bhlhe40*, *Bach2* and *Runx2*) are also involved in memory CD8$^+$ T cell development/function by examining the effect of overexpression and knockdown of these TFs on the recall proliferation of memory CD8$^+$ T cells *in vitro* and *in vivo*. CD8$^+$ T cells expressing the 2C TCR were activated with cognate peptide SIYRYYGL (SIY) and then transduced with retroviruses expressing GFP plus *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* or expressing GFP plus shRNA specific for one of the four TFs (Supplementary Table S3 and S4). The 2C T cells were then cultured in the presence of cytokine IL-7 to induce the development of memory CD8$^+$ T cells (Supplementary Fig. S6). To assay recall proliferation, the *in vitro* memory 2C T cells were restimulated with SIY and the number of transduced (GFP$^+$) and non-transduced (GFP$^-$) 2C T cells were quantified on day 4 and 6. Compared to the vector control, overexpression of *Sox4* or *Bach2* led to a significant increase in the proportions of GFP$^+$ cells (Fig. 3a), suggesting a higher recall proliferation. When the *in vitro* generated memory 2C T cells were adoptively transferred into C57BL/6 (B6) mice followed by activation through infection with influenza virus that express SIY (WSN-SIY virus) [26], a significant increase in the proportion of GFP$^+$ 2C T cells was also observed in the draining lymph nodes (DLN) (Fig. 3b), the blood, lung and spleen (Supplementary Fig. S7) 5 days post infection (dpi) if the transduced memory T cells expressed *Sox4* or *Bach2*. Conversely, knockdown of *Sox4* or *Bach2* (Supplementary Fig. S8) resulted in a significant inhibition of the recall proliferation of memory 2C T cells both *in vitro* and *in vivo* (Fig. 3c,d). Although overexpression of *Bhlhe40* and *Runx2* inhibited the *in vivo* recall proliferation of the transduced memory 2C T cells (Fig. 3b), no significant change was observed in *in vitro* recall response (Fig. 3a) and in knockdown assay (Fig. 3c,d). As positive controls, we tested in parallel known TFs: overexpression of *Eomes* promoted the recall proliferation whereas overexpression of *Klf2* inhibited the recall proliferation (Supplementary Fig. S9), consistent with previous reports [9, 11]. These results show that *Sox4* and *Bach2* likely promote the recall proliferation of memory CD8$^+$ T cells.

Enhanced memory T cell development by *Bach2* overexpression

To confirm the effect of overexpression of *Bach2* on recall proliferation of memory $CD8^+$ T cells, we activated 2C T cells *in vitro* for two days, transduced the activated T cells with retroviruses expressing GFP alone or GFP plus Bach2. The cells were cultured in the presence of IL-2 for two more days and then adoptively transferred into antigen-free B6 mice to induce *in vivo* memory 2C T cells (Fig. 4a). Twenty-three days after transfer, the frequency, phenotype and function of persisting 2C T cells were analyzed. Both transduced (GFP^+) and non-transduced (GFP^-) 2C cells persisted in the recipient mice (Fig. 4b), GFP^+ 2C cells exhibited a typical memory phenotype as indicated by expression of CD62L and IL-7 receptor (IL-7R), similar to the GFP^- 2C cells in the same recipient (Fig. 4c). The persisting memory 2C cells, both transduced and non-transduced, were rapidly induced to express IFNγ and TNFα following antigen stimulation (Fig. 4d). Furthermore, some recipient mice were infected with WSN-SIY virus and the recall proliferation of persisting 2C cells in the spleen and DLN were analyzed 7 days later. As shown in Fig. 4e, if the 2C cells were originally transduced with GFP-expressing retrovirus, the proportion of GFP^+ versus GFP^- 2C cells did not changed following WSN-SIY challenge. However, if the 2C cells were originally transduced with GFP and *Bach2*-expressing retrovirus, the proportion of GFP^+ cells was significantly higher in both DLN and spleen, suggesting a stronger recall proliferation by *Bach2*-expressing memory 2C cells.

The observed stronger recall response by *Bach2*-expressing memory 2C cells could be due to the generation of more memory T cells and/or that the *Bach2*-expressing memory T cells are more responsive to restimulation. To investigate these possibilities, we activated 2C T cells *in vitro* for two days and transduced them with retroviruses expressing GFP alone (vector) or GFP plus *Bach2* (Fig. 5a). Twenty-four hours later, the T cells were adoptively transferred into B6 mice followed by WSN-SIY virus infection. 2C T cell responses were analyzed by flow cytometry 7 dpi. Compared to the non-transduced 2C T cells, vector-transduced 2C T cells had the same expression profiles for CD62L, IL-7R, Klrg1 and CD27 in the same organs of the same mice (Fig. 5b, upper panel). In contrast, a significantly higher proportion of *Bach2*-transduced 2C T cells expressed CD62L and CD27, but a significantly lower fraction expressed Klrg1, in the DLN and spleen compared to non-transduced 2C cells in the same organs of the same recipients. When the persisting 2C T cells were analyzed 30 dpi, no significant differences in IFNγ and IL-2 expression were observed among non-transduced, vector-transduced and *Bach2*-transduced 2C T

cells in response to restimulation *in vitro* (Fig. 5c and Supplementary Fig. S10). Although no difference in CD62L and IL-7R expression was detected, CD27 was higher in *Bach2*-transduced than vector-transduced 2C T cells in the lung and DLN. (Fig. 5c and Supplementary Fig. S10). Importantly, significantly more *Bach2*-transduced 2C T cells persisted in the DLN, spleen and bone marrow (Fig. 5d,e). These result suggest that overexpression of *Bach2* likely promotes the generation of memory $CD8^+$ T cells *in vivo*.

We further investigated the effect of *Bach2* overexpression in naïve 2C T cells on memory $CD8^+$ T cell development *in vivo*. Bone marrow progenitor cells isolated from 2C TCR transgenic mice were transduced with retroviruses expressing GFP alone or GFP plus *Bach2* and adoptively transferred into sublethally irradiated $Rag2^{-/-}$ mice to generate naïve 2C T cells that express GFP or GFP plus *Bach2* (Fig. 6a). The resulting transduced and non-transduced naïve 2C cells were then adoptively transferred into B6 mice followed by infection with WSN-SIY virus. The frequency, phenotype and function of 2C T cells were analyzed 7 dpi. Vector-transduced and non-transduced 2C T cells from the same recipient mice had the same expression profile of CD62L, IL-7R, Klrg1, IFNγ and TNFα (Fig. 6b). While *Bach2*-transduced 2C T cells expressed similar levels of IFNγ and TNFα as non-transduced and vector-transduced 2C T cells, more cells expressed CD62L but fewer cells expressed Klrg1, resembling to $CD62L^{hi}Klrg1^{low}$ memory precursors (Fig. 6b and Supplementary Fig. S11). By 30 dpi, no significant differences were observed among non-transduced, vector-transduced and *Bach2*-transduced 2C T cells in expression of CD62L, IL-7R, Klrg1, IFNγ and TNFα (Fig. 6c). However, more *Bach2*-transduced 2C T cells were found in the blood, lung, DLN and spleen 30 dpi (Fig. 6d,e). Together, these data suggest that overexpression of *Bach2* promotes the development of memory $CD8^+$ T cells.

Diminished memory T cell development due to *Bach2* deficiency

We also examined the effect of Bach2 knockdown on memory T cell development. The approach was the same as outlined as in Figure 5a, except the retrovirus expressed shRNA specific for Bach2. Briefly, activated 2C T cells were transduced with shRNA-expressing retrovirus and adoptively transferred into B6 mice followed by infection with WSN-SIY influenza virus. The number and frequency of GFP^+ and GFP^- 2C T cells in various organs were analyzed 7 and 30

dpi. No significant difference was observed among non-transduced, vector-transduced and shRNA-transduced 2C T cells in terms of CD62L, IL-7R, Klrg1, CD27 and IFNγ expression (Supplemental Fig. S12). However, the proportion of shRNA-transduced 2C T cells was reduced significantly in the blood, lung, DLN, spleen and bone marrow 30 but not 7 dpi (Fig. 5f). Consistently, the number of shRNA-transduced 2C T cells in the spleen was lower as compared to the numbers of non-transduced and vector-transduced 2C T cells 30 dpi (Supplemental Fig. S12c).

To investigate the effect of Bach2 knockout on memory T cell development, we constructed chimeric mice where T and B cells were deficient in Bach2 by adoptively transferring bone marrow cells from Bach2 knockout mice into sublethally irradiated $Rag2^{-/-}$ mice. Three months after reconstitution, mice were infected with WSN-SIY influenza virus and analyzed for the presence of SIY-specific memory $CD8^+$ T cells 30 days later. The percentage of SIY-specific $CD8^+$ T cells was lower in the DLN and spleen of chimeric mice that were reconstituted with $Bach2^{-/-}$ than $Bach2^{+/+}$ bone marrow cells (Fig. 5g). Consistently, the number of SIY-specific $CD8^+$ T cells in the spleen was lower in mice reconstituted with $Bach2^{-/-}$ than $Bach2^{+/+}$ bone marrow cells. Considering that *Bach2* expression was down-regulated in effector and then up-regulated in memory T cells during naïve to effector to memory cell transition (Table 1), together these results show that Bach2 promotes memory $CD8^+$ T cell development.

Enhanced proliferation of T cells by *Bach2* overexpression

To further explore the mechanism underlying the observed effect of *Bach2* on memory T cell development and response, we examined whether *Bach2* affects T cell proliferation. 2C T cells were activated *in vitro* and transduced with either vector or *Bach2*-expressing retroviruses. The cells were cultured in the presence of either IL-2 or IL-7 and the proportion of transduced (GFP^+) versus non-transduced (GFP^-) cells in the same cultures was quantified over time. In the IL-7 culture, the proportion of transduced versus non-transduced 2C cells remained stable regardless whether the 2C T cells were transduced with vector or *Bach2* (Fig. 7a). Similarly, the proportion of vector-transduced versus non-transduced 2C cells remained stable in the IL-2 cultures. However, the proportion of *Bach2*-transduced 2C T cells increased significantly over time in the IL-2 cultures (Fig. 7a). When the cells were lableled with eFluor and followed over time, *Bach2*-

transduced 2C T cells diluted the flourescent dye more extensively than non-transduced 2C T cells (Fig. 7b). These data suggest that *Bach2* promotes proliferation of activated $CD8^+$ T cells.

We noticed a feed-forward regulatory motif of *Sox4-Bach2-Prdm1-Id3* in the perturbation network (Fig. 7c and Supplementary Fig. S4). We verified this regulatory motif by showing that overexpression of *Bach2* suppressed expression of *Prdm1* but stimulated expression of *Id3* at transcriptional (Table 3) and translational levels (Fig. 7d). ChIP-PCR analysis with anti-Bach2 antibody also confirmed that Bach2 binds directly to the promoter regions of *Prdm1* and *Id3* (Fig. 7e, f). Further supporting the regulatory motif, overexpression of *Prdm1* significantly reduced the *Bach2*-mediated proliferation of activated T cells (Supplementary Fig. S13). As *Prdm1* is known to inhibit T cell proliferation whereas *Id3* stimulates survival of effector T cells [15, 16], we determined the effect of *Bach2* overexpression on molecules that regulate cell cycle and survival. Overexpression of *Bach2* stimulated CDK4, CDK6 and Bcl6 expression and Rb phosphorylation at amino acid residues 780 and 795 but inhibited expression of p27kip (Fig. 7g). These results suggest that *Bach2* regulates memory T cell development and recall proliferation by regulating cell cycle control possibly through *Prdm1* and *Id3*.

Discussion

At the molecular level, development of memory $CD8^+$ T cells is the establishment of MSG expression program, which ultimately is controlled by TFs. Although several TFs have been described to regulate memory $CD8^+$ T cell development, for a comprehensive understanding of transcriptional regulation of memory $CD8^+$ T cell development, it is necessary to identify most, if not all, key TFs that regulate MSGs and construct a genome-wide transcriptional network that supports memory $CD8^+$ T cell development. Using systems biology approaches and publically available gene expression data, here we have assembled a genome-wide regulatory network associated with $CD8^+$ T cells of the mouse. Applying MRA to this network, we have identified twenty-one key TFs, down-narrowed from 1038 putative TFs, which regulate the expression of 70% of MSGs. Our approach is valid based on the following considerations. First, our method identified eight of the twelve TFs known to be involved in memory $CD8^+$ T cell development, including *Tcf7*, *Eomes*, *Prdm1*, *Klf2*, *Id2*, *Stat4*, *Id3* and *Tbx21* (Table 1). Although *Bcl6*, *Stat3* and *Myc* are known to regulate memory $CD8^+$ T cell development, they were not within the top

21 TFs identified using our methodology because CLR-inferred targets do not overlap with MSGs possibly due to limited data on MSGs or the network. *NF-κb* plays important roles in both effector and memory T cell development [27, 28]. We found that *NF-κb* was a key TF that regulates effector signature genes (ESG) when MRA was applied to the network and ESGs. Second, our method identified several TFs that are not known to function in the memory $CD8^+$ T cell development, including *Sox4*, *Bhlhe40*, *Bach2* and *Runx2* in the top 10 TFs (Table 1). Follow-up experimentation showed that these newly identified TFs indeed play important roles in memory $CD8^+$ T cell development and function. Overexpression of *Sox4* and *Bach2* promoted a significantly higher recall proliferation of memory $CD8^+$ T cells both *in vitro* and *in vivo*. Conversely knockdown of *Sox4* and *Bach2* inhibited the recall proliferation of the transduced memory T cells (Fig. 3). Overexpression of *Bhlhe40* and *Runx2* inhibited the *in vivo* recall proliferation, although no significant change was observed in *in vitro* recall response and in knockdown assay (Fig. 3). Further analysis showed that overexpression of *Bach2* also promotes memory $CD8^+$ T cell development (Fig. 4-6). Third, compared to the traditional method of differential gene expression analysis [2, 3], which generates a long list of candidates using fold-change-based approaches, our network methods identify and rank order TFs according to their statistical importance. Reduction of hundreds of TFs to two-dozen key TFs makes direct experimental validation more manageable. The network approach and methodologies developed here can be applied to any phenotypic transition, such as effector T cells and exhausted T cells, to identify novel transcriptional modules and TFs.

Studies have suggested that memory $CD8^+$ T cell development is coordinately regulated by several TFs (reviewed in Refs. [22, 24]), including *Eomes* and *Tbx21* [4]. Our network and perturbation studies have now greatly expanded the understanding of the mode of interactions to the top 21 TFs. Our analysis reveals a dense overlapping regulation (DOR) among the key TFs (Fig. 2b). This mode of regulation is essential for sensing multiple external signals and integrate them into distinct cell fate outcomes [29]. As two classes of TFs have been proposed to control the developmental potentials of effector and memory fates in a quantitative manner [22, 24, 30], the DOR might contribute to the quantitative regulations during effector to memory $CD8^+$ T cell development. Our analysis also shows complex regulations with both feedback and feed-forward motifs among the key TFs (Table 3 and Fig. 2b). In the regulatory motif of *Sox4-Tcf7-Eomes-*

Tbx21, Tcf7, Eomes and Tbx21 are known to be critical for memory $CD8^+$ T cell development. The association of *Sox4* with these three TFs and especially its "hub" position in this motif are intriguing. *Sox4* is not known to regulate memory $CD8^+$ T cell development. However, it stabilizes β-catenin to modulate Wnt-Tcf7 signaling [31, 32], which promotes memory $CD8^+$ T cell development. *Sox4* also regulates 'stemness' of cancer cells [33], a property shared by memory T cells. Furthermore, evidence suggests that *Sox4* might be a direct regulatory target of TGFβ signaling [34], which is essential for the differentiation of $CD8^+$ T cells [35]. These previous observations, together with our finding of the "hub" position of *Sox4* in the regulatory motif of *Sox4-Tcf7-Eomes-Tbx21*, suggest that Sox4 is a critical TFs regulating memory $CD8^+$ T cell development. While this hypothesis has yet to be validated, our finding that overexpression of Sox4 promotes recall proliferation of memory T cells suggests that Sox4 is involved in memory $CD8^+$ T cell development.

In the regulatory motif of *Sox4-Tcf7-Eomes-Tbx21*, we found that Tcf7 binds to Eomes promoter (Fig. 2a) and retroviral expression of Tcf7 leads to a downregulation of Eomes transcript (Table 3) and protein (Supplementary Fig. S5). The latter results contradict with the previous report showing that the level of Eomes transcript and protein are decreased in memory $CD8^+$ T cells from Tcf7-deficient mice [9]. Although we do not know the precise causes underlying the observed opposite effects, the following differences between the two studies may provide part of the explanation. First, our study was carried out in *in vitro*, using activated $CD8^+$ T cells that are in transition to memory T cells, whereas the previous study used memory $CD8^+$ T cells directly from mice. Second, in our study, we overexpressed Tcf7 for a short period (72 hrs) before assaying the effect on Eomes expression, whereas the previous study examined the accumulated effect of germline Tcf7 knockout on Eomes expression. The differences in the stage of T cells, *in vitro* vs. *in vivo*, overexpression vs. deficiency, and the length of Tcf7 overexpression or deficiency could all contribute to the observed differences in the two studies. Although the discrepancy raises concern of our approach, results from our perturbation study on the effect of *Bach2* on *Prdm1*, *Prdm1* on *Tcf7*, and *Id2* on *Id3* are all consistent with previous reports [17, 36, 37], suggesting the validity of our *in vitro* assay in most cases.

Our detailed analysis of the feed-forward motif of *Sox4-Bach2-Prdm1-Id3* (Fig. 7c) reveals new insight into memory $CD8^+$ T cell development. This regulatory motif includes two known TFs (*Id3* and *Prdm1*) and two unknown TFs (*Sox4* and *Bach2*) in memory $CD8^+$ T cell development. Through both overexpression and knockdown/knockout in $CD8^+$ T cells both *in vitro* and in mice, we provide extensive evidence showing that *Bach2* promotes memory $CD8^+$ T cell development (Fig. 4-6). One mechanism appears to be by stimulating the induction of memory T cell precursors as Bach2-expression in effector T cells leads to a phenotype of $CD62L^{hi}Klrg^{lo}CD27^+$ (Fig. 5b-c and Supplementary Fig. S10), which is considered as central memory T cell precursors with high proliferative potential [38, 39]. Recently, two studies report that Bach2 regulates $CD4^+$ T cell development and function by suppressing effector gene expression [40, 41]. Our observation that Klrg1 is suppressed by Bach2 suggests that suppression of effector function may also be important for the development of memory $CD8^+$ T cells. Another mechanism is by stimulating T cell proliferation. We showed that overexpression of *Bach2* enhances IL-2 driven T cell proliferation *in vitro* and recall proliferation *in vitro* and *in vivo*. In addition, when *in vitro* memory 2C T cells were labeled with eFluro dye and adoptively transferred into $Rag^{-/-}$ mice, *Bach2*-transduced T cells diluted the flourescent dye more extensively than the non-transduced and vector-transduced T cells (Supplemental Fig. S14). The enhanced proliferation could lead to development and/or survival of memory T cells.

Our study further sheds light on the mechanisms by which Bach2 promotes T cell proliferation. In the regulatory motif, *Bach2* promotes *Id3* expression but suppresses *Prdm1* expression through direct binding to their promoter regions (Fig. 7c-f), the latter is consistent with *Bach2* suppression of *Prdm1* expression in B cells [37]. *Id3* is known to promote cell cycle and recall proliferation of memory $CD8^+$ T cells by binding to and inhibiting E proteins [15, 42]. Thus, by promoting *Id3* expression, *Bach2* stimulates T cell proliferation. *Prdm1* is known to antagonize *Bcl6*, which promotes cell cycle by suppressing the expression of cell cycle inhibitor p27kip [43]. $Bcl6^{-/-}$ mice exhibit a profound deficiency of memory T cells [44, 45], whereas in the absence of p27kip, memory $CD8^+$ T cells exhibit enhanced homeostatic and recall proliferation [46]. Consistently, we show that overexpression of Bach2 promotes Bcl6 expression but inhibits p27kip expression (Fig. 7g). Thus, Bach2 also stimulates T cell proliferation by suppressing

Prdm1 expression. Together, these findings suggest that *Bach2* promotes memory $CD8^+$ T cell development and recall proliferation through Id3- and Prdm1-mediated cell cycle control.

Development of memory $CD8^+$ T cells requires integration of multiple external and internal signals to establish a new transcriptional program of MSGs that endows memory $CD8^+$ T cells with characteristic features in phenotype, tissue distribution, homeostasis and recall potentials. In this study, we have shown that integrated systems biology approaches can be effectively used to identify key TFs and their mode of interactions that underlies memory $CD8^+$ T cell differentiation and function. Further analysis of motifs in the regulatory network should help to elucidate in detail the molecular mechanisms underlying memory $CD8^+$ T cell development and function.

Methods

Regulatory network and master regulator analysis

386 public microarrays related to $CD8^+$ T cells from 35 independent GEO datasets (till September 2009) were downloaded from the NCBI database of Gene Expression Omnibus (GEO) (Supplementary Table S1). All raw image files were reprocessed to normalize the data using R program with a gcRMA method. Gene expression data was used to construct the regulatory network with the putative TFs using a reverse engineering algorithm CLR [19]. Among the 1,445 putative TFs identified according to the domain predictions of protein sequences [18], 1,038 were manually mapped to the latest mouse genome. To compare gene expression in different $CD8^+$ T cells, samples were grouped into naïve, effector and memory based the cell types from which the microarray analysis were done (Supplementary Table S1). Gene was considered as expressed in $CD8^+$ T cells if the average gene expression level in one of the three groups was more than 8 (gcRMA values).

To identify TFs associated with memory $CD8^+$ T cell development, we used Master Regulator Analysis (MRA) to compute the statistical significance of overlaps of all interactions of each TF (inferred by CLR) with MSGs or a control gene set by a binomial test. The MSGs were differentially expressed genes between memory $CD8^+$ T cells and naïve/effector $CD8^+$ T cells identified previously [3]. 332 background genes were identified from the 386 gene expression profiles based on high levels of gene expression (gcRMA value >10) but minimal variation

among 386 samples (variation from mean <0.5). This criterion minimizes the potential of the selected genes not being regulated by TFs in $CD8^+$ T cells. From the 1,038 putative TFs, MRA identified 60 MSG-specific TFs at $P<0.05$ (binomial test), all of which are expressed in $CD8^+$ T cells (Supplementary Table S2). These 60 candidates were filtered by removing those whose knockout does not have any immune system phenotype as defined in MGI [20]. The positive candidates were analyzed for enrichment of DNA-binding motifs in the promoter regions (-2000 to -1) of the MSGs using the program MatInspector [47] or differential expression among naïve, effector and memory $CD8^+$ T cells. This led to 21 key TFs that exhibit immune system phenotype with either an enrichment of DNA-binding motifs among MSGs or differential expression.

Mice and virus

The 2C TCR transgenic mice on $Rag2^{-/-}$ and C57BL/6 (B6 $Thy1.1^+$) background ($2C^+Rag^{-/-}$) were maintained in the animal facility at the Massachusetts Institute of Technology (MIT). These mice express the 2C TCR on $CD8^+$ T cells specific for SIYRYYGL peptide (SIY) in association with MHC class I K^b molecule [48]. B6 and $Rag2^{-/-}$ mice were from the Jackson Laboratory. Mice were used at 8-16 weeks of age. All animal studies and procedures were approved by the Massachusetts Institute of Technology's Committee for Animal Care. Recombinant WSN-SIY virus encoding the SIY epitope in the neuroaminidase stalk was constructed by plasmid-based reverse genetics and grown in Madin-Darby canine kidney cells [26]. For infection, mice were anesthetized and given 100 pfu (sublethal dose) intranasally.

Flow cytometry and cell sorting

Antibodies specific for CD8α, Thy1.1, Klrg1, CD62L, CD127 (IL-7R), CD27, IFNγ, TNFα, IL-2, Eomes and T-bet (Tbx21) were purchased from BioLegend or eBiosciences and used at the recommended concentration. Single cell suspensions were prepared from spleens and mediastinal (draining) lymph nodes (DLN), peripheral blood, and lung. Splenocytes and lymphocytes were collected in 8 ml HBSS by crushing the spleen and lymph node with frosted glass slides and filtering the cell suspension through 80 μm nylon filters, respectively. Lungs were harvested and ground through a cell strainer, followed by incubation with 2 ml of digestion buffer (RPMI 1640 medium containing 3mg/ml of collagenase A (Roche), 5% FBS and 10mM

HEPES) at 37°C for 1 hour. Red blood cells (RBCs) in the spleen, blood and lung were lysed with RBC lysis buffer (Gibco) and the cells were washed with complete RPMI. The cells were counted and 1-3 x10^6 cells were used for surface staining. Cells were washed twice with PBS plus 2% FBS before cytometry analsysis. For intracellular staining, splenocytes were stimulated with SIY peptide for 5 hours in the presence of GolgiPlug (BD Biosciences). Cells were washed twice with PBS with 2% FBS and stained with indicated antibodies. The cells were then fixed and stained with labeled antibodies using an intracellular staining kit (Cytofix/Cytoperm kit; BD Biosciences) according to the manufacture's instructions. Stained cells were analyzed on either a FACSCalibur or Accuri™ C6 flow cytometer (BD Biosciences). 0.5-2 x 10^6 events were collected and analyzed with FlowJo software. Cell sorting was carried out with a MoFlo cell sorter or FACSAria (BD Biosciences).

Retrovirus production and infection

Retroviral pMIGw-GW gateway vector was constructed by inserting a gateway cassette at EcoRI site of the pMIGw vector (Addgene #9044) using a gateway construction kit (Invitrogen). All ORFs encoding 12 TFs were amplified with primers (Supplementary Table S3) and cDNA from mouse splenocytes and cloned into pMIGw-GW using the gateway cloning technology. shRNAs for specific TFs (Supplementary Table S4) were chosen from the predicted TRC library and cloned into pMKO.1 GFP retroviral vector (Addgene #10676). Briefly, synthesized single-strand sense and antisense oligonucleotides were annealed into double-strand oligonucleotides for short hairpin RNA in the annealing buffer (10mM TrisCl (pH7.5), 50mM NaCl and 1mM EDTA). The double-strand oligonucleotides were directly treated with T4 polynucleotide kinase (NEB) and ligated into pMKO.1 GFP vector) between AgeI and EcoRI sites.

293FT cells were cultured to 60% confluency in 6-well plates. Cells were co-transfected with retroviral vector plasmid (4μg) and packing plasmids pCL-Eco (1μg) with 150μl DMEM and 15μl TransIT®-LT1 (Mirus) according to the manufacture's instructions. On the second day, the culture was replaced with fresh medium. On the third day, supernatant was collected and filtered through a 0.45 μm low-protein binding membrane (Pall Life Science). Fresh viral supernatants were used for spin infection of $CD8^+$ T cells in all experiments.

For infection, cells from spleens and lymph nodes were harvested from $2C^+RAG^{-/-}$ mice, pooled and cultured in 6-well plates in the presence of SIY peptide (1μg/ml) in the complete

RPMI medium (RPMI 1640 supplemented with 10% FBS, 5mM HEPES, 2 mM glutamine, 100U/ml penicillin, 100μg/ml streptomycin and 50μM β-mercaptoethanol (Invitrogen)). Two days later, activated 2C T cells were collected, washed and resuspended at 2 x 10^6 cells per ml in the complete RPMI medium. 1ml fresh retrovirus supernatants and 0.25ml 2C cells with a final concentration of 5μg/ml polybrene (American Bioanalytical) were added to one well of a 24-well plate and spun for 90min at 2500rpm at 32°C to infect T cells. 24 hours later, cells were collected for direct adoptive transfer, or resuspended and cultured in 3ml fresh RPMI medium with 100U/ml IL-2 (eBioscience). After culture for 24 hours, 2C T cells were analyzed for GFP expression by flow cytometry and prepared for injection into mice to generate *in vivo* memory T cells or further culture to generate *in vitro* memory T cells.

Generation and recall proliferation of memory T cells

To generate *in vitro* memory T cells, activated (and transduced) 2C T cells were cultured in complete RPMI medium supplemented with 5ng/ml IL-7 (Peprotech) for 7 days with change of fresh IL-7-supplemented medium every two days. Cells were analyzed for memory phenotype by flow cytometry on day 7. To test the recall proliferation, *in vitro* memory 2C T cells ($1x10^5$) were cultured with B6 splenocytes ($5x10^5$) in a 12 well plate in complete RPMI medium supplemented with 1 μg/ml SIY peptide and 100 units/ml IL-2. The numbers and phenotype of 2C T cells were analyzed by flow cytometry 4 and 6 days later. Alternatively, *in vitro* memory 2C T cells ($2x10^5$) were transferred to B6 recipients and challenged with WSN-SIY virus. 2C cells were analyzed by flow cytometry 5 dpi.

To generate *in vivo* memory T cells, activated and transduced 2C T cells were adoptively transferred into B6 recipients. Twenty-third days later, the frequency, phenotype and function of persisting 2C T cells were analyzed by flow cytometry. To assay for recall response, mice were infected with 100pfu WSN-SIY virus and the number and phenotype of 2C T cells in different organs were analyzed by flow cytometry 7 dpi.

Bone marrow chimera mice

Bone marrow cells were collected from the tiba and femur of $2C^+Rag^{-/-}$ mice. Stem and progenitor cells were enriched using a progenitor enrichment kit (Stemcell Technologies) according to the manufacture instructions. The enriched cells were cultured for 48 hours in

complete RPMI medium supplemented with IL-3 (30ng/ml), IL-6 (10ng/ml) and SCF (15ng/ml). The cells were resuspended at 2 x 10^6 cells per ml in complete RPMI. 600μl fresh retrovirus supernatants and 400μl cells plus a final concentration of 6 μg/ml polybrene were added to one well of a 24-well plate and spun for 90min at 2500rpm at 32°C. One the second day, cells were collected and washed and injected into $Rag^{-/-}$ mice that had been irradiated for 500rads 4 hours earlier. Eight weeks later, mice were bleeded to determine the reconstitution of $CD8^+$ T cells and GFP proportion by flow cytometry. Twelve weeks later, cells were collected from spleen and analyzed for 2C T cell percentage and phenotype. Splenocytes containing $5x10^4$ 2C T cells were adoptively transferred into B6 mice followed by WSN-SIY infection. The number, phenotype and function of 2C T cells in the recipient mice were analyzed 7 and 30 dpi. To generate $Bach2^{-/-}$ chimeric mice, bone marrow cells from $Bach2^{-/-}$ mice [49] (kindly gift of Dr. Kazuhiko Igarash of Tohoko University, Japan) were directly injected into sublethally irradiated $Rag2^{-/-}$ mice. Eight weeks later, reconstitution of $CD8^+$ T cells were verified by flow cytometry of peripheral blood mononuclear cells. Twelve weeks later, mice were infected with 50 pfu WSN-SIY virus and 30 dpi SIY-specific $CD8^+$ T cells in various tissues were identified by H-$2K^b$ DimerX (BD Biosciences) loaded with SIY peptide plus anti-CD8 by flow cytometry.

Gene perturbations and quantitative PCR

To perturb the network, selected TFs were overexpressed in $CD8^+$ 2C T cells by retrovirus transduction as described above. Transduced 2C T cells were cultured in the presence of IL-7 for 24 hours and GFP^+ 2C T cells were purified by sorting (>95% viable by PI staining). Total RNA was extracted from the purified 2C T cells using RNeasy micro kit (Qiagen) according to the manufacture's instructions. First strand cDNA was synthesized from 1 μg total RNA using the TaqMan® Reverse Transcription Reagents (ABI). 2 μl of diluted cDNA (total 200 μl) were used as template for the quantitative PCR with LightCycler®480 SYBR Green and LightCycler®480 machine (Roche). For each TF transduced $CD8^+$ T cells, the transcript levels of 12 TFs were measured by qPCR using gene specific primers (Supplementary Table S5). To measure the transcript level of TFs in naïve, effector and memory $CD8^+$ T cells, naïve 2C T cells were adoptively transferred into B6 mice followed with WSN-SIY virus infection, effector and memory 2C T cells were sorted from spleen 7 dpi and 30 dpi, respectively. Total RNA was

isolated from naïve, effector and memory 2C T cells and used for quantification of the transcript level of each TF by PCR.

ChIP and ChIP-PCR

A Millipore ChIP kit was used for chromatin immunoprecipitation. DNA-protein complexes were cross-linked with formaldehyde at a final concentration of 1%, sheared by sonication to 800~1000bp, followed by precipitation with nonspecific goat anti-IgG (Sigma) or rabbit anti-IgG (Cell Signaling Technology) or chromatin ChIP-grade anti-Sox4 (C-20, Santa Cruz Biotechnology), anti-Tcf7 (H-118, Santa Cruz Biotechnology), and anti-Eomes (ab23345, Abcam). DNA-protein complex was eluted, and ChIP DNA was purified by PCR purification kit (Qiagen). The promoter regions of the indicated TFs or MSGs were amplified using specific primers (Supplementary Table S6). Primers used to amplify the promoter regions were all within this 1 kb upstream of the transcription-starting site. For ChIP of Bach2 with anti-Bach2 (E-16, Santa Cruz Biotechnology), cross-linked DNA-protein complexes were digested to 400~600bp by micrococcal nuclease (Cell Signaling Technology).

Protein extraction and western blotting

Proteins were extracted from transduced 2C T cells with the CelLytic^{TM} Lysis Reagent (Sigma). Samples containing 20μg total protein (BCA™ Protein Assay Kit, Pierce Biotechnology) were resolved on a 10% SDS-PAGE gel and electro-transferred onto a PVDF membrane (Millipore Corporation). The membrane was blocked in 5% (w/v) fat-free milk in PBST (PBS containing 0.1% Tween-20). The blot was hybridized overnight with primary antibodies: anti-GAPDH (HRP-conjugated, Cell Signaling Technology, 1:2000), anti-Sox4 (C-20, Santa Cruz Biotechnology, 1:500), anti-Bach2 (AP10133b, Abgent, 1:500), anti-Id3 (6-1, CalBioreagents, 1:2500), anti-Blimp-1 (6D3, eBioscience, 1:1000), anti-Bcl2 (BioLegend, 1:500), anti-Bcl6 (BioLegend, 1:2000), anti-pFoxO1 (Cell Signaling Technology, 1:1000), anti-CDK4 (Cell Signaling Technology, 1:1000), anti-CDK6 (Cell Signaling Technology, 1:1000), anti-Rb-p780 (Cell Signaling Technology, 1:1000), anti-Rb-p795 (Cell Signaling Technology, 1:1000) and anti-p27kip (Cell Signaling Technology, 1:1000) according to the recommended dilution in 5% fat free milk. The blot was washed twice in PBST and then incubated with HRP-conjugated secondary antibody (Cell Signaling Technology: anti-Rabbit, 1:2000; anti-mouse,1:3000. Santa

Cruz Biotechnology: anti-Rat, 1:2000; anti-Goat 1:3000) in 5% fat-free milk. The membrane was washed twice in PBST and subjected to protein detection by ECL Plus Western Blotting Detection System (GE Healthcare) before being exposed to a Koda BioMax XAR film. The membrane was stripped and re-blotted with the rabbit anti-mouse HRP-conjugated anti-Gapdh antibody (Cell Signaling Technology) for protein loading control.

Statistical analysis

Statistical significance was determined with the two-tailed unpaired or paired Student's t-test. *P*-values for MRA and promoter enrichment results were calculated with a binomial test. The FDRs were computed with $q = p*n/i$, ($p = P$ value, n = total number of tests, i = sorted rank of P value).

References

1. Goldrath AW, Bevan MJ. Selecting and maintaining a diverse T-cell repertoire. *Nature* **402**, 255-261 (1999).

2. Kaech SM, Hemby S, Kersh E, Ahmed R. Molecular and functional profiling of memory CD8 T cell differentiation. *Cell* **111**, 837-851 (2002).

3. Wherry EJ*, et al.* Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity* **27**, 670-684 (2007).

4. Intlekofer AM*, et al.* Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. *Nat Immunol* **6**, 1236-1244 (2005).

5. Pearce EL*, et al.* Control of effector CD8+ T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041-1043 (2003).

6. Sullivan BM, Juedes A, Szabo SJ, von Herrath M, Glimcher LH. Antigen-driven effector CD8 T cell function regulated by T-bet. *Proc Natl Acad Sci U S A* **100**, 15818-15823 (2003).

7. Gattinoni L*, et al.* Wnt signaling arrests effector T cell differentiation and generates CD8+ memory stem cells. *Nat Med* **15**, 808-813 (2009).

8. Zhao DM*, et al.* Constitutive activation of Wnt signaling favors generation of memory CD8 T cells. *J Immunol* **184**, 1191-1199 (2010).

9. Zhou X, Yu S, Zhao DM, Harty JT, Badovinac VP, Xue HH. Differentiation and persistence of memory CD8(+) T cells depend on T cell factor 1. *Immunity* **33**, 229-240 (2010).

10. Schober SL, Kuo CT, Schluns KS, Lefrancois L, Leiden JM, Jameson SC. Expression of the transcription factor lung Kruppel-like factor is regulated by cytokines and correlates with survival of memory T cells in vitro and in vivo. *J Immunol* **163**, 3662-3667 (1999).

11. Bai A, Hu H, Yeung M, Chen J. Kruppel-like factor 2 controls T cell trafficking by activating L-selectin (CD62L) and sphingosine-1-phosphate receptor 1 transcription. *J Immunol* **178**, 7632-7639 (2007).

12. Carlson CM*, et al.* Kruppel-like factor 2 regulates thymocyte and T-cell migration. *Nature* **442**, 299-302 (2006).

13. Yang CY*, et al.* The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. *Nat Immunol* **12**, 1221-1229 (2011).

14. Cannarile MA*, et al.* Transcriptional regulator Id2 mediates CD8+ T cell immunity. *Nat Immunol* **7**, 1317-1325 (2006).

15. Ji Y*, et al.* Repression of the DNA-binding inhibitor Id3 by Blimp-1 limits the formation of memory CD8+ T cells. *Nat Immunol* **12**, 1230-1237 (2011).

16. Kallies A, Xin A, Belz GT, Nutt SL. Blimp-1 transcription factor is required for the differentiation of effector CD8(+) T cells and memory responses. *Immunity* **31**, 283-295 (2009).

17. Rutishauser RL*, et al.* Transcriptional repressor Blimp-1 promotes CD8(+) T cell terminal differentiation and represses the acquisition of central memory T cell properties. *Immunity* **31**, 296-308 (2009).

18. Gray PA*, et al.* Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* **306**, 2255-2257 (2004).

19. Faith JJ*, et al.* Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol* **5**, e8 (2007).

20. Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, Group MGD. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic acids research* **40**, D881-886 (2012).

21. Angelosanto JM, Wherry EJ. Transcription factor regulation of CD8+ T-cell memory and exhaustion. *Immunol Rev* **236**, 167-175 (2010).

22. Rutishauser RL, Kaech SM. Generating diversity: transcriptional regulation of effector and memory CD8 T-cell differentiation. *Immunol Rev* **235**, 219-233 (2010).

23. D'Cruz LM, Rubinstein MP, Goldrath AW. Surviving the crash: transitioning from effector to memory CD8+ T cell. *Semin Immunol* **21**, 92-98 (2009).

24. Kaech SM, Cui W. Transcriptional control of effector and memory CD8(+) T cell differentiation. *Nat Rev Immunol* **12**, 749-761 (2012).

25. Joshi NS*, et al.* Inflammation directs memory precursor and short-lived effector CD8(+) T cell fates via the graded expression of T-bet transcription factor. *Immunity* **27**, 281-295 (2007).

26. Shen CH, Ge Q, Talay O, Eisen HN, GarcÌa-Sastre A, Chen J. Loss of IL-7R and IL-15R expression is associated with disappearance of memory T cells in respiratory tract following influenza infection. *J Immunol* **180**, 171-178 (2008).

27. Hettmann T, Opferman JT, Leiden JM, Ashton-Rickardt PG. A critical role for NF-kappaB transcription factors in the development of CD8+ memory-phenotype T cells. *Immunol Lett* **85**, 297-300 (2003).

28. Teixeiro E*, et al.* Different T cell receptor signals determine CD8+ memory versus effector development. *Science* **323**, 502-505 (2009).

29. Alon U. Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**, 450-461 (2007).

30. Kaech SM, Wherry EJ. Heterogeneity and cell-fate decisions in effector and memory CD8+ T cell differentiation during viral infection. *Immunity* **27**, 393-405 (2007).

31. Bernard P, Harley VR. Acquisition of SOX transcription factor specificity through protein-protein interaction, modulation of Wnt signalling and post-translational modification. *Int J Biochem Cell Biol* **42**, 400-410 (2010).

32. Sinner D*, et al.* Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol Cell Biol* **27**, 7802-7815 (2007).

33. Ikushima H, Todo T, Ino Y, Takahashi M, Miyazawa K, Miyazono K. Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell Stem Cell* **5**, 504-514 (2009).

34. Kuwahara M*, et al.* The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF-β and suppresses T(H)2 differentiation. *Nat Immunol* **13**, 778-786 (2012).

35. Sanjabi S, Mosaheb MM, Flavell RA. Opposing effects of TGF-beta and IL-15 cytokines control the number of short-lived effector CD8+ T cells. *Immunity* **31**, 131-144 (2009).

36. Masson F*, et al.* Id2-mediated inhibition of E2A represses memory CD8+ T cell differentiation. *J Immunol* **190**, 4585-4594 (2013).

37. Ochiai K*, et al.* Plasmacytic transcription factor Blimp-1 is repressed by Bach2 in B cells. *J Biol Chem* **281**, 38226-38234 (2006).

38. Buchholz VR*, et al.* Disparate individual fates compose robust CD8+ T cell immunity. *Science* **340**, 630-635 (2013).

39. Gerlach C*, et al.* Heterogeneous differentiation patterns of individual CD8+ T cells. *Science* **340**, 635-639 (2013).

40. Roychoudhuri R*, et al.* BACH2 represses effector programs to stabilize T(reg)-mediated immune homeostasis. *Nature* **498**, 506-510 (2013).

41. Tsukumo S*, et al.* Bach2 maintains T cells in a naive state by suppressing effector memory-related genes. *Proc Natl Acad Sci U S A* **110**, 10735-10740 (2013).

42. Rivera R, Murre C. The regulation and function of the Id proteins in lymphocyte development. *Oncogene* **20**, 8308-8316 (2001).

43. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP, Staudt LM. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* **13**, 199-212 (2000).

44. Crotty S, Johnston RJ, Schoenberger SP. Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation. *Nat Immunol* **11**, 114-120 (2010).

45. Ichii H, Sakamoto A, Kuroda Y, Tokuhisa T. Bcl6 acts as an amplifier for the generation and proliferative capacity of central memory CD8+ T cells. *J Immunol* **173**, 883-891 (2004).

46. Singh A, Jatzek A, Plisch EH, Srinivasan R, Svaren J, Suresh M. Regulation of memory CD8 T-cell differentiation by cyclin-dependent kinase inhibitor p27Kip1. *Mol Cell Biol* **30**, 5145-5159 (2010).

47. Cartharius K*, et al.* MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**, 2933-2942 (2005).

48. Chen J, Eisen HN, Kranz DM. A model T-cell receptor system for studying memory T-cell development. *Microbes Infect* **5**, 233-240 (2003).

49. Muto A*, et al.* The transcriptional programme of antibody class switching involves the repressor Bach2. *Nature* **429**, 566-571 (2004).

50. Kuwahara M*, et al.* The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF-beta and suppresses T(H)2 differentiation. *Nat Immunol* **13**, 778-786 (2012).

51. Sun H, Lu B, Li RQ, Flavell RA, Taneja R. Defective T cell activation and autoimmune disorder in Stra13-deficient mice. *Nat Immunol* **2**, 1040-1047 (2001).

52. Vaillant F, Blyth K, Andrew L, Neil JC, Cameron ER. Enforced expression of Runx2 perturbs T cell development at a stage coincident with beta-selection. *J Immunol* **169**, 2866-2874 (2002).

53. Li Q, Eppolito C, Odunsi K, Shrikant PA. IL-12-programmed long-term CD8+ T cell responses require STAT4. *J Immunol* **177**, 7618-7625 (2006).

54. Rao RR, Li Q, Odunsi K, Shrikant PA. The mTOR kinase determines effector versus memory CD8+ T cell fate by regulating the expression of transcription factors T-bet and Eomesodermin. *Immunity* **32**, 67-78 (2010).

55. Cruz-Guilloty F*, et al.* Runx3 and T-box proteins cooperate to establish the transcriptional program of effector CTLs. *J Exp Med* **206**, 51-59 (2009).

56. Karwot R*, et al.* Protective role of nuclear factor of activated T cells 2 in CD8+ long-lived memory T cells in an allergy model. *J Allergy Clin Immunol* **121**, 992-999.e996 (2008).

57. Zheng W, Flavell RA. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89**, 587-596 (1997).

58. Cortes M, Wong E, Koipally J, Georgopoulos K. Control of lymphocyte development by the Ikaros gene family. *Curr Opin Immunol* **11**, 167-171 (1999).

59. Oestreich KJ, Yoon H, Ahmed R, Boss JM. NFATc1 regulates PD-1 expression upon T cell activation. *J Immunol* **181**, 4832-4839 (2008).

60. Isakov N, Altman A. Protein kinase C(theta) in T cell activation. *Annu Rev Immunol* **20**, 761-794 (2002).

61. Kim HJ, Nel AE. The role of phase II antioxidant enzymes in protecting memory T cells from spontaneous apoptosis in young and old mice. *J Immunol* **175**, 2948-2959 (2005).

End notes

Acknowledgements

The authors thank Ching-Hung Shen, Zhuyan Guo, and Camille M. Justino for technique assistance, Pete S. Bak and Herman Eisen for discussion, and Professor Kazuhiko Igarash for providing Bach$^{-/-}$ bone marrow cells. This work was supported in part by National Institutes of Health Grant AI69208, funds from the Singapore-MIT Alliance, and Ivan R. Cottrell Professorship and Research Fund (to J.C.), and the Koch Institute Support (core) Grant P30-CA14051 from the National Cancer Institute.

Author contributions: G.H. and C.J. conceived and designed the study. G.H. carried out all computations and experimentations. G.H. and C.J. performed data analysis and drafted the paper. All authors read and approved the final manuscript.

Competing financial interests: The authors declare no competing financial interests.

Figure Legends

Figure 1 | Construction of regulatory network of memory CD8$^+$ T cells. (**a)** Schematic diagram of regulatory network analysis for identifying key TFs. N, E, and M, naïve, effector and memory CD8$^+$ T cells, respectively. (**b**) The regulatory module of the top 10 TFs (orchid circles) and their MSGs (blue). **c**, ChIP-PCR analysis of Sox4, Tcf7 and Emoes-regulated MSGs. ChIP was carried out with CD8$^+$ T cells expressing the 2C TCR using antibodies specific for Sox4, Tcf7 or Eomes or control IgG antibodies. Promoter regions of the indicted genes were amplified using the precipitated DNA. Shown are PCR products after electrophoresis.

Figure 2 | Construction of perturbation network of TFs in CD8$^+$ T cells. (**a**) ChIP-PCR analysis. ChIP was carried out with 2C T cells using antibodies specific for Sox4, Tcf7 or Eomes or control IgG antibodies. Promoter regions of the indicted genes were then amplified using the precipitated DNA. Shown are PCR products after electrophoresis. (**b**) Perturbation network based on Table 3. (**c**) An example of network motifs from the perturbation network.

Figure 3 | Effect of overexpression and knockdown of TFs on memory CD8$^+$ T-cell recall proliferation. Naïve 2C T cells were activated *in vitro* with SIY peptide and then transduced with retroviruses expressing *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* or expressing shRNA specific for one of the TFs. The cells were cultured in the presence of IL-7 to induce memory T cell development. The resulting memory 2C T cells were either activated *in vitro* with SIY peptide or transferred into mice and activated by WSN-SIY virus infection. The proportion of GFP$^+$ (transduced) versus GFP$^-$ (non-transduced) 2C T cells was quantified 4 and 6 days post stimulation *in vitro* and in draining lymph node (DLN) 5 days post infection (dpi). Shown are proportion of GFP$^+$ 2C T cells that overexpressed *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* among total 2C T cells *in vitro* (**a**) and *in vivo* (**b**) or that expressed shRNA specific for *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* among total 2C T cells is *in vitro* (**c**) and *in vivo* (**d**). Each line was one independent experiment with one sample per time point for the *in vitro* experiments and one or two mice per in vivo experiment. Data shown are mean ± s.e.m. Pairwise two-tailed t-tests were used for statistical analyses. * P<0.05; ** P<0.01.

Figure 4 | *Bach2* promotes recall proliferation of memory CD8$^+$ T cells. (**a**) Scheme of experimental protocol. (**b-d)** Phenotype and function of persisting memory 2C cells. Twenty-two days post transfer, single cell suspension was prepared from spleen and analyzed for CD8, Thy1.1, GFP plus CD62L or IL-7R directly or stimulated *in vitro* with SIY peptide for 5 hours before staining for CD8, Thy1.1, GFP plus intracellular IFNγ or TNFα. Comparison of GFP versus Thy1.1 (**b**) staining profiles of live cells between vector control and *Bach2* overexpression group. Comparison of CD62L and IL-7R (**c**) or IFNγ and TNFα expression (**d**) between GFP$^+$ and GFP$^-$ 2C T cells. Gray trace, nontransduced (GFP$^-$) 2C T cells; black trace, transduced (GFP$^+$) 2C T cells; dash gray trace, isotype control for intracellular staining; dash black trace, IL-7R staining of naïve 2C T cells. (**e**) Recall responses of persisting memory 2C T cells *in vivo*. Some recipient mice were infected with WSN-SIY 23-25 days post transfer and the proportions of GFP$^+$ and GFP$^-$ 2C T cells in the DLN and spleen was quantified by flow cytometry 7 dpi. Comparison of proportions of GFP$^+$ 2C T cells in the DLN and spleen before (d23) and after antigen restimulation (7dpi). Representative data from three independent experiments with 2-3 mice per group per experiment are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses. ** P<0.01; n.s., not significant.

Figure 5 | *Bach2* promotes memory CD8$^+$ T cell development. (**a**) Scheme of experimental protocol for **b-f**. (**b-e)** Effect of Bach2 overexpression on memory T cell development. **b,c**, Persistence and phenotype of transferred 2C T cells over time. Seven and 30 dpi, 2C T cells in various organs were analyzed for CD62L, IL-7R, Klrg1, CD27 and IFNγ as in **Fig. 4**. Shown are histograms of CD62L, IL-7R, Klrg1 and CD27 expression of Thy1.1$^+$ CD8$^+$ 2C T cells 7 (**b**) and 30 dpi (**c**). Gray trace, nontransduced (GFP$^-$) 2C T cells; black trace, transduced (GFP$^+$) 2C T cells. Representative data from 6 mice in 3 independent experiments are shown. (**d**) Proportion of GFP$^+$ transduced 2C cells in different organs normalized to the average of the blood at 7 dpi. (**e**) Total Bach2-transduced 2C T cells (GFP$^+$) in the spleen 7 and 30 dpi. (**f**) Effect of Bach2 knockdown on memory T cell development. Proportion of Bach2-knockdown 2C T cells (GFP$^+$) in different organs normalized to the average of the blood at 7 dpi. (**g**) Effect of Bach2 knockout on memory T cell development. Chimera mice were constructed by injecting Bach2$^{-/-}$ and Bach2$^{+/+}$ bone marrow cells into sublethally irradieated Rag2$^{-/-}$ recipient mice. Following reconstitution (3 months later), mice were infected with WSN-SIY virus and 30 dpi cells from DLN and spleen were stained for H-2K^b-SIY and anti-CD8. Shown are staining profiles of H-2K^b-SIY versus CD8. The numbers in the plots indicate percentage of SIY-specific memory CD8$^+$ T cells. The numbers in the boxes indicate the number of SIY-specific memory CD8$^+$ T cells (top) and total CD8$^+$ T cells (bottom). Representative data from three independent experiments with 2-3 mice per group per experiment (d, e) and from two independent experiments with 3-4 mice per group per experiment (f) are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses. * P<0.05; ** P<0.01.

Figure 6 | *Bach2* promotes memory CD8$^+$ T cell development. (**a**) Scheme of experimental protocol. (**b-e**) Persistence and phenotype of transferred CD8$^+$ T cells over time. Seven and 30 dpi, 2C T cells in various organs were analyzed as in **Fig. 4**. Shown are histograms of CD62L, IL-7R, Klrg1, IFNγ and TNFα expression of Thy1.1$^+$ CD8$^+$ 2C T cells 7 (**b**) and 30 dpi (**c**). Gray trace, nontransduced (GFP$^-$) 2C T cells; black trace, transduced (GFP$^+$) 2C T cells; dash gray trace, isotype control for intracellular staining. Representative data from 9-12 mice in 3 independent experiments are shown. (**d**) Proportion of GFP$^+$ transduced 2C cells in different organs normalized to the average of the blood at 7 dpi. (**e)** Total Bach2-transduced 2C T cells

(GFP^+) in the spleen 7 and 30 dpi. Data from three independent experiments with 3-4 mice per group per experiment are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses. * P<0.05; ** P<0.01.

Figure 7 | *Bach2* promotes proliferation of $CD8^+$ T cells. **(a)** 2C cells were activated *in vitro* and transduced with retroviruses expressing GFP alone (vector) or GFP plus Bach2 and cultured in the presence of either IL-2 or IL-7. The proportion of GFP^+ 2C T cells was quantified. Shown (mean ± s.e.m.) are changes in proportion of GFP^+ cells over time from four independent experiments. Two-tailed student's t-tests were used for statistical analyses. ** P<0.01. (**b**) *Bach2* transduced 2C T cells (day 4 in **Fig. 7a**) were labeled with eFluor®670 and cultured in the presence of IL-2 for four days. Shown are representative histograms of eFluor®670 from one of the two experiments. Gray trace, nontransduced (GFP^-) 2C T cells; black trace, transduced (GFP^+) 2C T cells. (**c**) *Sox4-Bach2-Prdm1-Id3* regulatory motif identified in perturbation network. (**d**) Bach2 inhibits Prdm1 expression but promotes Id3 expression. Activated 2C T cells were transduced with retroviruses expressing GFP alone (vector) or GFP plus *Bach2*, *Sox4* or *Id3*. The levels of the indicated TFs were assayed by Western blotting. Shown are representative Western blotting and the average expression level quantified from three independent experiments. (**e-f**) Bach2 binds to *Prdm1* and *Id3* promoter. Activated 2C T cells were cultured with IL-7 for 24 hours and harvested for ChIP using anti-Bach2. DNA was used to amplify different parts of the *Prdm1* (**e**) and *Id3* (**f**) promoter region indicated by i-v. The location of the predicted Bach2 binding motif was indicted as triangle. Data are from two independent experiments, error bar: SEM. (**g**) *Bach2* affects Rb phosphorylation and p27kip expression. Activated 2C T cells were transduced with retroviruses expressing GFP alone (vector) or GFP plus Bach2 and cultured for 8 days (**Fig. 7a**). The levels of phosphorylated Rb and FoxO1, CDK4, CDK6, Bcl6, Bcl2 and p27kip in the transduced 2C T cells were assayed by Western blotting.

Table 1. Ranking of the 21 key TFs identified by master regulator analysis.

TF	#MSG	P-value	ISP*	DBM	FDR	ESG	MSG	qPCR E/N	qPCR M/N	Reference gene functions[#]	Ref.
Sox4	50	9.55E-10	YES	SORY	0.0001	YES	YES	-4.305	-5.845	Th2 differentiation	[50]
Tcf7	40	1.10E-06	YES	LEFF	0.0010	YES	NO	-3.71	-0.37	Promote memory	[7, 9]
Eomes	34	1.03E-06	YES	BRAC	0.0021	YES	YES	1.99	3.48	Effector and memory	[4, 5]
Bhlhe40	27	2.50E-05	YES	HESF	0.3353	YES	YES	2.72	4.215	CD8+ T cell activation	[51]
Prdm1	26	2.90E-05	YES	PRDF	-	YES	NO	4.185	4.5	Effector/inhibit memory	[16, 17]
Klf2	24	0.000587	YES	KLFS	0.1135	NO	NO	-2.01	-0.63	Homeostasis/memory	[10]
Bach2	22	0.000229	YES	AP1R	0.0151	YES	NO	-3.77	-0.935	-	
Runx2	21	0.000365	YES	HAML	0.0031	YES	YES	0.19	2.775	Early T cell development	[52]
Id2	20	0.000319	YES	-	-	YES	YES	0.265	2.645	KO promotes memory	[13, 14]
Stat4	17	0.002650	YES	STAT	0.2503	NO	NO	-2.33	0.935	Effector and memory	[53, 54]
Runx3	15	0.002869	YES	HAML	0.0031	NO	NO	-2.37	-1.925	Cooperation with Tbx21	[55]
Id3	14	0.005424	YES	-	-	NO	NO	-3.695	-1.995	Promote memory	[15]
Nfatc2	12	0.007564	YES	NFAT	0.0127	NO	NO	-	-	TCR signaling	[56]
Gata3	11	0.012973	YES	GATA	0.0425	NO	NO	-	-	Th2 differentiation	[57]
Ikzf1	11	0.038617	YES	IKRS	0.0208	NO	NO	-	-	IL2/TCR signaling	[58]
Nfatc1	11	0.038617	YES	NFAT	0.0127	NO	NO	-	-	PD-1/TCR signaling	[59]
Jun	10	0.019089	YES	AP1F	0.0046	NO	NO	-	-	AP-1 complex	[60]
Nfe2l2	10	0.019089	YES	AP1R	0.0151	NO	NO	-	-	Protect memory	[61]
Tbx21	9	0.036556	YES	BRAC	0.0021	NO	NO	6.215	3.165	Effector and memory	[4, 5]
Jund	9	0.028085	YES	AP1F	0.0046	NO	NO	-	-	AP-1 complex	[60]
Maf	9	0.025971	YES	AP1R	0.0151	YES	NO	-	-	-	

#MSG: the total number of MSG regulated directly by the TF; * ISP: immune system phenotype; DBM, DNA binding motif; FDR, false discovery rate; ESG, effector signature genes; MSG, memory signature genes; qPCR E/N, log2 gene expression fold-changes between effector and naïve $CD8^+$ T cells; qPCR M/N, log2 gene expression fold-changes between memory and naïve $CD8^+$ T cells; [#]gene function reported in T cells. Grey marked TFs that have been reported as playing important roles in memory $CD8^+$ T cell development or function.

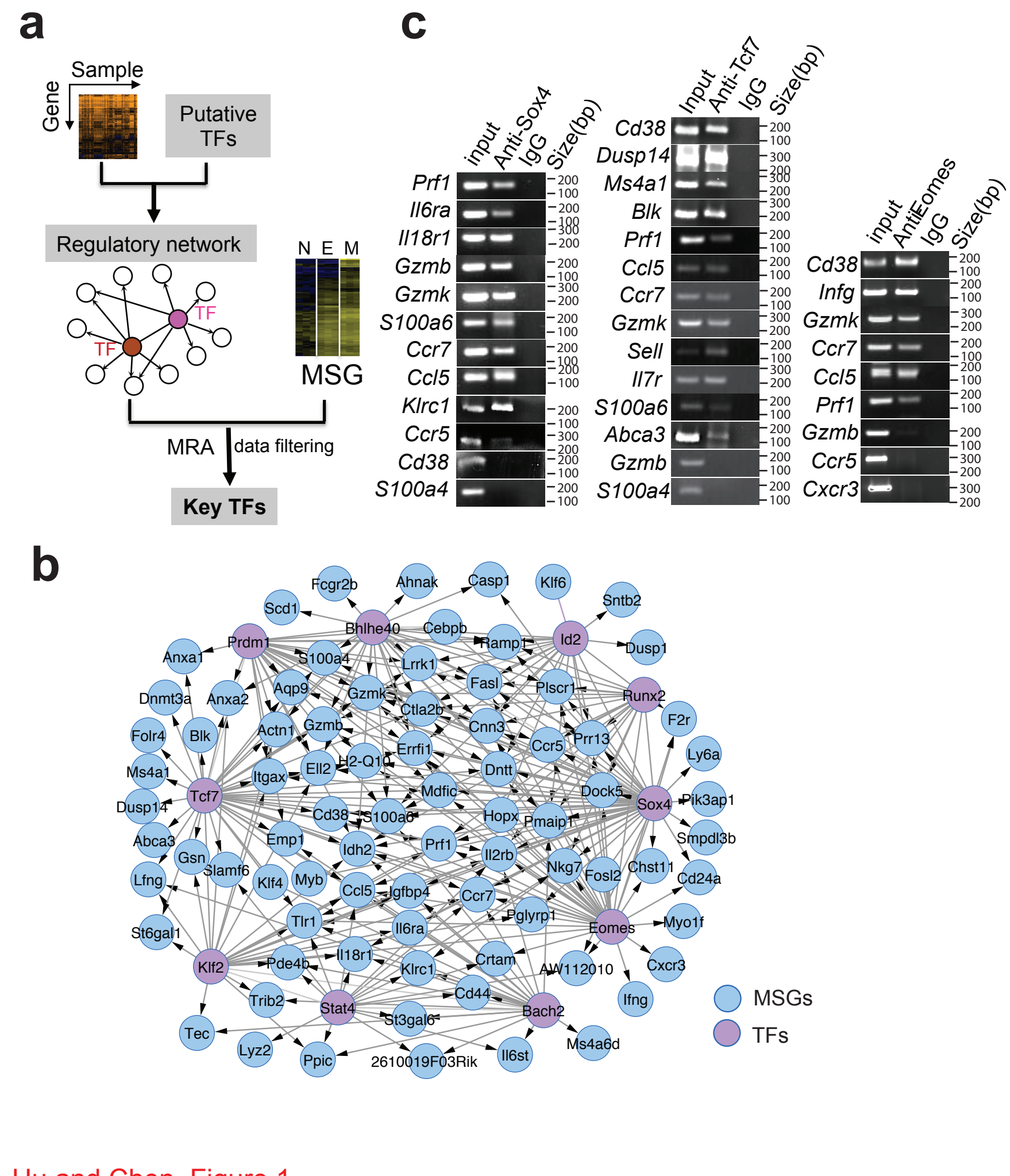
Table 2. Summary of the ChIP-PCR results.

TF	Sox4	Tcf7	Eomes	Total
No. MSG target	41	36	26	103
No. tested target	12	14	9	35
No. positive	10 (83.3%)	12 (85.7%)	6 (66.7%)	28 (80%)

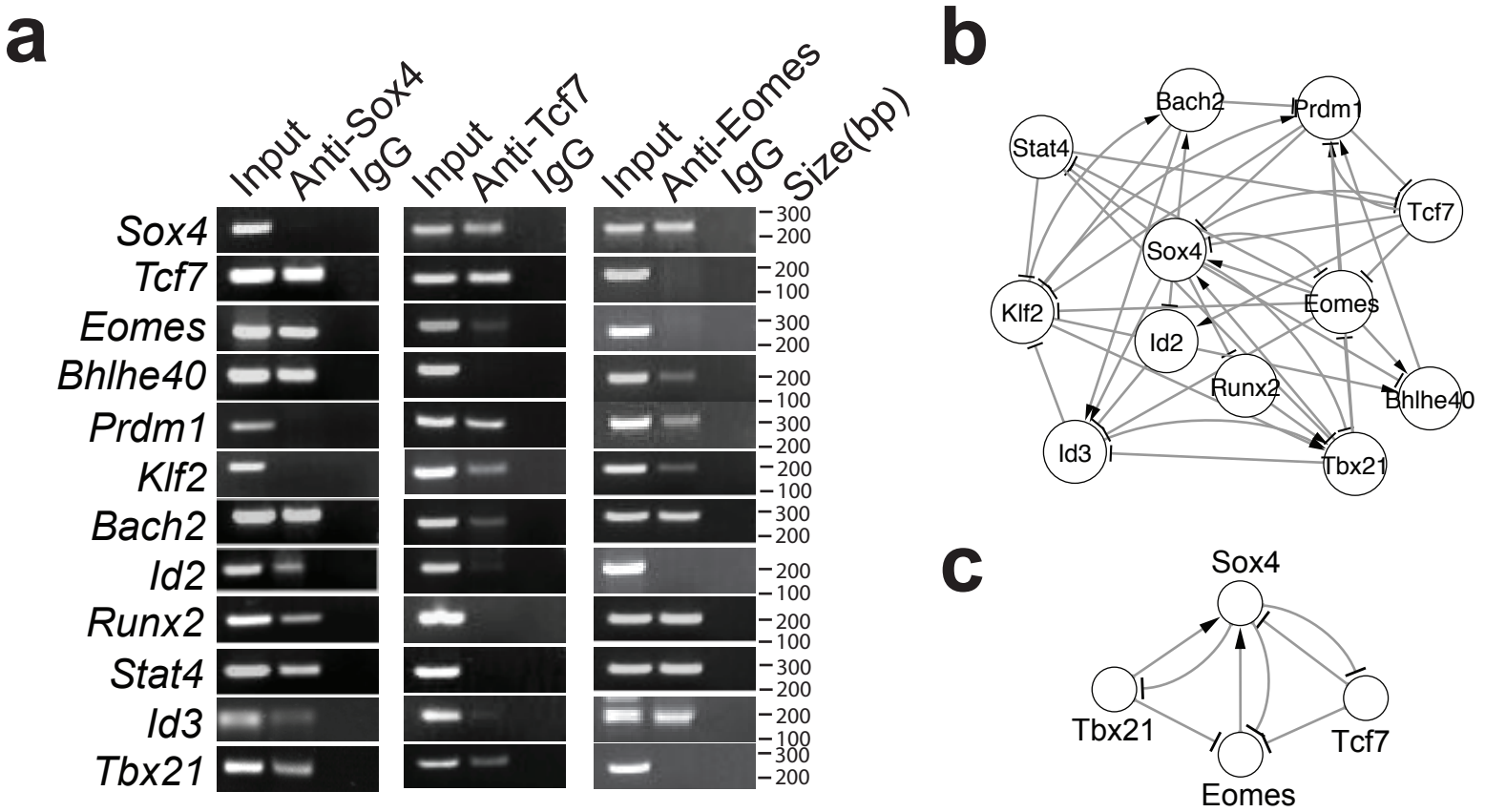
Table 3. Perturbation analysis.

Gene	***Sox4***	***Tcf7***	***Eomes***	***Bhlhe40***	***Prdm1***	***Klf2***	***Bach2***	***Id2***	***Runx2***	***Stat4***	***Id3***	***Tbx21****
Sox4	7.94	-1.38	1.77	-0.62	-3.17	-0.28	-0.93	-0.48	-0.05	-0.70	-0.83	1.21
Tcf7	-1.79	2.44	-0.03	-0.21	-2.01	-0.50	0.88	-0.39	0.04	-2.16	0.31	-0.64
Eomes	-2.28	-1.11	5.29	0.38	-0.08	-0.26	-0.64	-0.10	0.85	-0.55	0.54	-2.31
Bhlhe40	-2.03	0.41	1.09	3.88	0.02	1.99	0.32	0.31	0.14	-0.41	0.68	0.62
Prdm1	-0.14	-1.48	2.04	1.13	5.19	1.34	-1.62	-0.17	0.20	0.08	-0.39	1.30
Klf2	0.85	-0.78	-1.10	-0.55	-1.49	4.06	-1.52	-0.73	-0.07	-1.33	-1.51	-0.57
Bach2	1.37	0.75	0.35	0.00	0.22	1.40	4.41	-0.18	0.78	0.01	0.07	0.15
Id2	-2.24	1.10	0.46	0.38	0.24	0.66	0.09	4.98	0.93	-0.06	0.18	0.54
Runx2	-1.64	0.41	-0.51	-0.22	-0.02	0.91	-0.50	-0.82	3.71	-0.54	-0.03	-0.77
Stat4	-1.45	-0.26	-1.78	0.14	0.80	0.24	0.22	-0.01	0.85	2.32	0.41	-0.81
Id3	2.61	0.23	-1.32	0.47	-0.23	0.72	2.02	-1.03	0.25	0.73	7.44	-1.09
Tbx21	-1.68	-0.26	-0.22	0.12	-0.34	1.42	0.50	0.93	1.29	-2.29	1.17	6.54

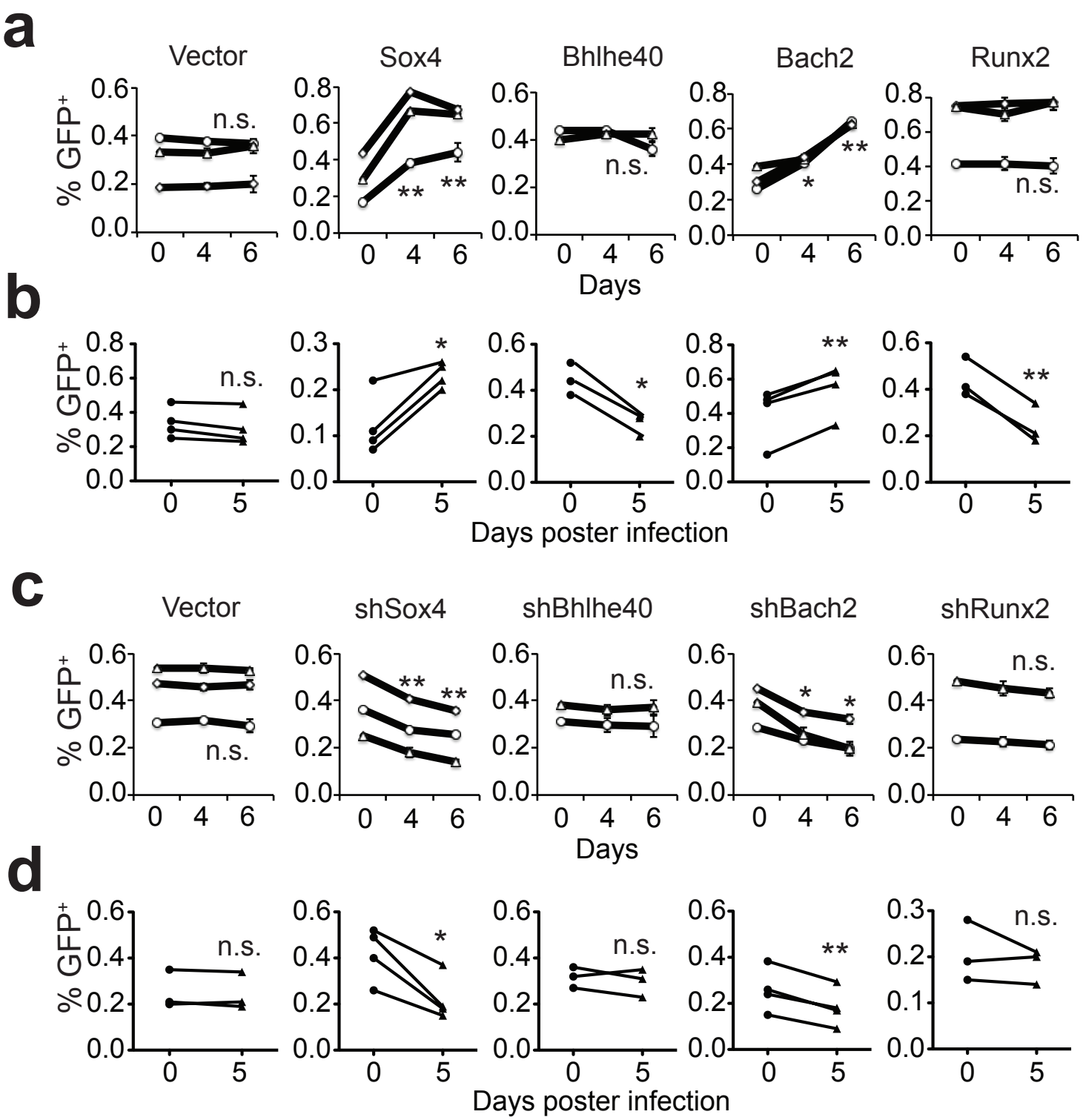
* The highlighted 12 TFs (first row) were overexpressed individually in 2C T cells and the level of their transcripts of each TFs was quantified by real-time PCR. Expression data was normalized to the empty vector control and then log2 transformed. Changes in transcript levels for ≥2 fold are marked orange (up-regulated) or green (down-regulated). The overexpressed TFs are marked red.



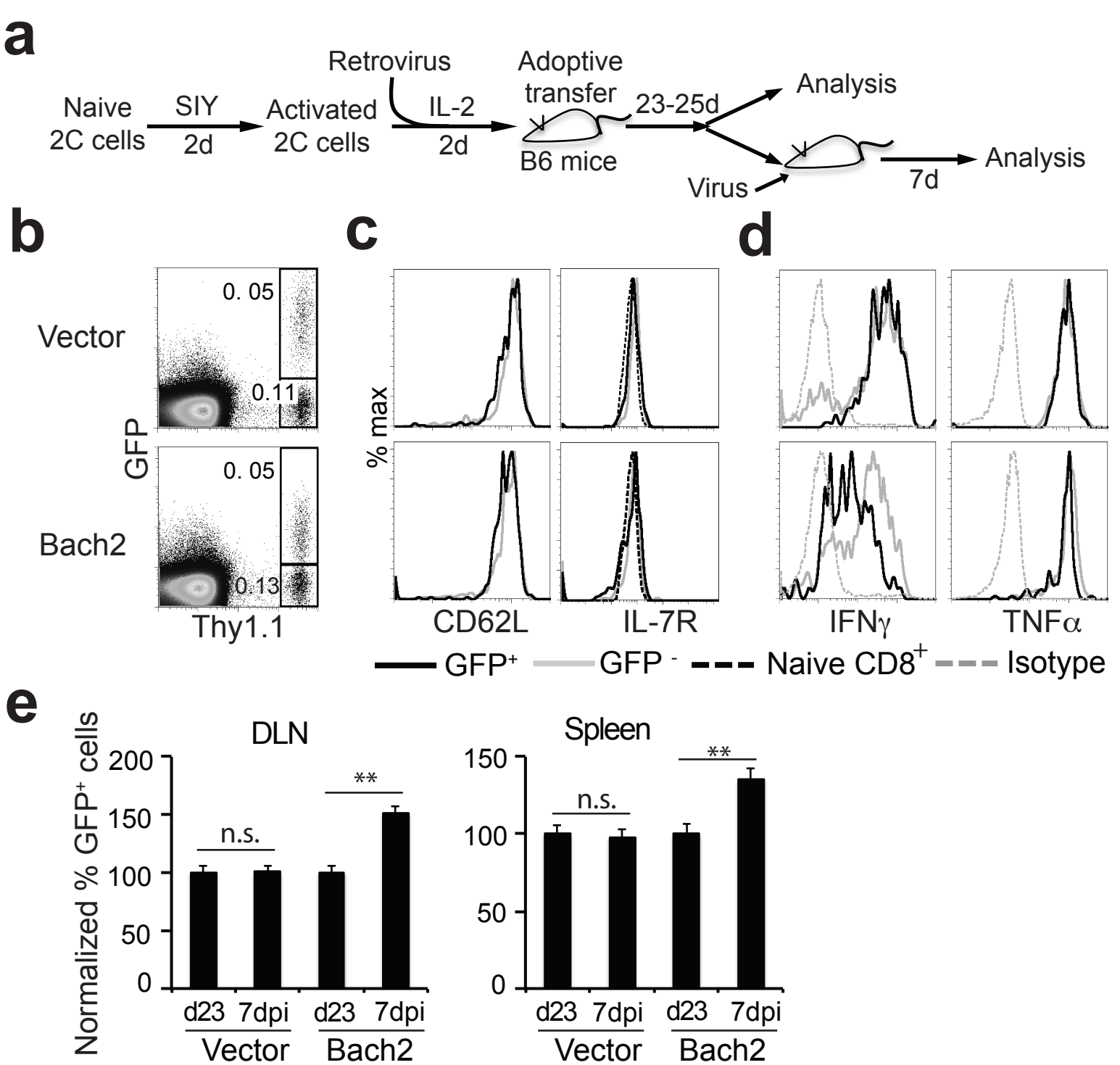
Hu and Chen, Figure 1



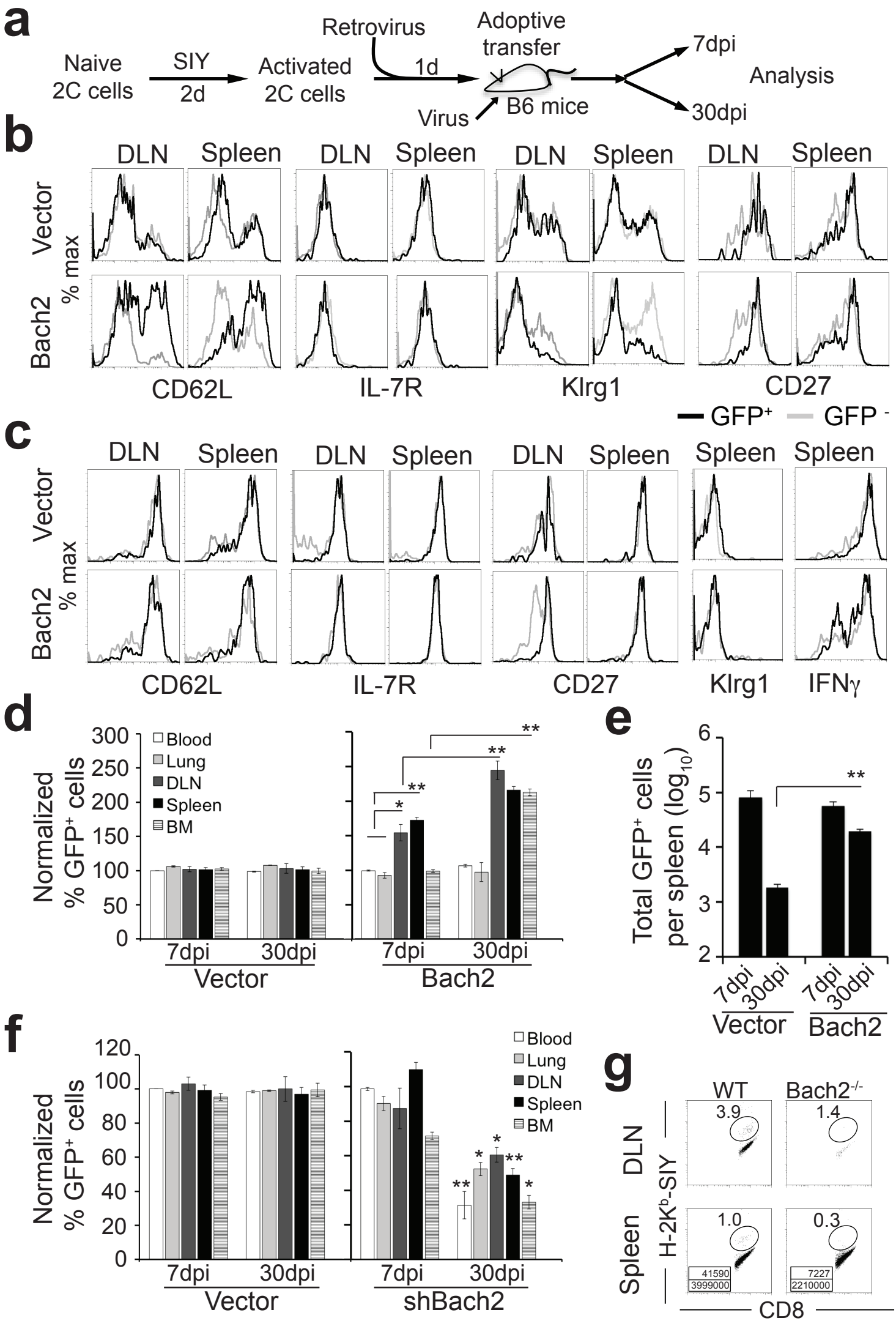
Hu and Chen, Figure 2



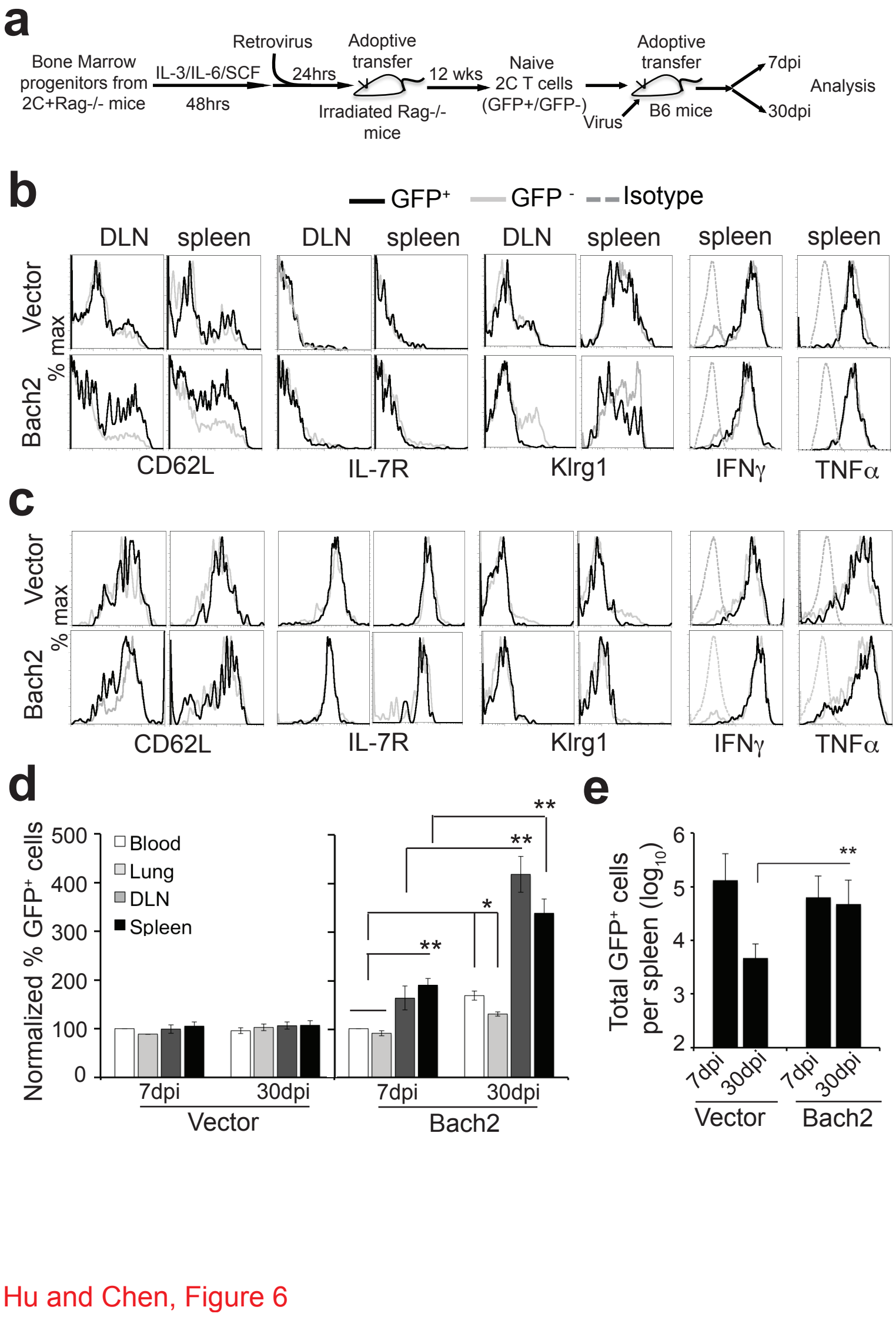
Hu and Chen, Figure 3



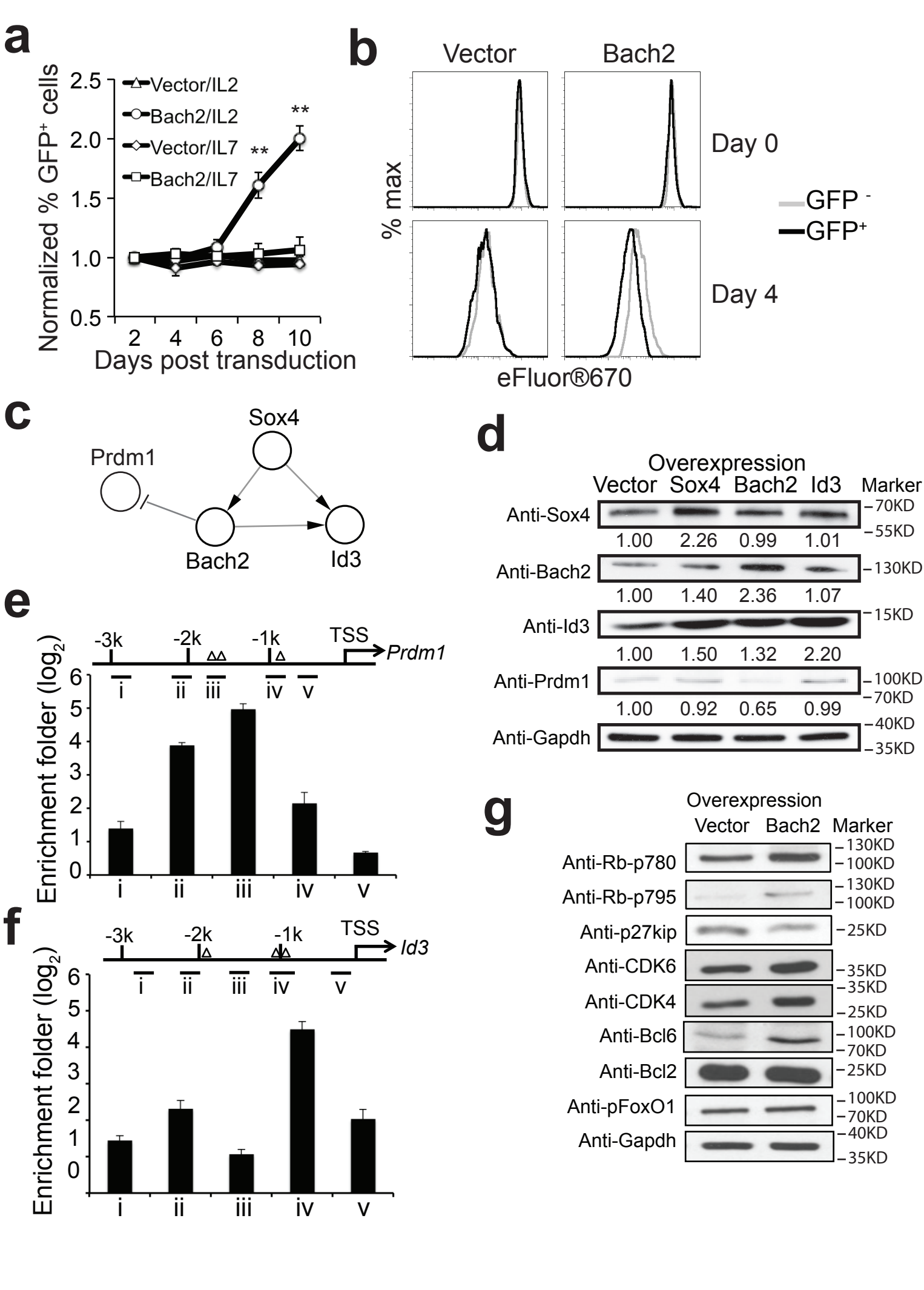
Hu and Chen, Figure 4



Hu and Chen, Figure 5

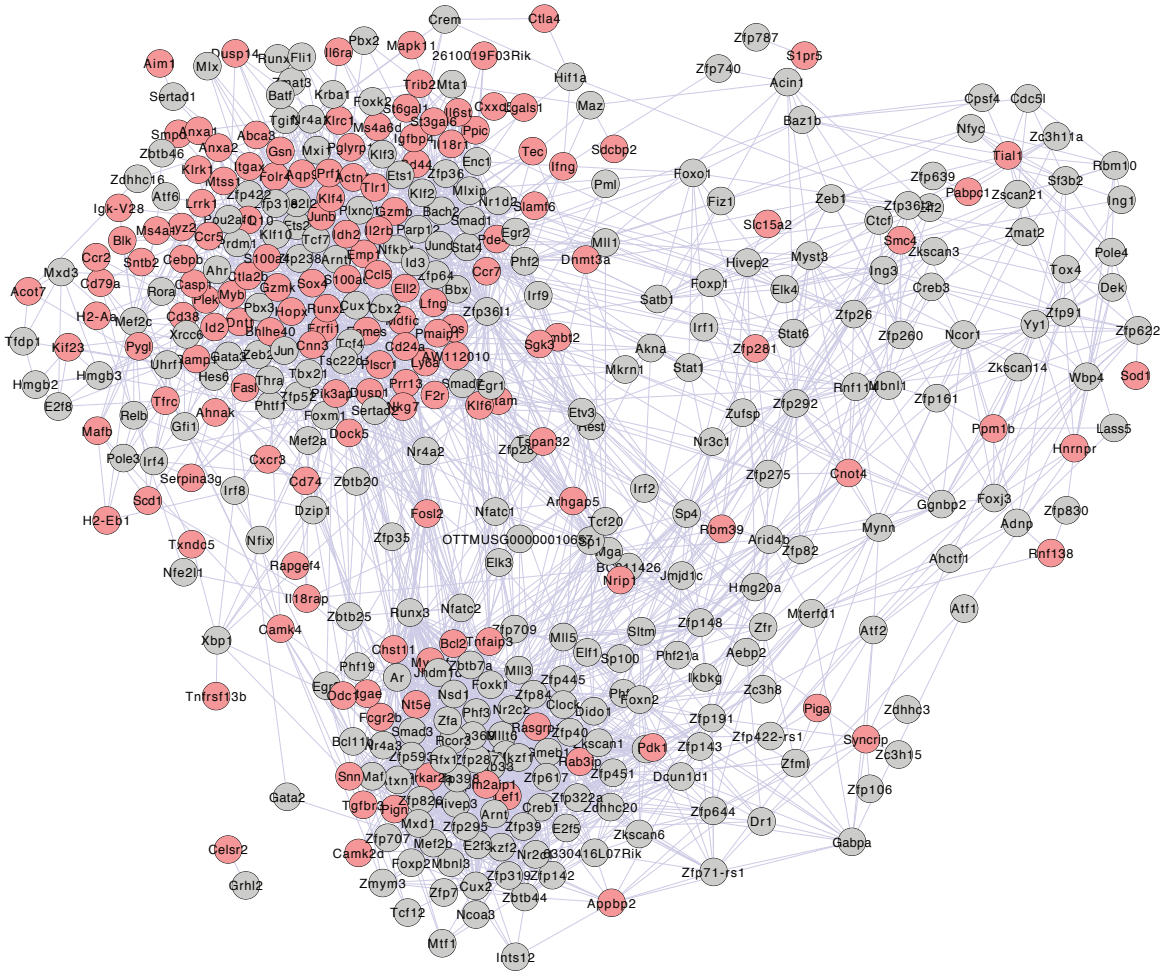


Hu and Chen, Figure 6

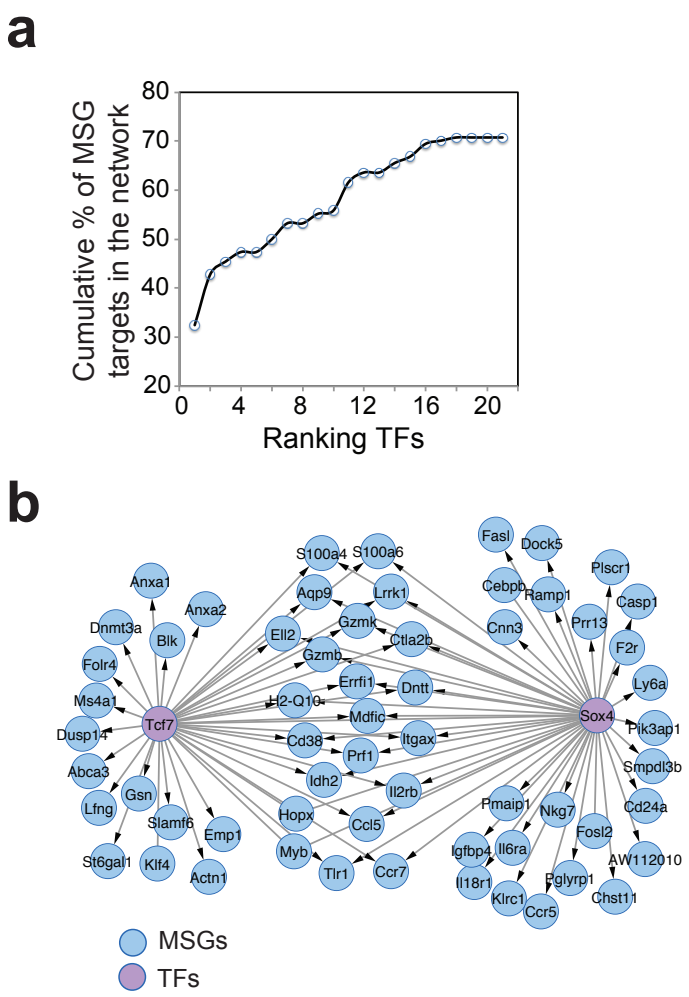


Hu and Chen, Figure 7

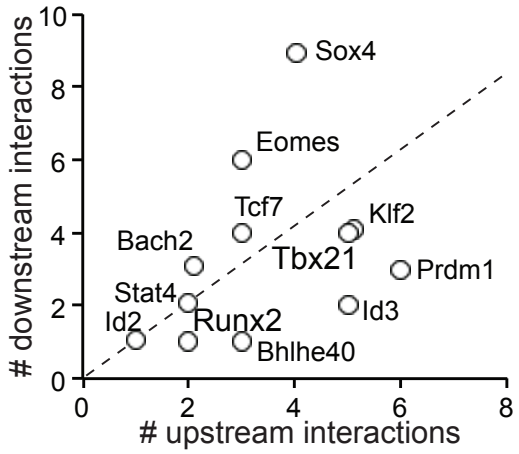
Supplementary information



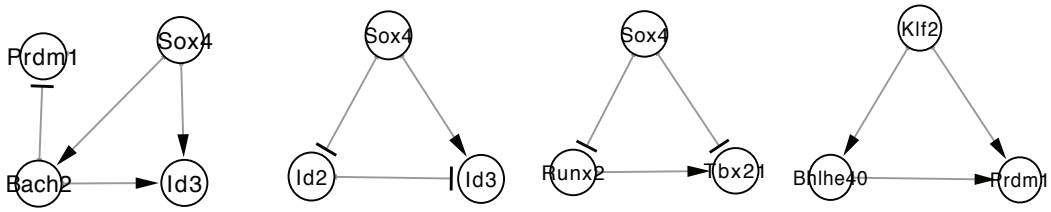
Supplementary Figure S1. The regulatory network of $CD8^+$ T cells showing 3,219 interactions involving 154 MSGs and 276 expressed putative TFs. The gray nodes indicate TFs and the red nodes MSGs.



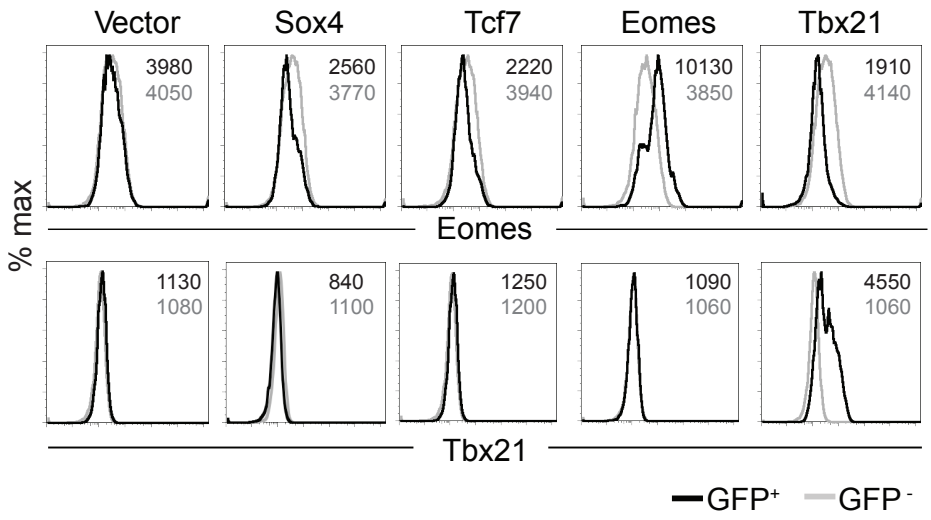
Supplementary Figure S2. (**a**) The cumulative percentage of MSGs that are regulated by the identified 21 TFs. The order that TFs were plotted is the same as they were ranked in Table 1. (**b**) The regulatory module of *Sox4* and *Tcf7* (orchid circles) and MSGs (blue circles).



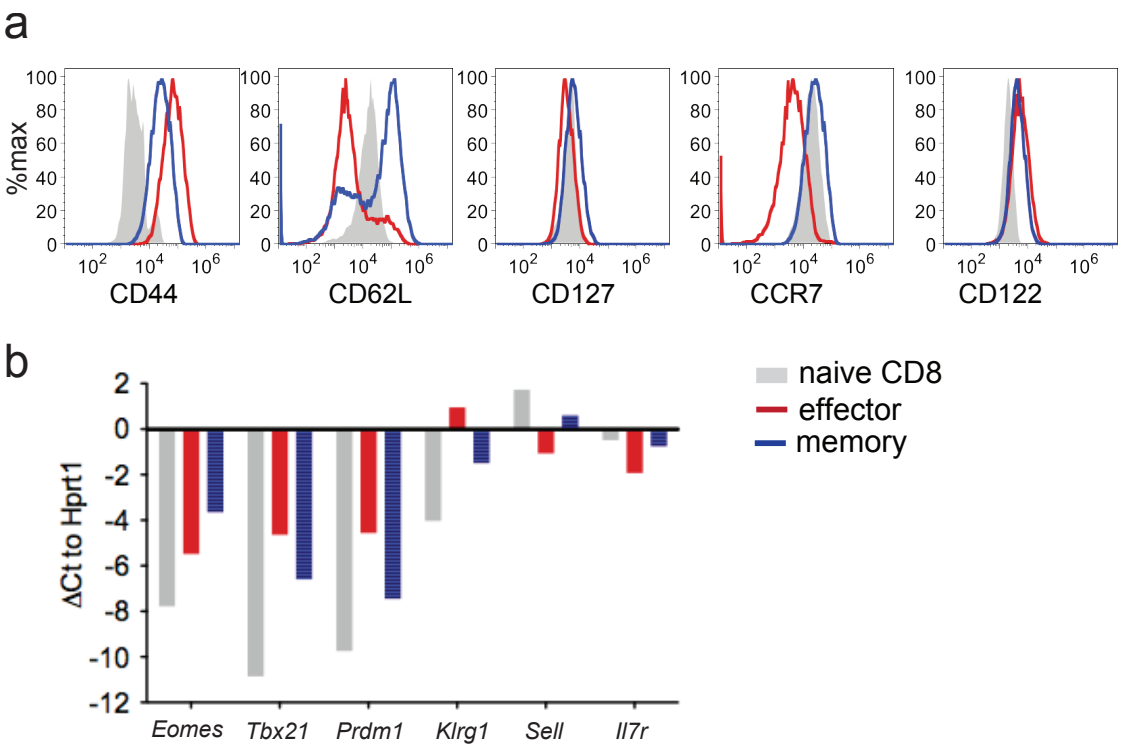
Supplementary Figure S3: Identification of TFs in the upstream of the perturbation network. Shown is the plotting of the number of downstream and upstream interacting TFs for each TF in the perturbation network.



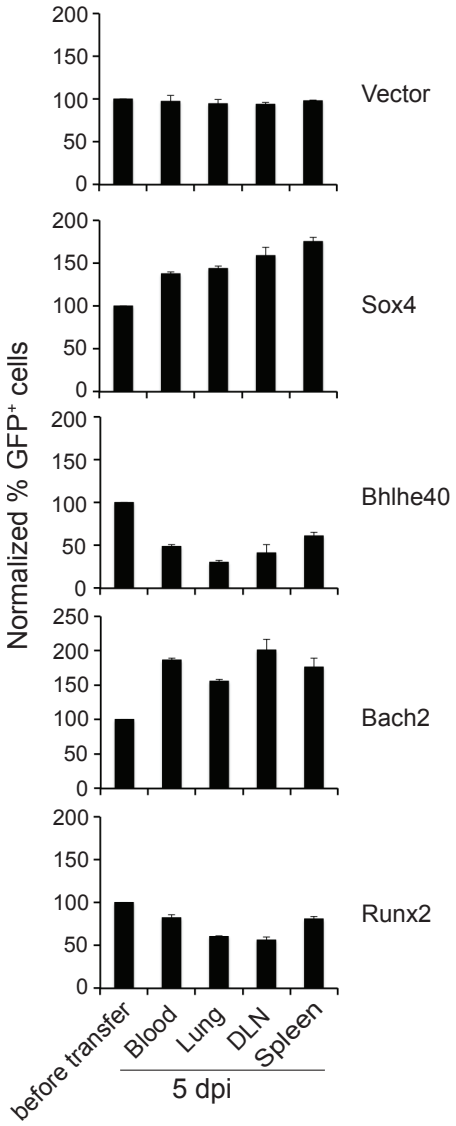
Supplementary Figure S4. Examples of network motifs from the perturbation network.



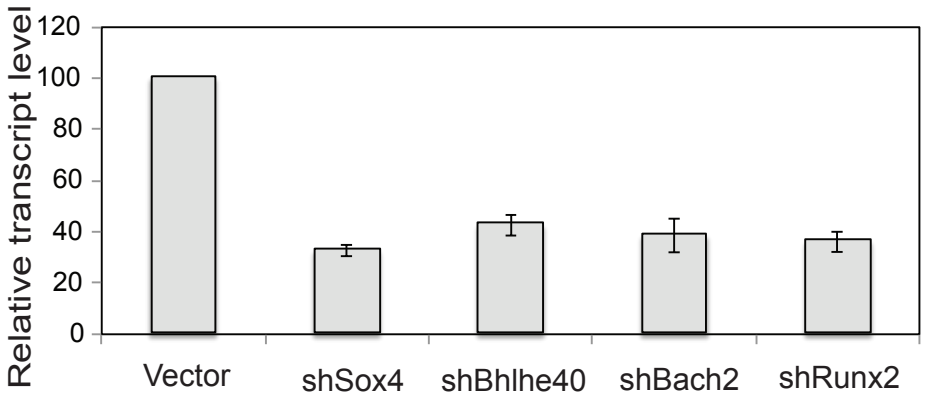
Supplementary Figure S5. Regulation of Eomes and Tbx21 by Sox4, Tcf7, Eomes and Tbx21. 2C T cells were transduced with retroviruses expressing GFP (vector) or GFP plus *Sox4*, *Tcf7*, *Eomes* or *Tbx21*. Expression of Eomes and Tbx21 was analyzed by intracellular staining and flow cytometry. Histograms compare Eomes and Tbx21 expression in non-transduced (GFP^-, gray lines) and transduced (GFP^+, black lines) T cells. The numbers indicate MFI.



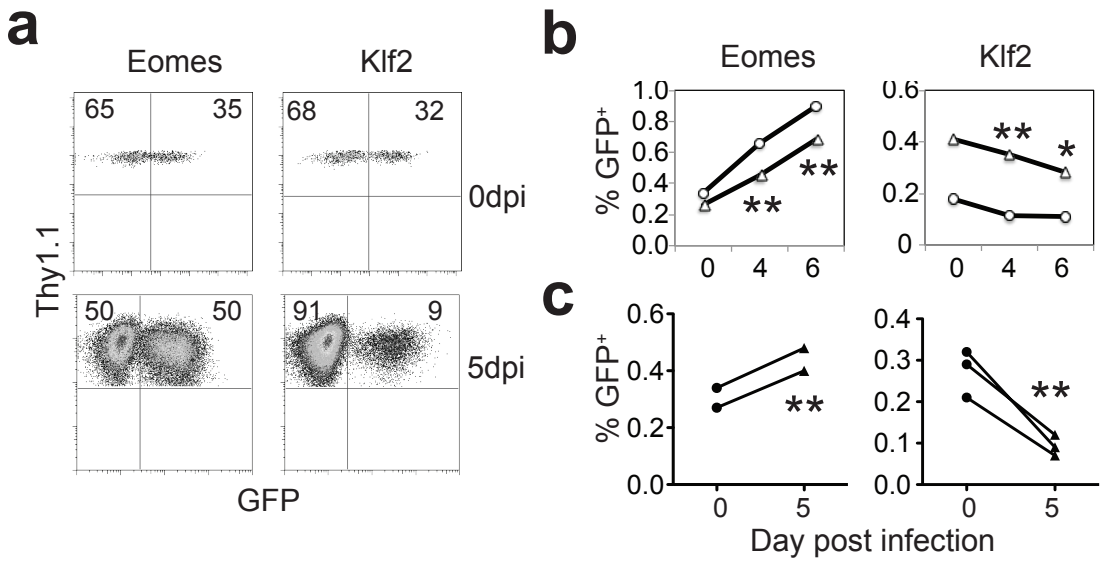
Supplementary Figure S6. Phenotype of *in vitro* memory T cells. Activated 2C T cells were cultured *in vitro* with IL-7 for 7 days to induce *in vitro* memory T cells. (**a**) Naïve, effector and *in vitro* memory 2C T cells were stained for CD44, CD62L, IL-7R, CCR7 and CD122 followed by cytometry. (**b**) Quantitative PCR analysis of gene expression of key memory genes (*Eomes*, *Tbx21*, *Prdm1*, *Il7r*, *Sell* and *Klrg1*) in naïve, activated and *in vitro* memory 2C T cells.



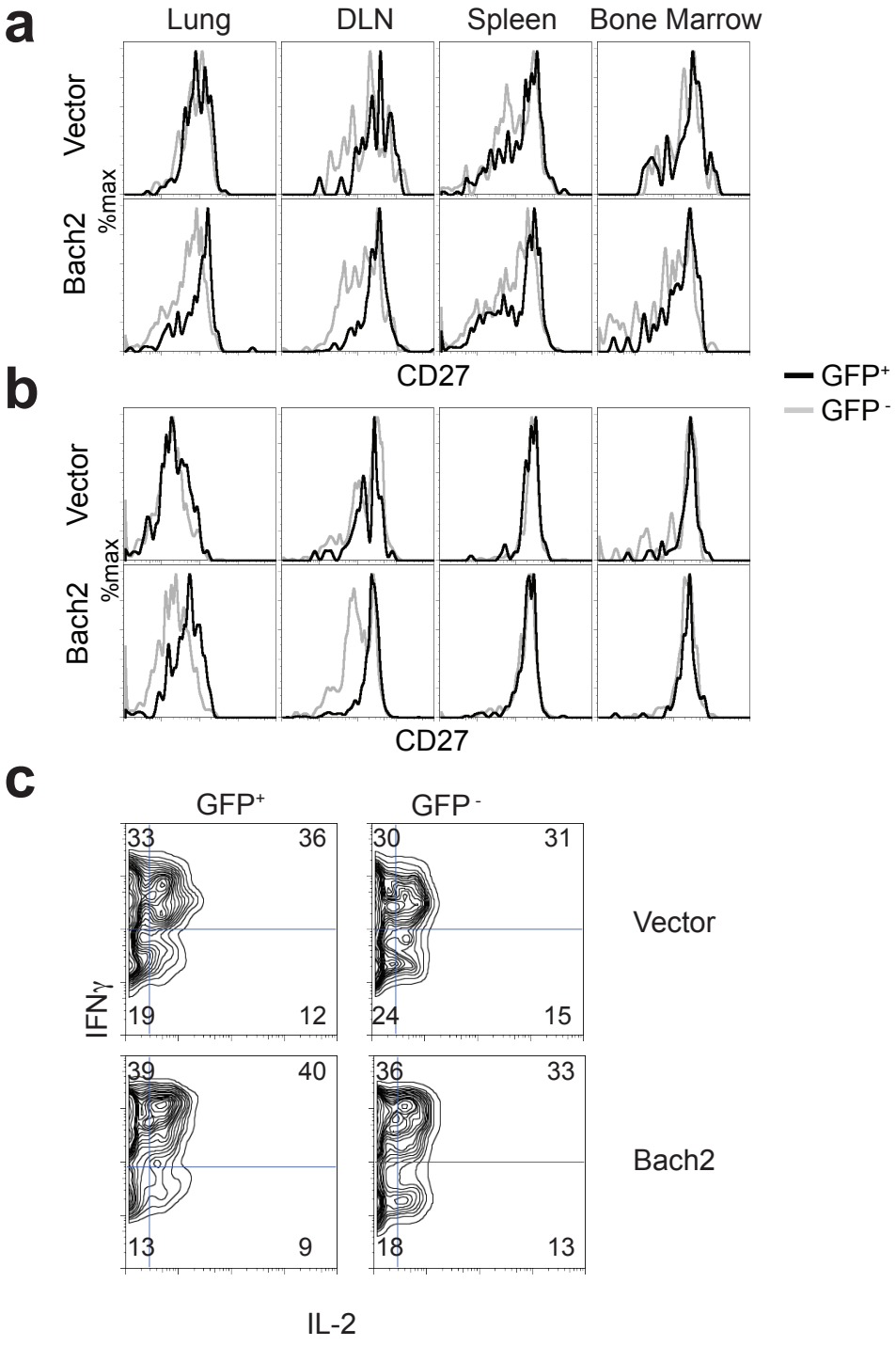
Supplementary Figure S7. Effect of overexpression of *Sox4*, *Bhlhe40*, *Bach2* and *Runx2* on recall proliferation of memory $CD8^+$ T cells *in vivo*. Naïve 2C T cells were activated *in vitro* with SIY peptide and then transduced with retroviruses expressing *Sox4*, *Bhlhe40*, *Bach2* or *Runx2*. The cells were cultured in the presence of IL-7 to derive memory T cells. The resulting memory 2C T cells were transferred into B6 mice and activated with WSN-SIY virus. The proportion of GFP^+ (transduced) versus GFP^- (non-transduced) 2C T cells was quantified in various organs 5 dpi. Shown is proportion of GFP^+ 2C T cells that overexpressed *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* among total 2C T cells *in vivo*. Data are from three independent experiments with 1-2 mice per group per experiment, shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses.



Supplementary Figure S8. Knowdown efficiency of shRNAs targeting *Sox4*, *Bhlhe40*, *Bach2* and *Runx2*. The indicated shRNA vectors were transduced individually in 2C T cells and the level of the target transcript was quantified by real-time PCR 48 hours later. Expression data was normalized to the empty vector control. Data from three independent experiments are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses.

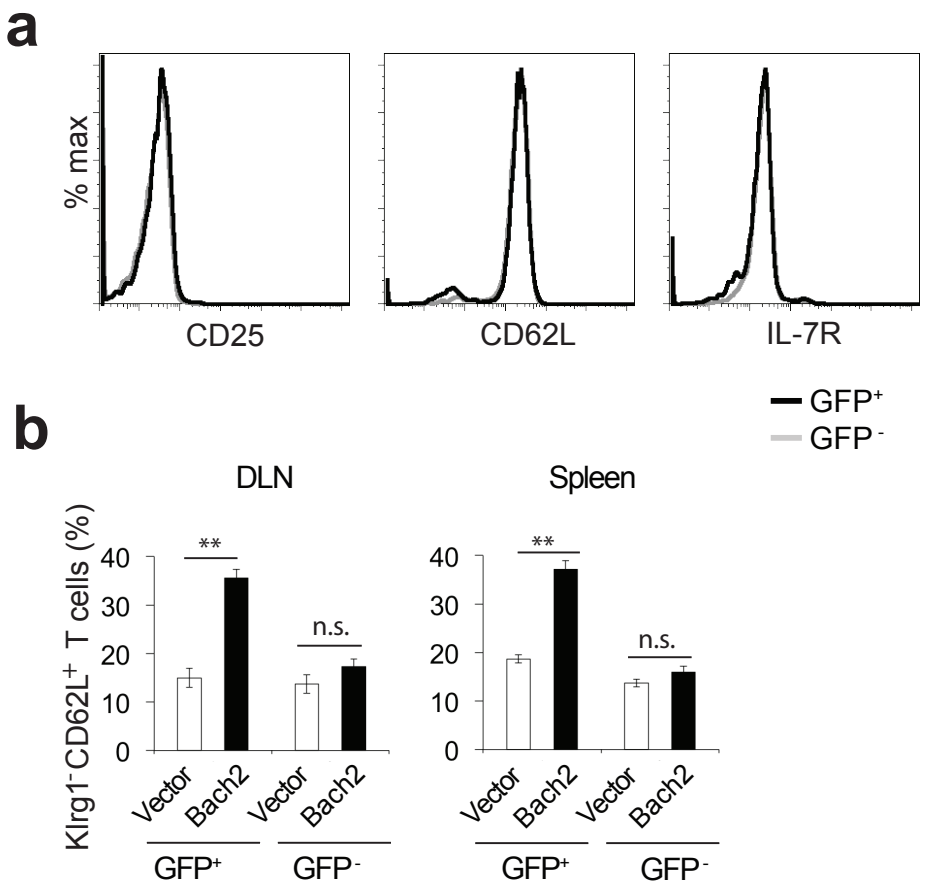


Supplementary Figure S9. Overexpression of *Eomes* and *Klf2* affects recall response of *in vitro* generated memory $CD8^+$ T cells. 2C T cells were activated and transduced with retroviruses expressing GFP plus *Eomes* or *Klf2*. The cells were cultured in the presence of IL-7 to derive memory T cells. The resulting memory 2C cells were either activated *in vitro* with SIY peptide or transferred into mice and activated with WSN-SIY virus. The proportion of GFP^+ (transduced) versus GFP^- (non-transduced) 2C T cells was quantified 4 and 6 days post stimulation *in vitro* and in DLN 5 dpi. Shown are Thy1.1 vs. GFP staining profiles of $CD8^+$ 2C T cells (**a**) and proportion of GFP^+ 2C T cells that overexpressed *Eomes* or *Klf2* among total 2C T cells *in vitro* (**b**) and *in vivo* (**c**). Each line was one independent experiment with one sample per time point for the *in vitro* experiments and one or two mice per *in vivo* experiment. * $P<0.05$; ** $P<0.01$. Pairwise two-tailed t-tests were used for statistical analyses.

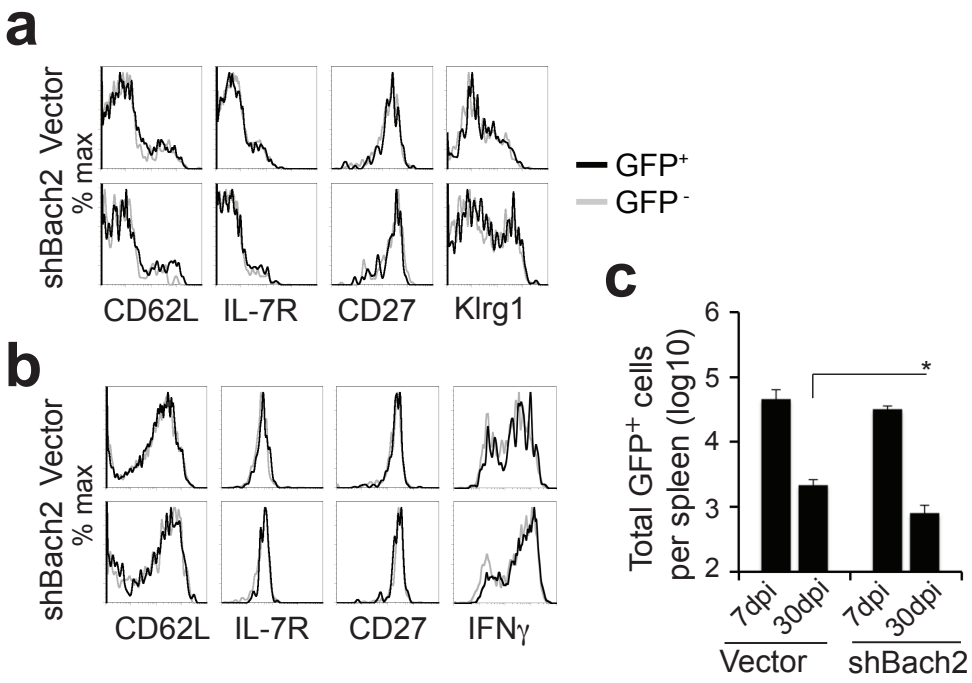


Supplementary Figure S10. *Bach2* promotes memory CD8$^+$ T cell development. Naïve 2C T cells were activated *in vitro* with SIY peptide and then transduced with retroviruses expressing *Bach2*. The cells were adoptively transferred into mice and activated with WSN-SIY virus. Phenotype of 2C T cells in various organs was analyzed for CD27.

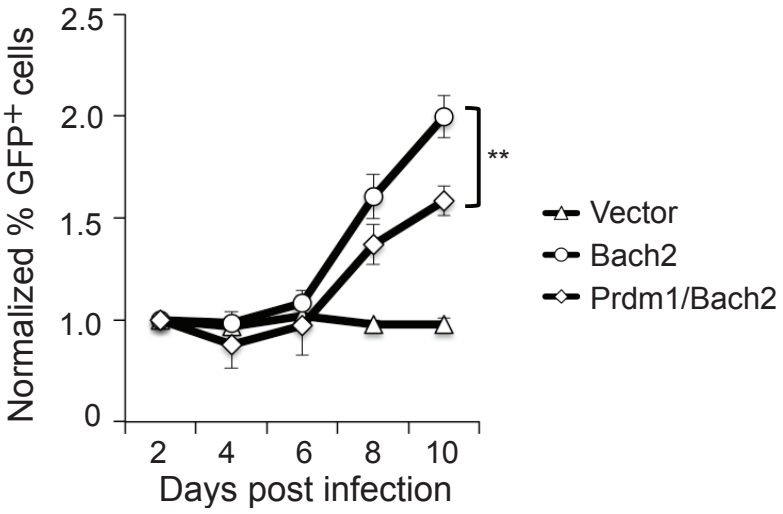
Shown are histograms CD27 expression of Thy1.1^{+} CD8^{+} 2C T cells 7 (**a**) and 30 dpi (**b**). Gray trace, nontransduced (GFP^{-}) 2C T cells; black trace, transduced (GFP^{+}) 2C T cells. IL-2 and IFNγ expression 2C T cells 30 dpi were analyzed as **Fig.4**. Shown is IL-2 vs. IFNγ staining profile of Thy1.1^{+} CD8^{+} 2C T cells (**c**). Representative data from 6 mice in 3 independent experiments are shown.



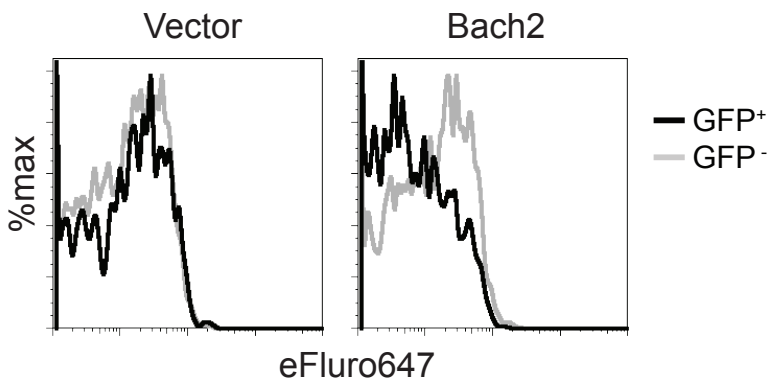
Supplementary Figure S11. *Bach2* promotes memory $CD8^+$ T cell development. Bone marrow progenitor cells isolated from 2C TCR transgenic mice were transduced with retroviruses expressing GFP alone (vector) or GFP plus *Bach2* and adoptively transferred into sublethally irradiated $RAG2^{-/-}$ mice to generate naïve 2C T cells that express GFP or GFP plus *Bach2*. Three to 4 months later, 2C T cells generated in $RAG2^{-/-}$ recipients were analyzed for Thy1.1, CD8 plus CD25, CD62L or IL-7R. (**a**) Comparison of CD25, CD62L and IL-7R expression between GFP^+ and GFP^- 2C T cells ($Thy1.1^+$ $CD8^+$). The resulting transduced and nontransduced naïve 2C cells were adoptively transferred into B6 mice followed by infection with WSN-SIY influenza virus. The frequency, phenotype and function of 2C T cells were analyzed 7 dpi. (**b)** Quantification of $CD62L^+Klrg1^-$ memory 2C T cell precursors in the DLN and spleen. Data from three independent experiments with 3-4 mice per group per experiment are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses. **P*<0.05, ***P*<0.01.



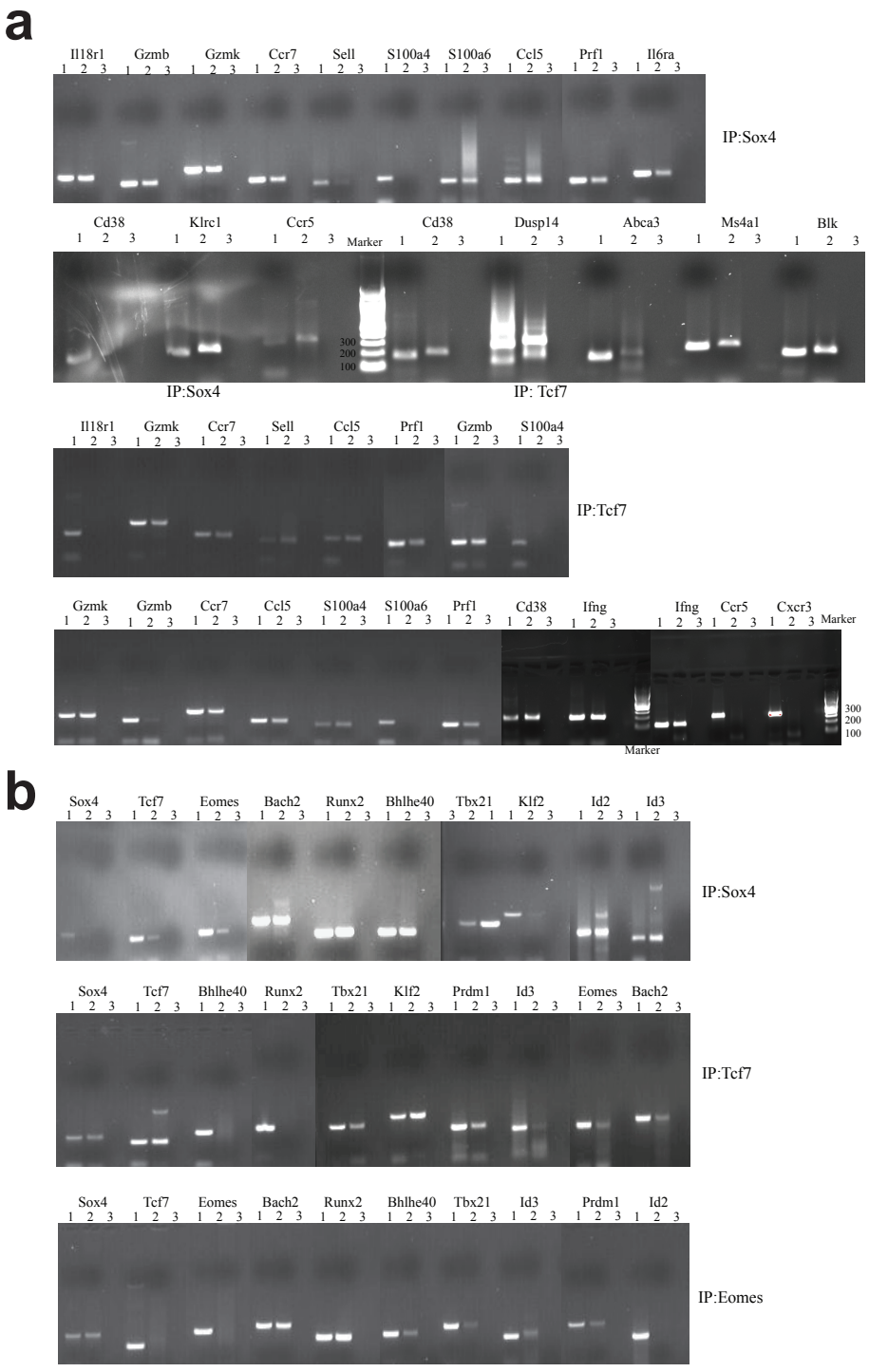
Supplementary Figure S12. *Bach2* is required for memory CD8$^+$ T cell development. Naïve 2C T cells were activated *in vitro* with SIY peptide and then transduced with retroviruses expressing the shRNA specific for *Bach2*. The cells were adoptively transferred into mice and activated with WSN-SIY virus. 2C T cells were analyzed 7 and 30 dpi as in **Fig. 4**. Shown are histograms of CD62L, IL-7R, CD27, Klrg1 and IFNγ expression of Thy1.1$^+$ CD8$^+$ 2C T cells in the spleen 7 (**a**) and 30 dpi (**b**). Gray trace, nontransduced (GFP$^-$) 2C T cells; black trace, transduced (GFP$^+$) 2C T cells. Representative data from 6-8 mice in two independent experiments are shown. (**c**) Total vector or shBach2 transduced 2C T cells (GFP$^+$) in the spleen 7 and 30 dpi. Representative data from 6-8 mice in two independent experiments are shown as mean ± s.e.m. Two-tailed student's t-tests were used for statistical analyses. **P*<0.05;



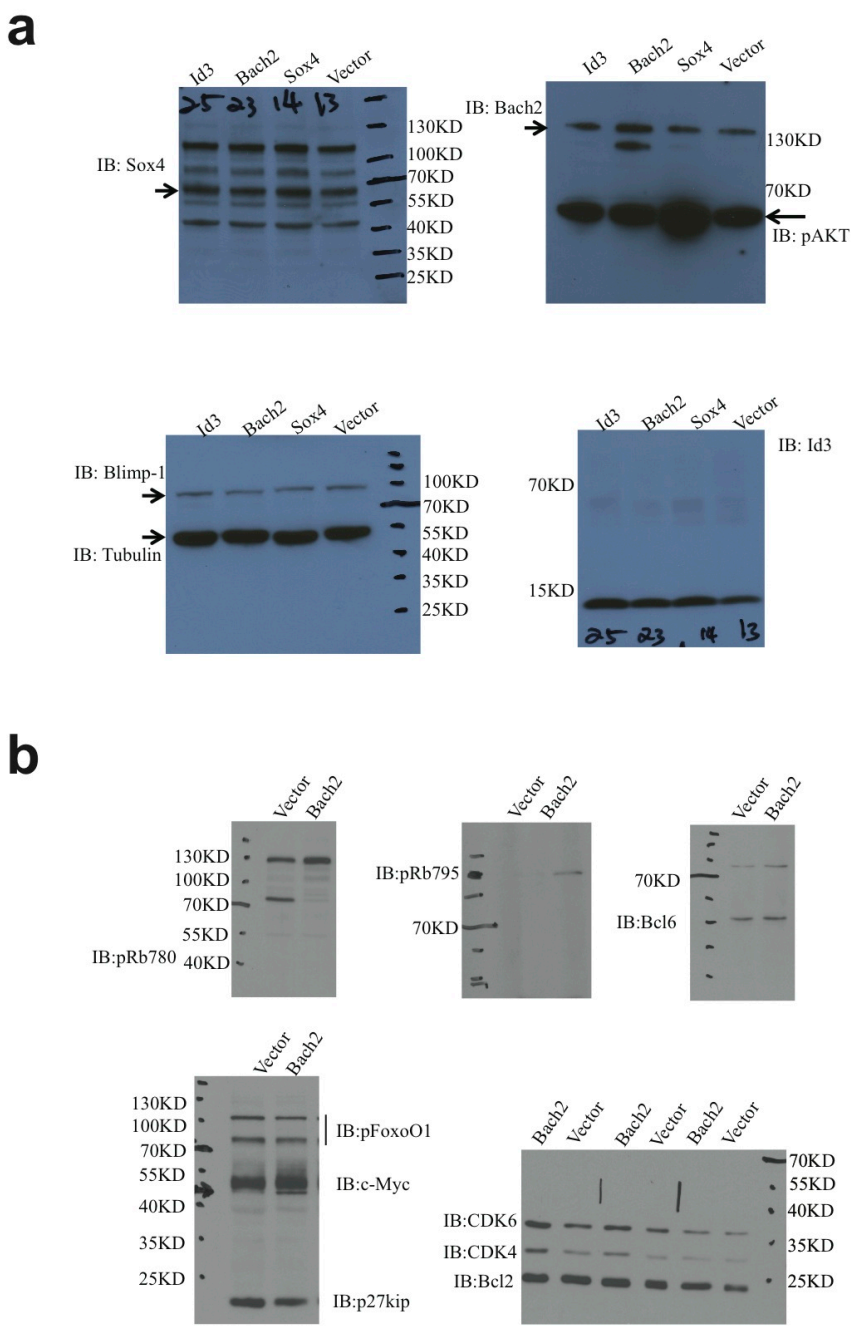
Supplementary Figure S13. *Prdm1* inhibits *Bach2*-mediated proliferation of $CD8^+$ T cells. 2C cells were activated *in vitro* and transduced with retroviruses expressing GFP alone (vector), GFP plus Bach2 or GFP plus Prdm1 and Bach2 (Prdm1-2A-Bach2) and cultured in the presence of IL-2. The proportion of GFP^+ 2C T cells was quantified. Shown are changes in proportion of GFP^+ cells over time from three independent experiments. Two-tailed student's t-tests were used for statistical analyses, error bar: s.e.m. ** $P<0.01$.



Supplementary Figure S14. Enhanced proliferation and survival of memory $CD8^+$ T cells. Naïve 2C cells were activated *in vitro* and transduced with retroviruses expressing GFP alone (vector) or GFP plus Bach2 and cultured in the presence of IL-7 to derive memory T cells. *In vitro* memory 2C T cells were labeled with eFluor®670 and adoptively transferred into $Rag2^{-/-}$ mice. 2C T cells in the spleen were analyzed 14 days later. Shown are representative histograms of eFluor®670 from one of the two experiments. Gray trace, nontransduced (GFP^-) 2C T cells; black trace, transduced (GFP^+) 2C T cells.



Supplementary Figure S15. The DNA gels of ChIP-PCR. (a) The DNA gels of ChIP-PCR for Fig. 1c. Representative data from 3 independent experiments. (b) The DNA gels of ChIP-PCR for Fig. 2a. Representative data from 3 independent experiments. 1, input chromatin; 2, IPed chromatin by TF Abs; 3, IPed chromatin by IgG.



Supplementary Figure S16. The Western blotting gels. (a) The protein gels of blotting Sox4, Bach2, Prdm1 and Id3 for Fig. 7d. Representative data from 3 independent experiments. (b) The protein gels of blotting pRb780, pRb795,pFoxO1, Bcl2,Bcl6, CDK4, CDK6 and p27kip for Fig. 7g. Marker is the PageRuler^{TM} Prestained Protein Ladder (Thermo Scientific). Representative data are from 3 independent experiments.

Supplementary Table S1. List of 386 gene expression profiles of CD8+ T cells used in the present study.

Index	Platform	GSE ID	ArrayID	Array Name	Title	Classification
1	GPL339	GSE2059	GSM35848	control 1	Control of naïve CD8+ T cells	N
2	GPL339	GSE2059	GSM37078	control 2	Control of naïve CD8+ T cells	N
3	GPL339	GSE2059	GSM37079	control 3	Control of naïve CD8+ T cells	N
4	GPL339	GSE2059	GSM37080	control 4	Control of naïve CD8+ T cells	N
5	GPL339	GSE2059	GSM37081	control 5	Control of naïve CD8+ T cells	N
6	GPL339	GSE2059	GSM37117	IL--15 1	IL15 of naïve CD8+ T cells	N
7	GPL339	GSE2059	GSM37118	IL--15 2	IL15 of naïve CD8+ T cells	N
8	GPL339	GSE2059	GSM37119	IL--15 3	IL15 of naïve CD8+ T cells	N
9	GPL339	GSE2059	GSM37120	IL--15 4	IL15 of naïve CD8+ T cells	N
10	GPL339	GSE2059	GSM37121	IL--15 5	IL15 of naïve CD8+ T cells	N
11	GPL339	GSE2059	GSM37122	IL--21 1	IL21 of naïve CD8+ T cells	N
12	GPL339	GSE2059	GSM37123	IL--21 2	IL21 of naïve CD8+ T cells	N
13	GPL339	GSE2059	GSM37124	IL--21 3	IL21 of naïve CD8+ T cells	N
14	GPL339	GSE2059	GSM37125	IL--21 4	IL21 of naïve CD8+ T cells	N
15	GPL339	GSE2059	GSM37126	IL--21 5	IL21 of naïve CD8+ T cells	N
16	GPL339	GSE2059	GSM37128	IL--15+IL-21 1	IL21 of naïve CD8+ T cells	N
17	GPL339	GSE2059	GSM37163	IL--15+IL-21 2	IL15+IL21 of naïve CD8+ T cells	N
18	GPL339	GSE2059	GSM37164	IL--15+IL-21 3	IL15+IL21 of naïve CD8+ T cells	N
19	GPL339	GSE2059	GSM37165	IL--15+IL-21 4	IL15+IL21 of naïve CD8+ T cells	N
20	GPL339	GSE2059	GSM37166	IL--15+IL-21 5	IL15+IL21 of naïve CD8+ T cells	N
21	GPL339	GSE4938	GSM111271	wt CD8 M	CD8 M of Gfi1-WT	N
22	GPL339	GSE4938	GSM111269	ko CD8 N	CD8 N of Gfi1-KO	N
23	GPL339	GSE4940	GSM111286	ko 501_0h	0hr T Gfi1-KO activated by anti-CD3&CD28	N
24	GPL339	GSE4940	GSM111287	ko 501_4h	4hr T Gfi1-KO activated by anti-CD3&CD28	E
25	GPL339	GSE4940	GSM111288	ko 501_8h	8hr T Gfi1-KO activated by anti-CD3&CD28	E
26	GPL339	GSE4940	GSM111289	ko 501_24h	24hr T Gfi1-KO activated by anti-CD3&CD28	E
27	GPL339	GSE4940	GSM111290	wt 502_0h	0hr T Gfi1-WT activated by anti-CD3&CD28	E
28	GPL339	GSE4940	GSM111291	wt 502_4h	4hr T Gfi1-WT activated by anti-CD3&CD28	E
29	GPL339	GSE4940	GSM111292	wt 502_8h	8hr T Gfi1-WT activated by anti-CD3&CD28	E
30	GPL339	GSE4940	GSM111293	wt 502_24h	24hr T Gfi1-WT activated by anti-CD3&CD28	E
31	GPL339	GSE10093	GSM254977	Spleen CD8 T cells DMSO1	Spleen CD8 T cells with DMSO	N
32	GPL339	GSE10093	GSM254983	Spleen CD8 T cells DMSO2	Spleen CD8 T cells with DMSO	N
33	GPL339	GSE10093	GSM254976	Spleen CD8 T cells TCDD1	Spleen CD8 T cells with TCDD	
34	GPL339	GSE10093	GSM254978	Spleen CD8 T cells TCDD2	Spleen CD8 T cells with TCDD	
35	GPL339	GSE3565	GSM81921	DUSP1-/- control replicate 1	DUSP1-/- control	
36	GPL339	GSE3565	GSM81922	DUSP1-/- control replicate 2	DUSP1-/- control	
37	GPL339	GSE3565	GSM81923	DUSP1-/- control replicate 3	DUSP1-/- control	
38	GPL339	GSE3565	GSM81924	DUSP1-/- LPS replicate 1	DUSP1-/- LPS	
39	GPL339	GSE3565	GSM81925	DUSP1-/- LPS replicate 2	DUSP1-/- LPS	
40	GPL339	GSE3565	GSM81926	DUSP1-/- LPS replicate 3	DUSP1-/- LPS	
41	GPL339	GSE3565	GSM81927	DUSP1+/+ control replicate 1	DUSP1+/+ control	
42	GPL339	GSE3565	GSM81928	DUSP1+/+ control replicate 2	DUSP1+/+ control	
43	GPL339	GSE3565	GSM81929	DUSP1+/+ control replicate 3	DUSP1+/+ control	
44	GPL339	GSE3565	GSM81930	DUSP1+/+ LPS replicate 1	DUSP1+/+ LPS	
45	GPL339	GSE3565	GSM81931	DUSP1+/+ LPS replicate 2	DUSP1+/+ LPS	
46	GPL339	GSE3565	GSM81932	DUSP1+/+ LPS replicate 3	DUSP1+/+ LPS	
47	GPL339	GSE5811	GSM135705	B1_saline_overexpressor_422	saline_overexpressor_422	
48	GPL339	GSE5811	GSM135706	B2_saline_overexpressor_422	saline_overexpressor_422	
49	GPL339	GSE5811	GSM135707	B3_saline_overexpressor_422	saline_overexpressor_422	
50	GPL339	GSE5811	GSM135708	B4_pseudomonas_overexpressor_422	pseudomonas_overexpressor_422	
51	GPL339	GSE5811	GSM135709	B5_pseudomonas_overexpressor_422	pseudomonas_overexpressor_422	
52	GPL339	GSE5811	GSM135710	B6_pseudomonas_overexpressor_422	pseudomonas_overexpressor_422	
53	GPL339	GSE5811	GSM135711	R1_saline_wildtype_422	saline_wildtype_422	
54	GPL339	GSE5811	GSM135712	R2_saline_wildtype_422	saline_wildtype_422	
55	GPL339	GSE5811	GSM135713	R4_saline_wildtype_422	saline_wildtype_422	
56	GPL339	GSE5811	GSM135714	R5_Pseudomonas_wildtype_422	Pseudomonas_wildtype_422	
57	GPL339	GSE5811	GSM135715	R6_Pseudomonas_wildtype_422	Pseudomonas_wildtype_422	
58	GPL339	GSE5811	GSM135716	R7_Pseudomonas_wildtype_422	Pseudomonas_wildtype_422	
59	GPL339	GSE5811	GSM135717	B2_saline_overexpressor_48	saline_overexpressor_48	
60	GPL339	GSE5811	GSM135718	B3_saline_overexpressor_48	saline_overexpressor_48	
61	GPL339	GSE5811	GSM135719	B5_Pseudomonas_overexpressor_48	Pseudomonas_overexpressor_48	
62	GPL339	GSE5811	GSM135720	B6_Pseudomonas_overexpressor_48	Pseudomonas_overexpressor_48	
63	GPL339	GSE5811	GSM135721	R2_saline_wildtype_48	saline_wildtype_48	
64	GPL339	GSE5811	GSM135722	R3_saline_wildtype_48	saline_wildtype_48	
65	GPL339	GSE5811	GSM135723	R5_Pseudomonas_wildtype_48	Pseudomonas_wildtype_48	
66	GPL339	GSE5811	GSM135724	R6_Pseudomonas_wildtype_48	Pseudomonas_wildtype_48	
67	GPL339	GSE5811	GSM135725	C1_CLP_overexpressor_12	CLP_overexpressor_12	
68	GPL339	GSE5811	GSM135726	C2_CLP_overexpressor_12	CLP_overexpressor_12	
69	GPL339	GSE5811	GSM135727	C3_CLP_overexpressor_12	CLP_overexpressor_12	
70	GPL339	GSE5811	GSM135728	C4_CLP_overexpressor_12	CLP_overexpressor_12	
71	GPL339	GSE5811	GSM135729	C5_CLP_overexpressor_12	CLP_overexpressor_12	
72	GPL339	GSE5811	GSM135730	C6_CLP_wildtype_12	CLP_wildtype_12	
73	GPL339	GSE5811	GSM135731	S1_sham_overexpressor_12	sham_overexpressor_12	
74	GPL339	GSE5811	GSM135732	S2_sham_overexpressor_12	sham_overexpressor_12	
75	GPL339	GSE5811	GSM135733	S3_sham_overexpressor_12	sham_overexpressor_12	
76	GPL339	GSE5811	GSM135734	S4_sham_overexpressor_12	sham_overexpressor_12	
77	GPL339	GSE5811	GSM135735	S5_sham_overexpressor_12	sham_overexpressor_12	
78	GPL339	GSE5811	GSM135736	S6_Sham_wildtype_12	Sham_wildtype_12	
79	GPL339	GSE5811	GSM135737	C1_CLP_wildtype_8	CLP_wildtype_8	
80	GPL339	GSE5811	GSM135738	C3_CLP_wildtype_8	CLP_wildtype_8	
81	GPL339	GSE5811	GSM135739	C4_CLP_wildtype_8	CLP_wildtype_8	
82	GPL339	GSE5811	GSM135740	C5star_CLP_wildtype_8	CLP_wildtype_8	
83	GPL339	GSE5811	GSM135741	C6star_CLP_wildtype_8	CLP_wildtype_8	
84	GPL339	GSE5811	GSM135742	S2_Sham_wildtype_8	Sham_wildtype_8	
85	GPL339	GSE5811	GSM135743	S3star_Sham_wildtype_8	Sham_wildtype_8	
86	GPL339	GSE5811	GSM135744	S4_Sham_wildtype_8	Sham_wildtype_8	
87	GPL339	GSE5811	GSM135745	S6_Sham_wildtype_8	Sham_wildtype_8	
88	GPL339	GSE7768	GSM188270	Ova, biological replicate 1	Ova	
89	GPL339	GSE7768	GSM188271	Ova, biological replicate 2	Ova	
90	GPL339	GSE7768	GSM188272	Ova, biological replicate 3	Ova	
91	GPL339	GSE7768	GSM188273	OvaLPS, biological replicate 1	OvaLPS	
92	GPL339	GSE7768	GSM188274	OvaLPS, biological replicate 2	OvaLPS	
93	GPL339	GSE7768	GSM188275	OvaLPS, biological replicate 3	OvaLPS	
94	GPL339	GSE7768	GSM188276	OvaMPL, biological replicate 1	OvaMP	
95	GPL339	GSE7768	GSM188277	OvaMPL, biological replicate 2	OvaMP	

96	GPL339	GSE7768	GSM188278	OvaMPL, biological replicate 3	OvaMP
97	GPL339	GSE3997	GSM91280	B10 TCR double positive rep 1	CD4+CD8+TCRhi
98	GPL339	GSE3997	GSM91281	B10 TCR double positive rep 2	CD4+CD8+TCRhi
99	GPL339	GSE3997	GSM91282	B10 TCR double positive rep 3	CD4+CD8+TCRhi
100	GPL339	GSE3997	GSM91283	NOD TCR double positive rep 1	CD4+CD8+TCRhi
101	GPL339	GSE3997	GSM91284	NOD TCR double positive rep 2	CD4+CD8+TCRhi
102	GPL339	GSE3997	GSM91285	NOD TCR double positive rep 3	CD4+CD8+TCRhi
103	GPL339	GSE3997	GSM91286	B10 Dbl double positive rep 1	CD4+CD8+CD69-TCRlow
104	GPL339	GSE3997	GSM91287	B10 Dbl double positive rep 2	CD4+CD8+CD69-TCRlow
105	GPL339	GSE3997	GSM91288	B10 Dbl double positive rep 3	CD4+CD8+CD69-TCRlow
106	GPL339	GSE3997	GSM91289	NOD Dbl double positive rep 1	CD4+CD8+CD69-TCRlow
107	GPL339	GSE3997	GSM91290	NOD Dbl double positive rep 2	CD4+CD8+CD69-TCRlow
108	GPL339	GSE3997	GSM91291	NOD Dbl double positive rep 3	CD4+CD8+CD69-TCRlow
109	GPL339	GSE2413	GSM45451	0h WT Sample 01	S49 T-lymphoma cell line
110	GPL339	GSE2413	GSM45452	0h WT Sample 13	S49 T-lymphoma cell line
111	GPL339	GSE2413	GSM45453	0h WT Sample 16	S49 T-lymphoma cell line
112	GPL339	GSE2413	GSM45454	0h WT Sample 36	S49 T-lymphoma cell line
113	GPL339	GSE2413	GSM45455	2h WT Cel.02	S49 T-lymphoma cell line
114	GPL339	GSE2413	GSM45456	2h WT Cel.17	S49 T-lymphoma cell line
115	GPL339	GSE2413	GSM45457	2h WT Cel.37	S49 T-lymphoma cell line
116	GPL339	GSE2413	GSM45458	6h WT Cel.03	S49 T-lymphoma cell line
117	GPL339	GSE2413	GSM45459	6h WT Cel.18	S49 T-lymphoma cell line
118	GPL339	GSE2413	GSM45460	6h WT cel.38	S49 T-lymphoma cell line
119	GPL339	GSE2413	GSM45461	24h WT Cel.04	S49 T-lymphoma cell line
120	GPL339	GSE2413	GSM45462	24h WT Cel.19	S49 T-lymphoma cell line
121	GPL339	GSE2413	GSM45463	24h WT Cel.39	S49 T-lymphoma cell line
122	GPL339	GSE2413	GSM45464	0h Kin Cel.05	S49 T-lymphoma cell line
123	GPL339	GSE2413	GSM45465	0h Kin Cel.14	S49 T-lymphoma cell line
124	GPL339	GSE2413	GSM45466	0h Kin Cel.40	S49 T-lymphoma cell line
125	GPL339	GSE2413	GSM45467	0h Kin Cel.44	S49 T-lymphoma cell line
126	GPL339	GSE2413	GSM45468	2h Kin Cel.06	S49 T-lymphoma cell line
127	GPL339	GSE2413	GSM45469	2h Kin- Cel.41	S49 T-lymphoma cell line
128	GPL339	GSE2413	GSM45470	2h Kin- Cel.45	S49 T-lymphoma cell line
129	GPL339	GSE2413	GSM45471	6h kin- Cel.07	S49 T-lymphoma cell line
130	GPL339	GSE2413	GSM45472	6h Kin- Cel.42	S49 T-lymphoma cell line
131	GPL339	GSE2413	GSM45473	6h Kin- Cel.46	S49 T-lymphoma cell line
132	GPL339	GSE2413	GSM45474	24h kin- Cel.08	S49 T-lymphoma cell line
133	GPL339	GSE2413	GSM45475	24h Kin- Cel.43	S49 T-lymphoma cell line
134	GPL339	GSE2413	GSM45476	24h Kin Cel.47	S49 T-lymphoma cell line
135	GPL339	GSE9727	GSM245893	S49_untreated_time0,rep1	S49 T-lymphoma cell line
136	GPL339	GSE9727	GSM245894	S49_untreated_time0,rep2	S49 T-lymphoma cell line
137	GPL339	GSE9727	GSM245895	S49_untreated_time0,rep3	S49 T-lymphoma cell line
138	GPL339	GSE9727	GSM245896	S49_untreated_time0,rep4	S49 T-lymphoma cell line
139	GPL339	GSE9727	GSM245897	S49_treated_2h,rep1	S49 T-lymphoma cell line
140	GPL339	GSE9727	GSM245898	S49_treated_2h,rep2	S49 T-lymphoma cell line
141	GPL339	GSE9727	GSM245899	S49_treated_2h,rep3	S49 T-lymphoma cell line
142	GPL339	GSE9727	GSM245900	S49_treated_6h,rep1	S49 T-lymphoma cell line
143	GPL339	GSE9727	GSM245901	S49_treated_6h,rep2	S49 T-lymphoma cell line
144	GPL339	GSE9727	GSM245902	S49_treated_6h,rep3	S49 T-lymphoma cell line
145	GPL339	GSE9727	GSM245903	S49_treated_24h,rep1	S49 T-lymphoma cell line
146	GPL339	GSE9727	GSM245904	S49_treated_24h,rep2	S49 T-lymphoma cell line
147	GPL339	GSE9727	GSM245905	S49_treated_24h,rep3	S49 T-lymphoma cell line
148	GPL339	GSE9727	GSM248218	S49_untreated_time0,rep1	S49 T-lymphoma cell line
149	GPL339	GSE9727	GSM248219	S49_untreated_time0,rep2	S49 T-lymphoma cell line
150	GPL339	GSE9727	GSM248220	S49_untreated_time0,rep3	S49 T-lymphoma cell line
151	GPL339	GSE9727	GSM248221	S49_untreated_time0,rep4	S49 T-lymphoma cell line
152	GPL339	GSE9727	GSM248222	S49_untreated_2h,rep1	S49 T-lymphoma cell line
153	GPL339	GSE9727	GSM248223	S49_untreated_2h,rep2	S49 T-lymphoma cell line
154	GPL339	GSE9727	GSM248224	S49_untreated_2h,rep3	S49 T-lymphoma cell line
155	GPL339	GSE9727	GSM248225	S49_untreated_6h,rep1	S49 T-lymphoma cell line
156	GPL339	GSE9727	GSM248226	S49_untreated_6h,rep2	S49 T-lymphoma cell line
157	GPL339	GSE9727	GSM248227	S49_untreated_6h,rep3	S49 T-lymphoma cell line
158	GPL339	GSE9727	GSM248228	S49_untreated_24h,rep1	S49 T-lymphoma cell line
159	GPL339	GSE9727	GSM248229	S49_untreated_24h,rep2	S49 T-lymphoma cell line
160	GPL339	GSE9727	GSM248230	S49_untreated_24h,rep3	S49 T-lymphoma cell line
161	GPL1261	GSE15324	GSM384857	WT1 naive CD8 T cells	WT naive CD8 T cells	N
162	GPL1261	GSE15324	GSM384858	WT2 naive CD8 T cells	WT naive CD8 T cells	N
163	GPL1261	GSE15324	GSM384859	WT1 activated CD8 T cells	WT activated CD8 T cells	E
164	GPL1261	GSE15324	GSM384866	WT2 activated CD8 T cells	WT activated CD8 T cells	E
165	GPL1261	GSE15324	GSM384867	Elf4 KO1 naive CD8 T cells	Elf4 KO naive CD8 T cells	N
166	GPL1261	GSE15324	GSM384868	Elf4 KO2 naive CD8 T cells	Elf4 KO naive CD8 T cells	N
167	GPL1261	GSE15324	GSM384869	Elf4 KO1 activated CD8 T cells	Elf4 KO activated CD8 T cells	E
168	GPL1261	GSE15324	GSM384906	Elf4 KO2 activated CD8 T cells	Elf4 KO activated CD8 T cells	E
169	GPL1261	GSE6084	GSM140919	T-cell_Cycloheximide_30min_rep1	T cell_Cycloheximide_30min	
170	GPL1261	GSE6084	GSM140920	T-cell_Cycloheximide_30min_rep2	T cell_Cycloheximide_30min	
171	GPL1261	GSE6084	GSM140921	T-cell_Cycloheximide_30min_rep3	T cell_Cycloheximide_30min	
172	GPL1261	GSE6084	GSM140922	T-cell_Cycloheximide_30min_rep4	T cell_Cycloheximide_30min	
173	GPL1261	GSE6084	GSM108343	T-cell_IL-2_Cycloheximide_4 hour_rep1	T cell_IL-2_Cycloheximide_4 hour	
174	GPL1261	GSE6084	GSM108344	T-cell_IL-2_Cycloheximide_4 hour_rep2	T cell_IL-2_Cycloheximide_4 hour	
175	GPL1261	GSE6084	GSM108345	T-cell_IL-2_Cycloheximide_4 hour_rep3	T cell_IL-2_Cycloheximide_4 hour	
176	GPL1261	GSE6084	GSM108346	T-cell_IL-2_Cycloheximide_4 hour_rep4	T cell_IL-2_Cycloheximide_4 hour	
177	GPL1261	GSE6084	GSM183635	T-cell_Cycloheximide_4.5h_rep1	T cell_Cycloheximide_4.5h	
178	GPL1261	GSE6084	GSM183636	T-cell_Cycloheximide_4.5h_rep2	T cell_Cycloheximide_4.5h	
179	GPL1261	GSE6084	GSM183637	T-cell_Cycloheximide_4.5h_rep3	T cell_Cycloheximide_4.5h	
180	GPL1261	GSE6084	GSM183638	T-cell_Cycloheximide_4.5h_rep4	T cell_Cycloheximide_4.5h	
181	GPL1261	GSE11677	GSM296650	CD8 memory T cell_Expansion_rep1	CD8 memory T cell_Expansion	M
182	GPL1261	GSE11677	GSM296652	CD8 memory T cell_Expansion_rep2	CD8 memory T cell_Expansion	M
183	GPL1261	GSE11677	GSM296654	CD8 memory T cell_Expansion_rep3	CD8 memory T cell_Expansion	M
184	GPL1261	GSE11677	GSM296656	CD8 memory T cell_Expansion_rep4	CD8 memory T cell_Expansion	M
185	GPL1261	GSE11677	GSM296651	CD8 memory T cell_PolyclonalAged_rep1	CD8 memory T cell_PolyclonalAged	M
186	GPL1261	GSE11677	GSM296653	CD8 memory T cell_PolyclonalAged_rep2	CD8 memory T cell_PolyclonalAged	M
187	GPL1261	GSE11677	GSM296655	CD8 memory T cell_PolyclonalAged_rep3	CD8 memory T cell_PolyclonalAged	M
188	GPL1261	GSE11677	GSM296657	CD8 memory T cell_PolyclonalAged_rep4	CD8 memory T cell_PolyclonalAged	M
189	GPL1261	GSE12388	GSM310897	Spleen Aire-wild-type A	Spleen Aire-wild-type	
190	GPL1261	GSE12388	GSM310900	Spleen Aire-wild-type B	Spleen Aire-wild-type	
191	GPL1261	GSE12388	GSM310898	Spleen Aire-knockout A	Spleen Aire-knockout	
192	GPL1261	GSE12388	GSM310901	Spleen Aire-knockout B	Spleen Aire-knockout	
193	GPL1261	GSE12388	GSM310903	Lymph node Aire-wild-type A	Lymph node Aire-wild-type	

194	GPL1261	GSE12388	GSM310899	Lymph node Aire-wild-type B	Lymph node Aire-wild-type
195	GPL1261	GSE12388	GSM310904	Lymph node Aire-knockout A	Lymph node Aire-knockout
196	GPL1261	GSE12388	GSM310902	Lymph node Aire-knockout B	Lymph node Aire-knockout
197	GPL1261	GSE11884	GSM300153	Naive_Furin_Wild-type_1	Naive_Furin_Wild-type_1
198	GPL1261	GSE11884	GSM300154	Naive_Furin_Wild-type_2	Naive_Furin_Wild-type_1
199	GPL1261	GSE11884	GSM300155	Naive_Furin_Knockout_1	Naive_Furin_Knockout_1
200	GPL1261	GSE11884	GSM300156	Naive_Furin_Knockout_2	Naive_Furin_Knockout_1
201	GPL1261	GSE11446	GSM288809	CD8 T cells_IL-2 complex treatment_0 h	CD8 T cells_IL-2 complex treatment_0 h N
202	GPL1261	GSE11446	GSM288810	CD8 T cells_IL-2 complex treatment_1 h	CD8 T cells_IL-2 complex treatment_1 h N
203	GPL1261	GSE11446	GSM288811	CD8 T cells_IL-2 complex treatment_3 h	CD8 T cells_IL-2 complex treatment_3 h N
204	GPL1261	GSE10239	GSM257826	CD8 P14 Naive cells-1	CD8 P14 Naive cells N
205	GPL1261	GSE10239	GSM257827	CD8 P14 Naive cells-2	CD8 P14 Naive cells N
206	GPL1261	GSE10239	GSM257828	CD8 P14 Naive cells-3	CD8 P14 Naive cells N
207	GPL1261	GSE10239	GSM257829	CD8 P14 Memory cells-1	CD8 P14 Memory cells M
208	GPL1261	GSE10239	GSM257830	CD8 P14 Memory cells-2	CD8 P14 Memory cells M
209	GPL1261	GSE10239	GSM257831	CD8 P14 Memory cells-3	CD8 P14 Memory cells M
210	GPL1261	GSE10239	GSM257832	CD8 P14 day 4.5 post-infection KLRG-1 Int sorted cells-1	CD8 P14 day4.5pi KLRG-1 Int sorted cells E
211	GPL1261	GSE10239	GSM257833	CD8 P14 day 4.5 post-infection KLRG-1 Int sorted cells-2	CD8 P14 day4.5pi KLRG-1 Int sorted cells E
212	GPL1261	GSE10239	GSM257834	CD8 P14 day 4.5 post-infection KLRG-1 Int sorted cells-3	CD8 P14 day4.5pi KLRG-1 Int sorted cells E
213	GPL1261	GSE10239	GSM257835	CD8 P14 day 4.5 post-infection KLRG-1 Hi sorted cells-1	CD8 P14 day4.5pi KLRG-1 Hi sorted cells E
214	GPL1261	GSE10239	GSM257836	CD8 P14 day 4.5 post-infection KLRG-1 Hi sorted cells-2	CD8 P14 day4.5pi KLRG-1 Hi sorted cells E
215	GPL1261	GSE10239	GSM257837	CD8 P14 day 4.5 post-infection KLRG-1 Hi sorted cells-3	CD8 P14 day4.5pi KLRG-1 Hi sorted cells E
216	GPL1261	GSE9997	GSM252716	Naive_CD8+_T_cells	Naive_CD8+_T_cells N
217	GPL1261	GSE9997	GSM252717	Activated_CD8+_T_cells	Activated_CD8+_T_cells E
218	GPL1261	GSE6085	GSM108334	T-cell_IL-2_0hour_rep1	cytotoxic T cell_IL-2_0hour
219	GPL1261	GSE6085	GSM108335	T-cell_IL-2_0hour_rep2	cytotoxic T cell_IL-2_0hour
220	GPL1261	GSE6085	GSM108336	T-cell_IL-2_0hour_rep3	cytotoxic T cell_IL-2_0hour
221	GPL1261	GSE6085	GSM108337	T-cell_IL-2_0hour_rep4	cytotoxic T cell_IL-2_0hour
222	GPL1261	GSE6085	GSM108338	T-cell_IL-2_0hour_rep5	cytotoxic T cell_IL-2_0hour
223	GPL1261	GSE6085	GSM183455	T-cell_No IL-2_4 hour_rep1	cytotoxic T cell_No IL-2_4 hour
224	GPL1261	GSE6085	GSM183456	T-cell_No IL-2_4 hour_rep2	cytotoxic T cell_No IL-2_4 hour
225	GPL1261	GSE6085	GSM183457	T-cell_No IL-2_4 hour_rep3	cytotoxic T cell_No IL-2_4 hour
226	GPL1261	GSE6085	GSM183458	T-cell_No IL-2_4 hour_rep4	cytotoxic T cell_No IL-2_4 hour
227	GPL1261	GSE6085	GSM183459	T-cell_No IL-2_8 hour_rep1	cytotoxic T cell_No IL-2_8 hour
228	GPL1261	GSE6085	GSM183460	T-cell_No IL-2_8 hour_rep2	cytotoxic T cell_No IL-2_8 hour
229	GPL1261	GSE6085	GSM183461	T-cell_No IL-2_8 hour_rep3	cytotoxic T cell_No IL-2_8 hour
230	GPL1261	GSE6085	GSM140923	T-cell_IL-2_0.5 hour_rep1	cytotoxic T cell_IL-2_0.5 hour
231	GPL1261	GSE6085	GSM140924	T-cell_IL-2_0.5 hour_rep2	cytotoxic T cell_IL-2_0.5 hour
232	GPL1261	GSE6085	GSM140925	T-cell_IL-2_0.5 hour_rep3	cytotoxic T cell_IL-2_0.5 hour
233	GPL1261	GSE6085	GSM140926	T-cell_IL-2_1 hour_rep1	cytotoxic T cell_IL-2_1 hour
234	GPL1261	GSE6085	GSM140927	T-cell_IL-2_1 hour_rep2	cytotoxic T cell_IL-2_1 hour
235	GPL1261	GSE6085	GSM140928	T-cell_IL-2_1 hour_rep3	cytotoxic T cell_IL-2_1 hour
236	GPL1261	GSE6085	GSM140929	T-cell_IL-2_2 hour_rep1	cytotoxic T cell_IL-2_2 hour
237	GPL1261	GSE6085	GSM140930	T-cell_IL-2_2 hour_rep2	cytotoxic T cell_IL-2_2 hour
238	GPL1261	GSE6085	GSM140931	T-cell_IL-2_2 hour_rep3	cytotoxic T cell_IL-2_2 hour
239	GPL1261	GSE6085	GSM108339	T-cell_IL-2_4 hour_rep1	cytotoxic T cell_IL-2_4 hour
240	GPL1261	GSE6085	GSM108340	T-cell_IL-2_4 hour_rep2	cytotoxic T cell_IL-2_4 hour
241	GPL1261	GSE6085	GSM108341	T-cell_IL-2_4 hour_rep3	cytotoxic T cell_IL-2_4 hour
242	GPL1261	GSE6085	GSM108342	T-cell_IL-2_4 hour_rep4	cytotoxic T cell_IL-2_4 hour
243	GPL1261	GSE6085	GSM140932	T-cell_IL-2_6 hour_rep1	cytotoxic T cell_IL-2_6 hour
244	GPL1261	GSE6085	GSM140933	T-cell_IL-2_6 hour_rep2	cytotoxic T cell_IL-2_6 hour
245	GPL1261	GSE6085	GSM140934	T-cell_IL-2_6 hour_rep3	cytotoxic T cell_IL-2_6 hour
246	GPL1261	GSE6085	GSM140935	T-cell_IL-2_8 hour_rep1	cytotoxic T cell_IL-2_8 hour
247	GPL1261	GSE6085	GSM140936	T-cell_IL-2_8 hour_rep2	cytotoxic T cell_IL-2_8 hour
248	GPL1261	GSE6085	GSM140937	T-cell_IL-2_8 hour_rep3	cytotoxic T cell_IL-2_8 hour
249	GPL1261	GSE6085	GSM140938	T-cell_IL-2_10 hour_rep1	cytotoxic T cell_IL-2_10 hour
250	GPL1261	GSE6085	GSM140939	T-cell_IL-2_10 hour_rep2	cytotoxic T cell_IL-2_10 hour
251	GPL1261	GSE6085	GSM140940	T-cell_IL-2_10 hour_rep3	cytotoxic T cell_IL-2_10 hour
252	GPL1261	GSE6085	GSM140941	T-cell_IL-2_12 hour_rep1	cytotoxic T cell_IL-2_12 hour
253	GPL1261	GSE6085	GSM140942	T-cell_IL-2_12 hour_rep2	cytotoxic T cell_IL-2_12 hour
254	GPL1261	GSE6085	GSM140943	T-cell_IL-2_12 hour_rep3	cytotoxic T cell_IL-2_12 hour
255	GPL1261	GSE6085	GSM140944	T-cell_IL-2_16 hour_rep1	cytotoxic T cell_IL-2_16 hour
256	GPL1261	GSE6085	GSM140945	T-cell_IL-2_16 hour_rep2	cytotoxic T cell_IL-2_16 hour
257	GPL1261	GSE6085	GSM140946	T-cell_IL-2_16 hour_rep3	cytotoxic T cell_IL-2_16 hour
258	GPL1261	GSE6085	GSM140947	T-cell_IL-2_24 hour_rep1	cytotoxic T cell_IL-2_24 hour
259	GPL1261	GSE6085	GSM140948	T-cell_IL-2_24 hour_rep2	cytotoxic T cell_IL-2_24 hour
260	GPL1261	GSE6085	GSM140949	T-cell_IL-2_24 hour_rep3	cytotoxic T cell_IL-2_24 hour
261	GPL1261	GSE8039	GSM198411	C1 peripheral lymph node	peripheral lymph node
262	GPL1261	GSE8039	GSM198412	C2 peripheral lymph node	peripheral lymph node
263	GPL1261	GSE8039	GSM198413	C3 peripheral lymph node	peripheral lymph node
264	GPL1261	GSE8039	GSM198414	C4 peripheral lymph node	peripheral lymph node
265	GPL1261	GSE8039	GSM198409	T1 CD2-CD30L peripheral lymph node	CD2-CD30L peripheral lymph node
266	GPL1261	GSE8039	GSM198410	T2 CD2-CD30L peripheral lymph node	CD2-CD30L peripheral lymph node
267	GPL1261	GSE8039	GSM198415	T3 CD2-CD30L peripheral lymph node	CD2-CD30L peripheral lymph node
268	GPL1261	GSE8039	GSM198416	T4 CD2-CD30L peripheral lymph node	CD2-CD30L peripheral lymph node
269	GPL1261	GSE8039	GSM198423	C1 mesenteric lymph node	mesenteric lymph node
270	GPL1261	GSE8039	GSM198424	C2 mesenteric lymph node	mesenteric lymph node
271	GPL1261	GSE8039	GSM198417	C3 mesenteric lymph node	mesenteric lymph node
272	GPL1261	GSE8039	GSM198418	C4 mesenteric lymph node	mesenteric lymph node
273	GPL1261	GSE8039	GSM198419	T3 mesenteric lymph node	mesenteric lymph node
274	GPL1261	GSE8039	GSM198420	T4 mesenteric lymph node	mesenteric lymph node
275	GPL1261	GSE8039	GSM198421	T1 CD2-CD30L mesenteric lymph node	CD2-CD30L mesenteric lymph node
276	GPL1261	GSE8039	GSM198422	T2 CD2-CD30L mesenteric lymph node	CD2-CD30L mesenteric lymph node
277	GPL1261	GSE8039	GSM198427	C1 spleen	spleen
278	GPL1261	GSE8039	GSM198428	C2 spleen	spleen
279	GPL1261	GSE8039	GSM198431	C3 spleen	spleen
280	GPL1261	GSE8039	GSM198432	C4 spleen	spleen
281	GPL1261	GSE8039	GSM198425	T1 CD2-CD30L spleen	CD2-CD30L spleen
282	GPL1261	GSE8039	GSM198426	T2 CD2-CD30L spleen	CD2-CD30L spleen
283	GPL1261	GSE8039	GSM198429	T3 CD2-CD30L spleen	CD2-CD30L spleen
284	GPL1261	GSE8039	GSM198430	T4 CD2-CD30L spleen	CD2-CD30L spleen
285	GPL1261	GSE8039	GSM198439	C1 Thymus	Thymus
286	GPL1261	GSE8039	GSM198440	C2 Thymus	Thymus
287	GPL1261	GSE8039	GSM198433	C3 Thymus	Thymus
288	GPL1261	GSE8039	GSM198434	C4 Thymus	Thymus
289	GPL1261	GSE8039	GSM198437	T1 CD2-CD30L Thymus	CD2-CD30L Thymus
290	GPL1261	GSE8039	GSM198438	T2 CD2-CD30L Thymus	CD2-CD30L Thymus
291	GPL1261	GSE8039	GSM198435	T3 CD2-CD30L Thymus	CD2-CD30L Thymus

292	GPL1261	GSE8039	GSM198436	T4 CD2-CD30L Thymus	CD2-CD30L Thymus	
293	GPL1261	GSE8678	GSM215223	Effector CD8 T cells, IL-7Rlo, rep1	Effector CD8 T cells, IL-7Rlo	E
294	GPL1261	GSE8678	GSM215224	Effector CD8 T cells, IL-7Rlo, rep2	Effector CD8 T cells, IL-7Rlo	E
295	GPL1261	GSE8678	GSM215225	Effector CD8 T cells, IL-7Rlo, rep3	Effector CD8 T cells, IL-7Rlo	E
296	GPL1261	GSE8678	GSM215226	Effector CD8 T cells, IL-7Rhi, rep1	Effector CD8 T cells, IL-7Rhi	MP
297	GPL1261	GSE8678	GSM215227	Effector CD8 T cells, IL-7Rhi, rep2	Effector CD8 T cells, IL-7Rhi	MP
298	GPL1261	GSE8678	GSM215228	Effector CD8 T cells, IL-7Rhi, rep3	Effector CD8 T cells, IL-7Rhi	MP
299	GPL1261	GSE2924	GSM62977	CD8+Tet-_04	CD8+Tet-	E
300	GPL1261	GSE2924	GSM62978	CD8+Tet-_02	CD8+Tet-	E
301	GPL1261	GSE2924	GSM62979	CD8+Tet-_03	CD8+Tet-	E
302	GPL1261	GSE2924	GSM62980	CD8+Tet-_01	CD8+Tet-	E
303	GPL1261	GSE2924	GSM62981	CD8+Tet-_05	CD8+Tet-	E
304	GPL1261	GSE2924	GSM62982	CD8+HYTet+_01	CD8+HYTet+	E
305	GPL1261	GSE2924	GSM62983	CD8+HYTet+_02	CD8+HYTet+	E
306	GPL1261	GSE2924	GSM62984	CD8+HYTet+_03	CD8+HYTet+	E
307	GPL1261	GSE2924	GSM62985	CD8+HYTet+_04	CD8+HYTet+	E
308	GPL1261	GSE2924	GSM62986	CD8+H7aTet+_01	CD8+H7aTet+	E
309	GPL1261	GSE2924	GSM62987	CD8+H7aTet+_02	CD8+H7aTet+	E
310	GPL1261	GSE2924	GSM62988	CD8+H7aTet+_03	CD8+H7aTet+	E
311	GPL1261	GSE2924	GSM62876	CD8+H7aTet+_04	CD8+H7aTet+	E
312	GPL1261	GSE13493	GSM340091	DP thymocytes 1	CD4+CD8+ thymocytes	
313	GPL1261	GSE13493	GSM340092	DP thymocytes 2	CD4+CD8+ thymocytes	
314	GPL1261	GSE13493	GSM340093	CD4intCD8+ thymocytes 1	CD4intCD8+ thymocytes	
315	GPL1261	GSE13493	GSM340094	CD4intCD8+ thymocytes 2	CD4intCD8+ thymocytes	
316	GPL1261	GSE13493	GSM340095	CD8SP thymocytes 1	CD8SP thymocytes	N
317	GPL1261	GSE13493	GSM340096	CD8SP thymocytes 2	CD8SP thymocytes	N
318	GPL1261	GSE10813	GSM273000	Fresh CD8 rep1	Fresh CD8	N
319	GPL1261	GSE10813	GSM273001	Fresh CD8 rep2	Fresh CD8	N
320	GPL1261	GSE10813	GSM273005	Stimulated CD8 rep1	Stimulated CD8	E
321	GPL1261	GSE10813	GSM273006	Stimulated CD8 rep2	Stimulated CD8	E
322	GPL1261	GSE10813	GSM273002	Expanded CD8 Tcells with Suppression function rep2	Expanded CD8 Tcells with suppression function	
323	GPL1261	GSE10813	GSM272995	Expanded CD8 Tcells with Suppression function rep1	Expanded CD8 Tcells with suppression function	
324	GPL1261	GSE10813	GSM272996	Expanded CD8 Tcells with Suppression function rep3	Expanded CD8 Tcells with suppression function	
325	GPL1261	GSE10813	GSM272997	Expanded CD8 Tcells with Suppression function rep4	Expanded CD8 Tcells with suppression function	
326	GPL1261	GSE10813	GSM272998	Expanded CD8 Tcells with Suppression function rep5	Expanded CD8 Tcells with suppression function	
327	GPL1261	GSE14699	GSM366905	OT-I_T_cell_Naive_rep1	OT-I_T_cell_Naive	N
328	GPL1261	GSE14699	GSM366912	OT-I_T_cell_Naive_rep2	OT-I_T_cell_Naive	N
329	GPL1261	GSE14699	GSM366911	OT-I_T_cell_Rag_rep1	OT-I_T_cell_Rag	N
330	GPL1261	GSE14699	GSM366907	OT-I_T_cell_RIP-OVAhi_rep1	OT-I_T_cell_RIP-OVAhi	
331	GPL1261	GSE14699	GSM366914	OT-I_T_cell_RIP-OVAhi_rep2	OT-I_T_cell_RIP-OVAhi	
332	GPL1261	GSE14699	GSM366909	OT-I_T_cell_OCS/LPS_rep1	OT-I_T_cell_OCS/LPS	E
333	GPL1261	GSE14699	GSM366916	OT-I_T_cell_OCS/LPS_rep2	OT-I_T_cell_OCS/LPS	E
334	GPL1261	GSE4178	GSM95588	NK1.1 minus TGFbeta RII KO T cells T_KO	NK1.1- TGFbeta RII KO T cells T_KO	N
335	GPL1261	GSE4178	GSM95589	NK1.1 minus control T cells T_WT	NK1.1- control T cells T_WT	N
336	GPL1261	GSE6810	GSM155205	Control replicate 1	Control	
337	GPL1261	GSE6810	GSM155248	Control replicate 2	Control	
338	GPL1261	GSE6810	GSM155249	B. abortus 2308 replicate 1	B. abortus 2308	
339	GPL1261	GSE6810	GSM155251	B. abortus 2308 replicate 2	B. abortus 2308	
340	GPL1261	GSE6810	GSM155252	B. abortus BA41 replicate 1	B. abortus BA41	
341	GPL1261	GSE6810	GSM155253	B. abortus 2308 BA41 replicate 2	B. abortus BA41	
342	GPL1261	GSE6810	GSM155254	B. abortus 2308 ADH4.2 replicate 1	B. abortus 2308 ADH4.2	
343	GPL1261	GSE6810	GSM155255	B. abortus 2308 ADH4.2 replicate 2	B. abortus 2308 ADH4.2	
344	GPL1261	GSE6810	GSM155256	B. melitensis 16M replicate 1	B. melitensis 16M	
345	GPL1261	GSE6810	GSM155257	B. melitensis 16M replicate 2	B. melitensis 16M	
346	GPL1261	GSE15037	GSM375666	Wild-type CD8 T cell	Wild-type CD8 T cell	N
347	GPL1261	GSE15037	GSM375667	Foxo1KO CD8 T cell	Foxo1KO CD8 T cell	N
348	GPL1261	GSE1566	GSM26964	LNT-1	LN T cells-WT	N
349	GPL1261	GSE1566	GSM26965	LNT-2	LN T cells-WT	N
350	GPL1261	GSE1566	GSM26966	LN-T-3	LN T cells-WT	N
351	GPL1261	GSE1566	GSM26967	LNT-4	LN T cells-Ezh2 KO	N
352	GPL1261	GSE1566	GSM26968	LNT-5	LN T cells-Ezh2 KO	N
353	GPL1261	GSE1566	GSM26969	LNT-6	LN T cells-Ezh2 KO	N
354	GPL1261	GSE7694	GSM186666	C57BL/6J female number 1, untreated control ear (#1_Wt-C-3)	cutaneous T cell	
355	GPL1261	GSE7694	GSM186667	C57BL/6J female number 2, untreated control ear (#2_Wt-C-4)	cutaneous T cell	
356	GPL1261	GSE7694	GSM186668	C57BL/6J female number 3, untreated control ear (#3_Wt-C-7)	cutaneous T cell	
357	GPL1261	GSE7694	GSM186709	C57BL/6J female number 1, DNFB-treated allergic ear (#4_Wt-Tr-3)	cutaneous T cell	
358	GPL1261	GSE7694	GSM186710	C57BL/6J female number 2, DNFB-treated allergic ear (#5_Wt-Tr-4)	cutaneous T cell	
359	GPL1261	GSE7694	GSM186711	C57BL/6J female number 3, DNFB-treated allergic ear (#6_Wt-Tr-7)	cutaneous T cell	
360	GPL1261	GSE7694	GSM186743	Cnr1-/-/Cnr2-/- female number 1, untreated control ear (#7_Ko-C-4)	cutaneous T cell	
361	GPL1261	GSE7694	GSM187134	Cnr1-/-/Cnr2-/- female number 2, untreated control ear (#8_Ko-C-5)	cutaneous T cell	
362	GPL1261	GSE7694	GSM187135	Cnr1-/-/Cnr2-/- female number 3, untreated control ear (#9_Ko-C-7)	cutaneous T cell	
363	GPL1261	GSE7694	GSM187136	Cnr1-/-/Cnr2-/- female number 1, DNFB-treated allergic ear (#10_Ko-Tr-4)	cutaneous T cell	
364	GPL1261	GSE7694	GSM187137	Cnr1-/-/Cnr2-/- female number 2, DNFB-treated allergic ear (#11_Ko-Tr-5)	cutaneous T cell	
365	GPL1261	GSE7694	GSM187138	Cnr1-/-/Cnr2-/- female number 3, DNFB-treated allergic ear (#12_Ko-Tr-7)	cutaneous T cell	
366	GPL1261	GSE6506	GSM149585	CD8+ naive T-cells t8n_1	CD8+ T-cells	N
367	GPL1261	GSE6506	GSM149586	CD8+ naive T-cells t8n_2	CD8+ T-cells	N
368	GPL1261	GSE6506	GSM149589	CD8+ activated T-cells t8a_1	CD8+ T-cells	E
369	GPL1261	GSE6506	GSM149590	CD8+ activated T-cells t8a_2	CD8+ T-cells	E
370	GPL1261	GSE7050	GSM161520	thymus_control1	thymus	
371	GPL1261	GSE7050	GSM161521	thymus_control2	thymus	
372	GPL1261	GSE7050	GSM161526	thymus_control3	thymus	
373	GPL1261	GSE7050	GSM161527	thymus_control4	thymus	
374	GPL1261	GSE7050	GSM161528	thymus_control5	thymus	
375	GPL1261	GSE7050	GSM161522	thymus_lymphoma1	thymus	
376	GPL1261	GSE7050	GSM161523	thymus_lymphoma2	thymus	
377	GPL1261	GSE7050	GSM161514	thymus_lymphoma3	thymus	
378	GPL1261	GSE7050	GSM161515	thymus_lymphoma4	thymus	
379	GPL1261	GSE7050	GSM161516	thymus_lymphoma5	thymus	
380	GPL1261	GSE7050	GSM161517	thymus_lymphoma6	thymus	
381	GPL1261	GSE7050	GSM161518	thymus_lymphoma7	thymus	
382	GPL1261	GSE7050	GSM161519	thymus_lymphoma8	thymus	
383	GPL1261	GSE12581	GSM315613	T-cell control sample	T cell	
384	GPL1261	GSE12581	GSM315600	T-cell sample 1	T-Cell leukemia	
385	GPL1261	GSE12581	GSM315601	T-cell sample 2	T-Cell leukemia	
386	GPL1261	GSE12581	GSM315602	T-cell sample 3	T-Cell leukemia	

* collected till Sept 2009

** Classification: N, naïve; E,effector; M, memory;MP,memory precursor

Supplementary Table S2. List of the 60 TFs identified by master regulator analysis.

TF Name	#MSG	#Bkg	P-value	ISP*	DNA Motif	FDR	ESG	MSG	qPCR.E/N	qPCR.M/N
Sox4	50	4	9.55E-10	YES	SORY	0.00012	YES	YES	-4.305	-5.845
Zfp36l1	41	10	7.05E-07	NO	-	-	NO	NO	-	-
Tcf7	40	10	1.10E-06	YES	LEFF	0.00104	YES	NO	-3.71	-0.37
Hopx	37	7	9.01E-07	NO	-	-	YES	YES	-	-
Eomes	34	4	1E-06	YES	BRAC	0.00209	YES	YES	1.99	3.48
Ets2	28	3	1.17E-05	NO	ETSF	0.16476	YES	NO	-	-
Bhlhe40	27	1	2.50E-05	YES	HESF	0.33531	YES	YES	2.72	4.215
Prdm1	26	2	2.9E-05	YES	PRDF	-	YES	NO	4.185	4.5
Klf2	24	9	0.00059	YES	KLFS	0.11349	NO	NO	-2.01	-0.63
Pou2af1	23	3	9.14E-05	YES	-	-	NO	NO	-	-
Bach2	22	5	0.00023	YES	AP1R	0.01508	YES	NO	-3.77	-0.935
Runx2	21	0	0.00036	YES	HAML	0.00308	YES	YES	0.19	2.775
Tcf4	21	9	0.00191	YES	EBOX	0.29778	NO	NO	-	-
Id2	20	2	0.00032	YES	-	-	YES	YES	0.265	2.645
Smad1	20	5	0.00052	NO	SMAD	0.76295	YES	NO	-	-
Stat4	17	6	0.00265	YES	STAT	0.25029	NO	NO	-2.33	0.935
Nr4a2	16	1	0.00187	NO	NBRE	-	NO	NO	-	-
Mef2c	16	4	0.00191	YES	MEF2	-	NO	NO	-	-
Zfp318	15	3	0.00234	-	-	-	YES	NO	-	-
Runx3	15	4	0.00287	YES	HAML	0.00308	NO	NO	-2.37	-1.925
Id3	14	0	0.00542	YES	-	-	NO	NO	-3.695	-1.995
Fosl2	12	1	0.00881	NO	AP1F	0.00459	NO	YES	-	-
Nfatc2	12	2	0.00756	YES	NFAT	0.01269	NO	NO	-	-
Zfp422	12	3	0.00781	-	-	-	NO	NO	-	-
Ikzf2	12	4	0.00961	-	IKRS	0.0208	NO	NO	-	-
Zfp52	12	5	0.01318	-	-	-	NO	NO	-	-
Jhdm1d	12	6	0.01878	-	-	-	NO	NO	-	-
Etv3	12	7	0.02675	-	ETSF	0.16476	NO	NO	-	-
Zfp238	11	0	0.01707	-	RP58	-	YES	NO	-	-
Gata3	11	1	0.01297	YES	GATA	0.04247	NO	NO	-	-
Bbx	11	2	0.01122	-	-	-	NO	NO	-	-
Enc1	11	3	0.01166	-	-	-	NO	NO	-	-
Cux1	11	4	0.01435	YES	CLOX	0.6281	NO	NO	-	-
Nsd1	11	4	0.01435	NO	-	-	NO	NO	-	-
Ikzf1	11	7	0.03862	YES	IKRS	0.0208	NO	NO	-	-
Nfatc1	11	7	0.03862	YES	NFAT	0.01269	NO	NO	-	-
Phf3	11	7	0.03862	-	-	-	NO	NO	-	-
Jun	10	1	0.01909	YES	AP1F	0.00459	NO	NO	-	-
Nr4a1	10	1	0.01909	YES	NBRE	-	NO	NO	-	-
Nfe2l2	10	1	0.01909	YES	AP1R	0.01508	NO	NO	-	-
Egr1	10	4	0.0214	NO	EGRF	0.24216	NO	NO	-	-
Mxi1	9	0	0.03656	YES	EBOX	0.29778	NO	NO	-	-
Pbx3	9	0	0.03656	NO	-	-	YES	NO	-	-
Tbx21	9	0	0.03656	YES	BRAC	0.00209	NO	NO	6.215	3.165
Jund	9	1	0.02808	YES	AP1F	0.00459	NO	NO	-	-
Klf10	9	1	0.02808	NO	SP1F	0.9321	NO	NO	-	-
Tsc22d1	9	2	0.02469	NO	-	-	NO	NO	-	-
Ar	9	2	0.02469	YES	GREF	-	NO	NO	-	-
Maf	9	3	0.02597	YES	AP1R	0.01508	YES	NO	-	-
Rfx1	9	3	0.02597	-	XBBF	0.19098	NO	NO	-	-
Klf3	9	5	0.04257	NO	KLFS	0.11349	NO	NO	-	-
Arntl	8	1	0.04132	NO	HIFF	0.34596	NO	NO	-	-
Zfp709	8	1	0.04132	-	-	-	NO	NO	-	-
Foxm1	8	2	0.03666	NO	FKHD	0.00021	NO	NO	-	-
Xrcc6	8	3	0.03877	YES	-	-	NO	NO	-	-
Smad3	8	3	0.03877	YES	SMAD	0.76295	NO	NO	-	-
Atxn1	8	3	0.03877	NO	-	-	NO	NO	-	-
Zfp826	8	4	0.04742	YES	-	-	NO	NO	-	-
Sp1	8	4	0.04742	NO	SP1F	0.9321	NO	NO	-	-
Klf4	7	1	0.0608	NO	KLFS	0.11349	NO	YES	-	-

* ISP: immune system phenotype

marked gray:known TFs that regulate memory CD8+ T cell development

Supplementary Table S3. Primers for gateway cloning of TFs

attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCT
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGT
Sox4Fgw	AAAAAGCAGGCTACCATGGTACAACAGACCAACAACGCG
Sox4Rgw	AGAAAGCTGGGTTCAGTAGGTGAAGACCAGGTTAGAG
Tcf7Fgw	AAAAAGCAGGCTACCATGTACAAAGAGACTGTCTACTCTG
Tcf7Rgw	AGAAAGCTGGGTCTAGAGCACTGTCATCGGAAGG
Klf2Fgw	AAAAAGCAGGCTGCCATGGCGCTCAGCGAGCCTATCTTG
Klf2Rgw	AGAAAGCTGGGTCTACATGTGTCGCTTCATGTGC AAG
EomesFgw	AAAAAGCAGGCTAGCATGCAGTTGGGAGAGCAG
EomesRgw	AGAAAGCTGGGTCTAGGGACTTGTGTAAAAAGCATAATAAGC
Id2Fgw	AAAAAGCAGGCTagcATGAAAGCCTTCAGTCCGGTG
Id2Rgw	AGAAAGCTGGGTTTAGCCACAGAGTACTTTGCTATCAT
Prdm1Fgw	AAAAAGCAGGCTCAGATGAGAGAGGCTTATCTCAGATGTTG
Prdm1Rgw	AGAAAGCTGGGTTCTTAAGGATCCATCGGTTCAACTGT
Bach2Fgw	AAAAAGCAGGCTGGC ATGTCTGTGGATGAGAAGC
Bach2Rgw	AGAAAGCTGGGTGCTAGGCATAATCTTTCCTGGGC
Tbx21Fgw	AAAAAGCAGGCTCGGATGGGCATCGTGGAGC
Tbx21Rgw	AGAAAGCTGGGTCAGCGGCATTTTCTCAGTTGGG
Runx2Fgw	AAAAAGCAGGCTATGCTTCATTCGCCTCACAAACAAC
Runx2Rgw	AGAAAGCTGGGTTCAATATGGCCGCCAAACAGACTC
Id3Fgw	AAAAAGCAGGCTAACATGAAGGCGCTGAGCCCG
Id3Rgw	AGAAAGCTGGGTCCGGGTCAGTGGCAAAAGCTCC
Bhlhe40Fgw	AAAAAGCAGGCTATCATGGAACGGATCCCCAGC
Bhlhe40Rgw	AGAAAGCTGGGTCCCTCCAGAGTTTAGTCTTTGGTTTCTAAGTT
Stat4Fgw	AAAAAGCAGGCTAGCATGTCTCAGTGGAATCAAG
Stat4Rgw	AGAAAGCTGGGTCCGTCATTCAGCAGAATATGG

Supplementary Table S4. Sequences for shRNAs

shSox4F	CCGGGCATCGTTCTCTCCAGAGCAA CTCGAG TTGCTCTGGAGAGAACGATGCTTTTTG
shSox4R	AATTCAAAAAGCATCGTTCTCTCCAGAGCAA CTCGAG TTGCTCTGGAGAGAACGATGC
shBhlhe40F	CCGGGCGAGGTTACAGTGTTTATAT CTCGAG ATATAAACACTGTAACCTCGCTTTTTG
shBhlhe40R	AATTCAAAAAGCGAGGTTACAGTGTTTATAT CTCGAG ATATAAACACTGTAACCTCGC
shBach2aF	CCGGAAGAGCGAATTTGGTGCACAC CTCGAG GTGTGCACCAAATTCGCTCTT TTTTG
shBach2aR	AATTCAAAAAAGAGCGAATTTGGTGCACAC CTCGAG GTGTGCACCAAATTCGCTCTT
shRunx2F	CCGGGCACGCTATTAAATCCAAATT CTCGAG AATTTGGATTTAATAGCGTGCTTTTTG
shRunx2R	AATTCAAAAAGCACGCTATTAAATCCAAATT CTCGAG AATTTGGATTTAATAGCGTGC

Supplementary Table S5. Primers for real-time PCR.

Gapdh_137	AGTATGACTCCACTCACGGC
Gapdh_402	GTTCACACCCATCACAAACA
Hprt_22	GTCGTGATTAGCGATGATGA
Hprt_245	ATGTAATCCAGCAGGTCAGC
Sox4FRT	GAGAACACTGAGGCTCTGCT
Sox4RRT	TGAACGGAATCTTGTCGCTG
Tcf7FRT	GTCTACTCTGCCTTCAATCTG
Tcf7RRT	GGGAAGTGCTGTCTATATCCG
Klf2FRT	CTATCTTGCCGTCCTTTGCC
Klf2RRT	CTCCGGGTAGTAGAAGGCAG
Eomes_1504	GTCAACACTTTGCCTCAAGC
Emoes_1862	AAGACAGGTGGGCTCATTCT
Tbx21_817	GAGGTGAATGATGGAGAGCC
Tbx21_1183	CATAACTGTGTTCCCGAGGT
Bhlh40_152	AAACTTACAAACTGCCGCAC
Bhlh40_502	ATTTCAGGTCCCGAGTGTTC
Id2_92	ACCCGATGAGTCTGCTCTAC
Id2_336	GATGCTGATGTCCGTGTTCA
Id3_107	AGCCTCTTAGCCTCTTGGAC
Id3_235	CGAGGATGTAGTCTATGACACG
Prdm1_1227	GAATGTTTCCTATGGTTCCG
Prdm1_1447	GGAGGTTACTGTAGACGGGAT
Bach2_1076	CTCAGCAACCCTTAGTCAGG
Bach2_1353	GCTCACCGCAGAGTATGAAT
Runx2_1409	CAGCCACCTTTACCTACACC
Runx2_1711	CGTCAACACCATCATTCTGG
Stat4_1063	TATCAGGTGAAAGTAAAGGCG
Stat4_1443	AACCAAGTTCTGGGAGTCG

Supplementary Table S6. Primers for ChIP-PCR.

Pimer	sequence	disance to TSS (bp)
Hprt_603	GCCTTGAACTCAGAAATCCG	-397
Hprt_806	CCCAGATAATCACTCCGCA	-194
Gapdh_738	CCTATCAGTTCGGAGCCCA	-262
Gapdh_902	GCCCTGCTTATCCAGTCCTA	-98
Sox4_915	GAAGCGTTAGTTACAGCGGC	-85
Sox4_1136	TGTTTCTCCCTCCCTCCTG	136
Tcf7_692	GCCCAGGTGACTGACTAATCC	-308
Tcf7_876	ACTTGACAGGGAACAGCGAC	-124
Eomes_604	ACCAGGATTGATTCTCAGTGG	-396
Eomes_857	GACCGCTTGGAAACTTGTG	-143
Tbx21_519	GACAAGAGACAGCGAGCAAAC	-481
Tbx21_777	AATCTGGGAAAGAGTCAACCTG	-223
Prdm1_182	GGATGTCCTTCCTTCTCCAGA	-818
Prdm1_509	GGATTTCATTTCTCCTGCTAACTG	-491
Klf2_566	CAAAGATTTCACAGAGCGTTCC	-434
Klf2_760	AACAAACCAGGCACAACAGC	-240
Bach2a_79	GCTGTGGAAACTCAGTGGAAG	-921
Bach2a_364	TTCATCAAAGGGAAAGCCG	-636
Bhlhe40_744	AGTGTTCGTGGAAGTGACCG	-256
Bhlhe40_951	GCTACAACCAATCAGCGGG	-49
Id2_630	GTCTGCTGTGCGGAGAGAGT	-370
Id2_848	AGCCAATGCCTGTAGGGTG	-152
Id3_834	CAGACCAGCCTAAGGAAGCC	-166
Id3_1026	GAGACACCTAAAGCAGCAAACA	26
Runx2_883	GGCAGTCCCACTTTACTTTGAG	-117
Runx2_1087	AAACGCCAGAGCCTTCTTG	87
Gzmb_707	CAAGAGATGAGCCCAAATGC	-293
Gzmb_889	GAACCCACAGCGAATGACTC	-111
Sell_810	TGGTCCACTACTGTGTTCTTGG	-190
Sell_960	TGCTTTCTTTCATTTCTCCAGC	-40
Il7r_813	TGAAGTCTACCACCTTGCTCAA	-187
Il7r_1024	CTCTGTGCCTGCTAAACCAC	24
Ccr7_652	CTGCTGGGACAATAGCCACT	-348
Ccr7_829	TGTAGAGACTCAGGGAGGAGC	-171
Prf1_888	CAGGGCAGGAAGTAGTAATGATATG	-112
Prf1_1036	CTTCCTCCTCCTTACCTGAAGTC	36
Ccl5_808	AGGGCAGTTAGAGGCAGAGTC	-192
Ccl5_963	CCAGGGTAGCAGAGGAAGTG	37
Gzmk_708	GTGACCCACAGGTTGAGAGC	-292
Gzmk_973	ACGGACAAACTGCTTGCTTC	-27
Il18r1_638	GTGGTTGAGTAGACAGGCTGC	-362
Il18r1_865	GCTTTCAGTTCTGGGAAGACC	-135
Il6ra_427	ACACTGGTTTCGTTGCCTC	-573
Il6ra_611	GGTGCTTACGGGAATCAAA	-389
S100a4_737	TTCCTAAACTTCTGGCTGAGC	-263
S100a4_906	TGATGTAGTAAATGTCATAGCACCC	-94
S100a6_781	ACCACCACAACTTGAAGAACAG	-219
S100a6_936	GGACTGGGAACCACACTGAT	-64
Stat4_612	AGAGAAGAATCCAGGTGGCAC	-388
Stat4_895	AGACATTGAGGACCAGGCAG	-105
Ifng_645	CTGTGCTGTGCTCTGTGGAT	-355
Ifng_823	GACTCCTTGGGCTCTCTGAC	-177
Cxcr3_557	ATGCCAGGTCTGATTCAACC	-443
Cxcr3_857	ACTTGGGACTGTTACAAAGCCT	-143
Abca3_497	TTGGTGACAAGAGTGAGATGG	-503
Abca3_672	AACAGTAGGCGTTAGTTTGGG	-328
Dusp14_595	GCCAACCTGGTCTACATAGTGAG	-405
Dusp14_880	CACCTGGACTTACTCGCAGA	-120
Blk_801	AGTCCAGTCACATCTGTTCTGC	-199
Blk_1002	CCCTTGGTTGGTTTGTTAGG	2
Ms4a1_666	TACCTTCTCAGGGATTCCGAC	-334
Ms4a1_913	ACCTTGCTTTGCCTTACCAG	-87
Ccr5_779	TTCTGATTTCCAACGAAGTGTG	-221
Ccr5_1070	GATGTCTCACCTCCTCTGGC	70
Klrc1_843	TTGTTGTTACCACAGCACTCG	-157
Klrc1_1049	AGTTCTTTGCCACCTTCAGTTC	49
Cd38_841	TCTGGAGTCTGGAAGTAAGCAGT	-159
Cd38_1017	GAAGAGAGCACAGGGCTGAC	17