

ARTICLE

Received 7 Jun 2013 | Accepted 28 Oct 2013 | Published 12 Dec 2013

DOI: 10.1038/ncomms3830

# A genome-wide regulatory network identifies key transcription factors for memory CD8<sup>+</sup> T-cell development

Guangan Hu<sup>1</sup> & Jianzhu Chen<sup>1</sup>

Memory CD8<sup>+</sup> T-cell development is defined by the expression of a specific set of memory signature genes. Despite recent progress, many components of the transcriptional control of memory CD8<sup>+</sup> T-cell development are still unknown. To identify transcription factors and their interactions in memory CD8<sup>+</sup> T-cell development, we construct a genome-wide regulatory network and apply it to identify key transcription factors that regulate memory signature genes. Most of the known transcription factors having a role in memory CD8<sup>+</sup> T-cell development are rediscovered and about a dozen new ones 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<sup>+</sup> T cells through *Prdm1* and *Id3*. Gene perturbation study identifies the interactions between the transcription factors, with *Sox4* positioned as a hub. The identified transcription factors and insights into their interactions should facilitate further dissection of molecular mechanisms underlying memory CD8<sup>+</sup> T-cell development.

<sup>1</sup>David H. Koch Institute for Integrative Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. Correspondence and requests for materials should be addressed to J.C. (email: jchen@mit.edu).

Immunological memory refers to faster and stronger responses to re-encounter 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<sup>+</sup> T cells are usually generated following antigen-stimulated T-cell activation and expansion. In a typical CD8<sup>+</sup> T-cell response, naive CD8<sup>+</sup> T cells are activated to undergo clonal expansion when stimulated by appropriate antigen<sup>1</sup>. 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<sup>+</sup> T cells. Memory CD8<sup>+</sup> T cells are often divided into two subsets. Effector memory T cells (T<sub>EM</sub>) are CD62L<sup>lo</sup>CCR7<sup>lo</sup> and capable of rapid expression of effector functions following antigen stimulation to confer faster memory response. Central memory T cells (T<sub>CM</sub>) are CD62L<sup>hi</sup>CCR7<sup>hi</sup> and proliferate extensively upon antigen restimulation to confer stronger memory response.

Memory CD8<sup>+</sup> T cells are developmentally programmed as they express a specific set of memory signature genes (MSGs)<sup>2,3</sup>, which confer them with characteristic memory phenotype and function. Like many developmental processes, memory CD8<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T cells<sup>4–6</sup>. *Tcf7* is a TF downstream of the Wnt signalling. Consistent with the observation that activation of Wnt/ $\beta$ -catenin signalling promotes memory CD8<sup>+</sup> T-cell development by suppressing terminal differentiation of effector T cells<sup>7,8</sup>, *Tcf7*-deficiency in CD8<sup>+</sup> T cells impairs T<sub>CM</sub> differentiation<sup>9</sup>. *Klf2* has been shown to be associated with memory CD8<sup>+</sup> T-cell development<sup>10</sup> probably by directly controlling the expression of cell surface receptors S1P1 and CD62L<sup>11,12</sup>. *Id2* and *Id3*, the E-box-containing transcription suppressors, appear to regulate the development of memory CD8<sup>+</sup> T-cell subsets<sup>13</sup>. *Id2*-knockout mice are deficient in memory CD8<sup>+</sup> T cells, but effector T cells generated in these mice are CD127<sup>hi</sup>CD62L<sup>hi</sup>, a phenotype similar to T<sub>CM</sub><sup>14</sup>. Deficiency in *Id3* inhibits, whereas overexpression of *Id3* promotes memory CD8<sup>+</sup> T-cell development<sup>15</sup>. The B-cell transcriptional repressor Blimp-1 (encoded by *Prdm1*) promotes the terminal differentiation of effector CD8<sup>+</sup> T cells and is required for recall response of memory T cells<sup>16,17</sup>. Despite these progresses, the current understanding of transcriptional regulation of memory CD8<sup>+</sup> 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 programme for memory CD8<sup>+</sup> T-cell development.

In this study, we assemble a genome-wide regulatory network associated with the development of CD8<sup>+</sup> 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<sup>+</sup> T-cell development using the master regulator analysis (MRA) of the MSGs of CD8<sup>+</sup> T cells. The inferred TFs include most of the known TFs as well as a dozen new TFs with limited functional information in CD8<sup>+</sup> 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<sup>+</sup> T-cell development. Four of the newly identified key TFs (*Sox4*, *Bhlhe40*, *Bach2* and *Runx2*) are experimentally validated to regulate memory CD8<sup>+</sup> T-cell development and function. *Bach2* is shown to promote memory CD8<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T-cell development.

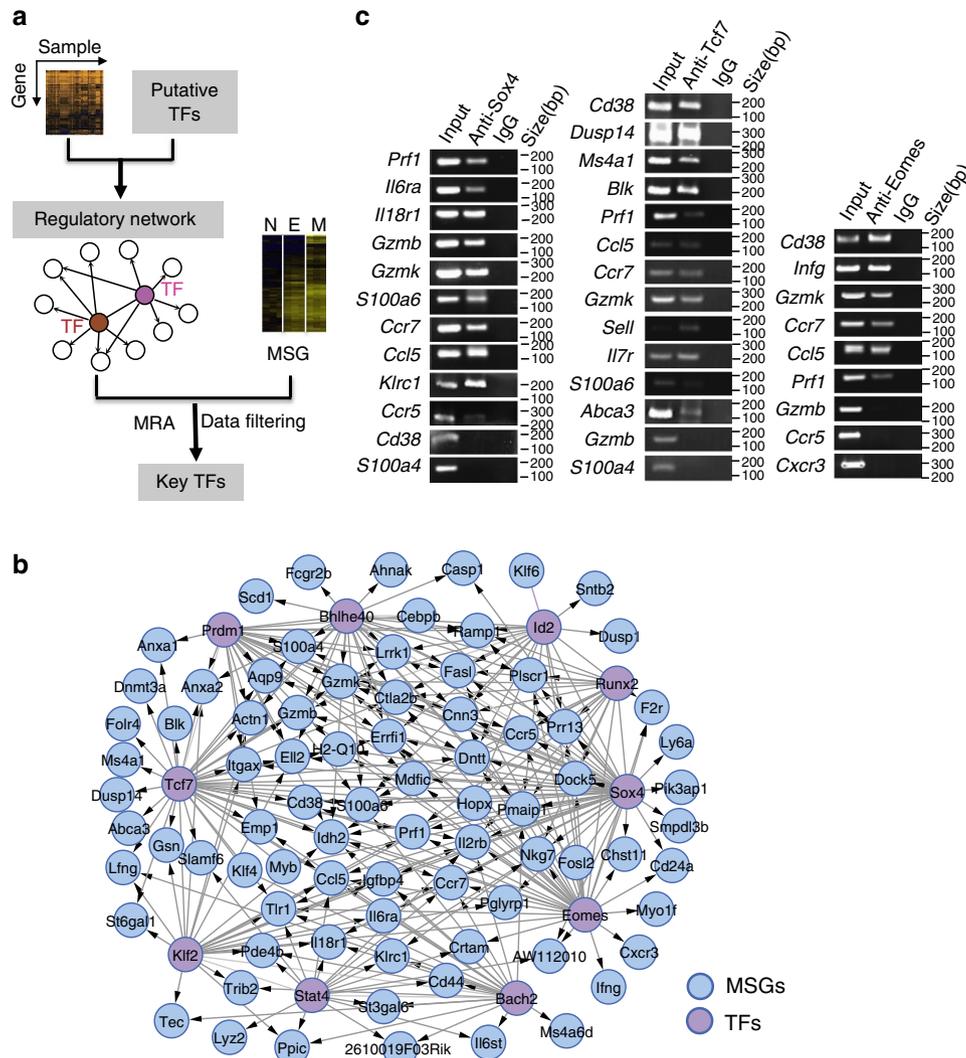
## Results

### Identification of TFs associated with memory CD8<sup>+</sup> T cells.

We collected 386 gene expression profiles of naive, effector and memory CD8<sup>+</sup> T cells of the mouse from 35 independent GEO data sets (Supplementary Table S1). A total of 1,445 genes coding putative TFs<sup>18</sup> 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 naive to effector to memory CD8<sup>+</sup> T-cell development (see Methods for detail). Using a reverse-engineering algorithm CLR (context likelihood of relatedness)<sup>19</sup>, 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. In all, 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<sup>3</sup> (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<sup>+</sup> 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 (Mouse Genome Informatics)<sup>20</sup>. The positive candidates were then analysed for enrichment of DNA-binding motifs among MSGs or differential expression among naive, effector and memory CD8<sup>+</sup> 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<sup>+</sup> T-cell development and function<sup>21–24</sup>, were identified by our analysis. These results show that our systematic approach is valid for identifying TFs that regulate memory CD8<sup>+</sup> T-cell development.

**Validation of a regulatory module for MSGs.** 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 three TFs, *Sox4*, *Tcf7* and *Eomes*, in CD8<sup>+</sup> 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,d (Supplementary Fig. S15), promoter regions of



**Figure 1 | Construction of regulatory network of memory CD8<sup>+</sup> T cells.** (a) Schematic diagram of regulatory network analysis for identifying key TFs. N, E, and M, naive, effector and memory CD8<sup>+</sup> 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 Eomes-regulated MSGs. ChIP was carried out with CD8<sup>+</sup> T cells expressing the 2C TCR using antibodies specific for Sox4, Tcf7 or Eomes or control IgG antibodies. Promoter regions of the indicated genes were amplified using the precipitated DNA. Shown are PCR products after electrophoresis.

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 two 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)<sup>6,15,25</sup> were perturbed in CD8<sup>+</sup> 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 more than twofold were taken

as directional regulations, the perturbation results identified 41 regulations among the 12 × 12 matrix (31%). Notably, the top three 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 (Supplementary Fig. S15), 18 of the 19 regulations were confirmed. ChIP-PCR also identified four more regulations that were not observed in the perturbation study. Thus, compared with 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 three 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

**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 <sup>†</sup>	Ref.
<i>Sox4</i>	50	9.55E-10	Yes	SORY	0.0001	Yes	Yes	-4.305	-5.845	Th2 differentiation	50
<i>Tcf7</i>	40	1.10E-06	Yes	LEFF	0.0010	Yes	No	-3.71	-0.37	Promote memory	7,9
<i>Eomes</i>	34	1.03E-06	Yes	BRAC	0.0021	Yes	Yes	1.99	3.48	Effector and memory	4,5
<i>Bhlhe40</i>	27	2.50E-05	Yes	HESF	0.3353	Yes	Yes	2.72	4.215	CD8+ T-cell activation	51
<i>Prdm1</i>	26	2.90E-05	Yes	PRDF	—	Yes	No	4.185	4.5	Effector/inhibit memory	16,17
<i>Klf2</i>	24	0.000587	Yes	KLFS	0.1135	No	No	-2.01	-0.63	Homeostasis/memory	10
<i>Bach2</i>	22	0.000229	Yes	AP1R	0.0151	Yes	No	-3.77	-0.935	—	—
<i>Runx2</i>	21	0.000365	Yes	HAML	0.0031	Yes	Yes	0.19	2.775	Early T-cell development	52
<i>Id2</i>	20	0.000319	Yes	—	—	Yes	Yes	0.265	2.645	KO promotes memory	13,14
<i>Stat4</i>	17	0.002650	Yes	STAT	0.2503	No	No	-2.33	0.935	Effector and memory	53,54
<i>Runx3</i>	15	0.002869	Yes	HAML	0.0031	No	No	-2.37	-1.925	Cooperation with Tbx21	55
<i>Id3</i>	14	0.005424	Yes	—	—	No	No	-3.695	-1.995	Promote memory	15
<i>Nfatc2</i>	12	0.007564	Yes	NFAT	0.0127	No	No	—	—	TCR signalling	56
<i>Gata3</i>	11	0.012973	Yes	GATA	0.0425	No	No	—	—	Th2 differentiation	57
<i>Ikzf1</i>	11	0.038617	Yes	IKRS	0.0208	No	No	—	—	IL-2/TCR signalling	58
<i>Nfatc1</i>	11	0.038617	Yes	NFAT	0.0127	No	No	—	—	PD-1/TCR signalling	59
<i>Jun</i>	10	0.019089	Yes	AP1F	0.0046	No	No	—	—	AP-1 complex	60
<i>Nfe2l2</i>	10	0.019089	Yes	AP1R	0.0151	No	No	—	—	Protect memory	61
<i>Tbx21</i>	9	0.036556	Yes	BRAC	0.0021	No	No	6.215	3.165	Effector and memory	4,5
<i>Jund</i>	9	0.028085	Yes	AP1F	0.0046	No	No	—	—	AP-1 complex	60
<i>Maf</i>	9	0.025971	Yes	AP1R	0.0151	Yes	No	—	—	—	—

DBM, DNA-binding motif; FDR, false discovery rate; ESG, effector signature genes; ISP, immune system phenotype; MSG, memory signature genes; qPCR E/N, log<sub>2</sub> gene expression fold-changes between effector and naïve CD8<sup>+</sup> T cells; qPCR M/N, log<sub>2</sub> gene expression fold-changes between memory and naïve CD8<sup>+</sup> T cells.

Grey marked TFs that have been reported as playing important roles in memory CD8<sup>+</sup> T-cell development or function.

\*MSG: the total number of MSG regulated directly by the TF.

†Gene function reported in T cells.

**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%)

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<sup>+</sup> T-cell development.

#### Validation of *Sox4* and *Bach2* in memory CD8<sup>+</sup> T cells.

Among the top 10 TFs (Table 1), 6 are known to have important roles in memory CD8<sup>+</sup> T-cell development and/or function. We then investigated whether the other four TFs (*Sox4*, *Bhlhe40*, *Bach2* and *Runx2*) are also involved in memory CD8<sup>+</sup> T-cell development/function by examining the effect of overexpression and knockdown of these TFs on the recall proliferation of memory CD8<sup>+</sup> T cells *in vitro* and *in vivo*. CD8<sup>+</sup> T cells expressing the 2C TCR were activated with cognate peptide SIYRYGL (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 Tables S3 and S4). The 2C T cells were then cultured in the presence of cytokine IL-7 to induce the development of memory CD8<sup>+</sup> 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<sup>+</sup>) and non-transduced (GFP<sup>-</sup>) 2C T cells were quantified on days 4 and 6. Compared with the vector control, overexpression of *Sox4* or *Bach2* led to a significant increase in the proportions of GFP<sup>+</sup> 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)<sup>26</sup>, a significant increase in the proportion of GFP<sup>+</sup> 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<sup>9,11</sup>. These results show that *Sox4* and *Bach2* likely promote the recall proliferation of memory CD8<sup>+</sup> T cells.

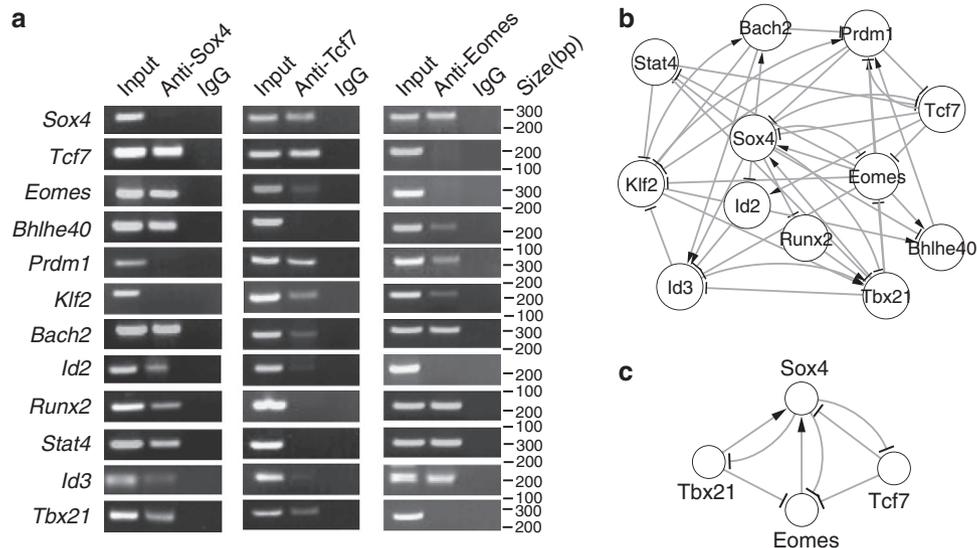
#### Enhanced memory T-cell development by *Bach2* overexpression.

To confirm the effect of overexpression of *Bach2* on recall proliferation of memory CD8<sup>+</sup> T cells, we activated 2C T cells *in vitro* for 2 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 2 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 analysed. Both transduced (GFP<sup>+</sup>) and non-transduced (GFP<sup>-</sup>) 2C cells persisted in the recipient mice (Fig. 4b). GFP<sup>+</sup> 2C cells exhibited a typical memory phenotype as indicated by expression of CD62L and IL-7 receptor (IL-7R), similar to the GFP<sup>-</sup> 2C cells in the same recipient (Fig. 4c). The persisting memory 2C cells, both transduced and non-transduced, were rapidly

**Table 3 | Perturbation analysis.**

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

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

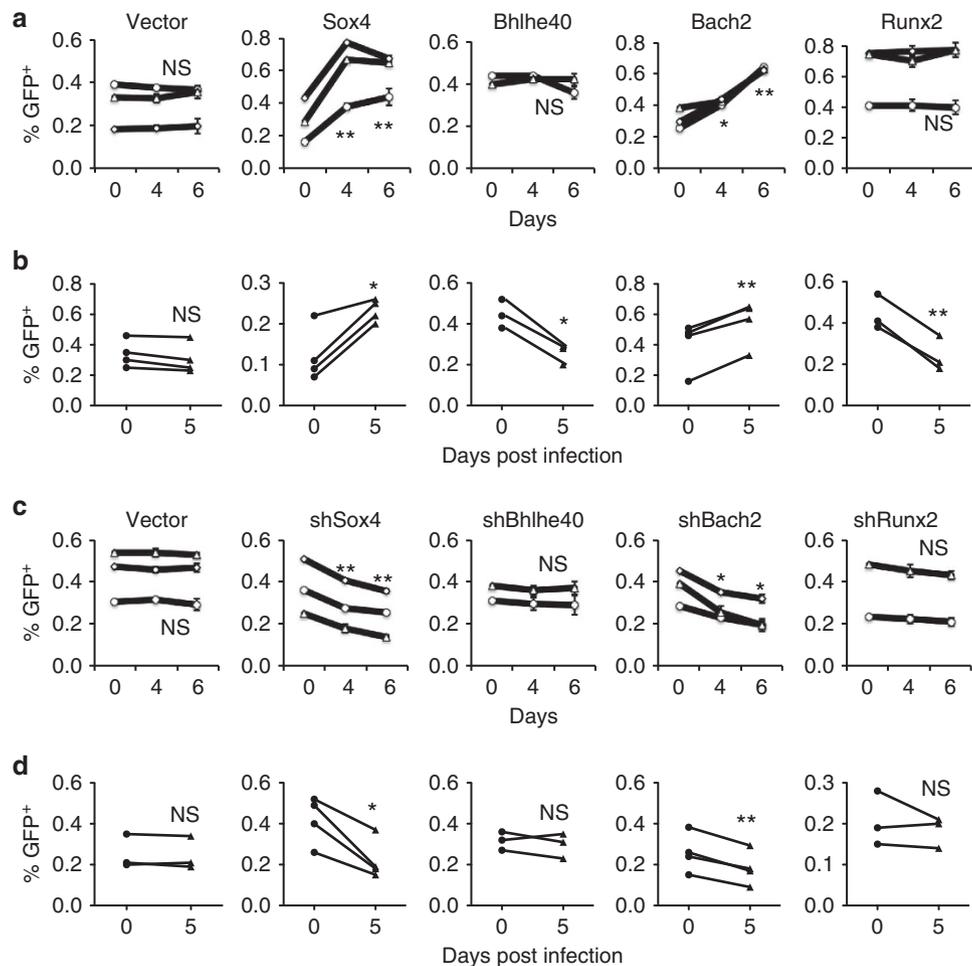


**Figure 2 | Construction of perturbation network of TFs in CD8<sup>+</sup> T cells.** (a) ChIP-PCR analysis. ChIP was carried out with 2C T cells using antibodies specific for Sox4, Tcf7 or Eomes or control IgG antibodies. Promoter regions of the indicated genes were then amplified using the precipitated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

induced to express IFN $\gamma$  and TNF $\alpha$  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 analysed 7 days later. As shown in Fig. 4e, if the 2C cells were originally transduced with GFP-expressing retrovirus, the proportion of GFP<sup>+</sup> versus GFP<sup>-</sup> 2C cells did not change following WSN-SIY challenge. However, if the 2C cells were originally transduced with GFP and *Bach2*-expressing retrovirus, the proportion of GFP<sup>+</sup> 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 2 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 analysed by flow cytometry 7 dpi. Compared with the non-transduced 2C T cells, vector-transduced 2C T cells had the same expression profiles for CD62L, IL-7R, Klr1 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 Klr1, in the DLN and spleen compared with non-transduced 2C cells in the same organs of the same recipients. When the persisting 2C T cells were analysed on 30 dpi, no significant differences in IFN $\gamma$  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



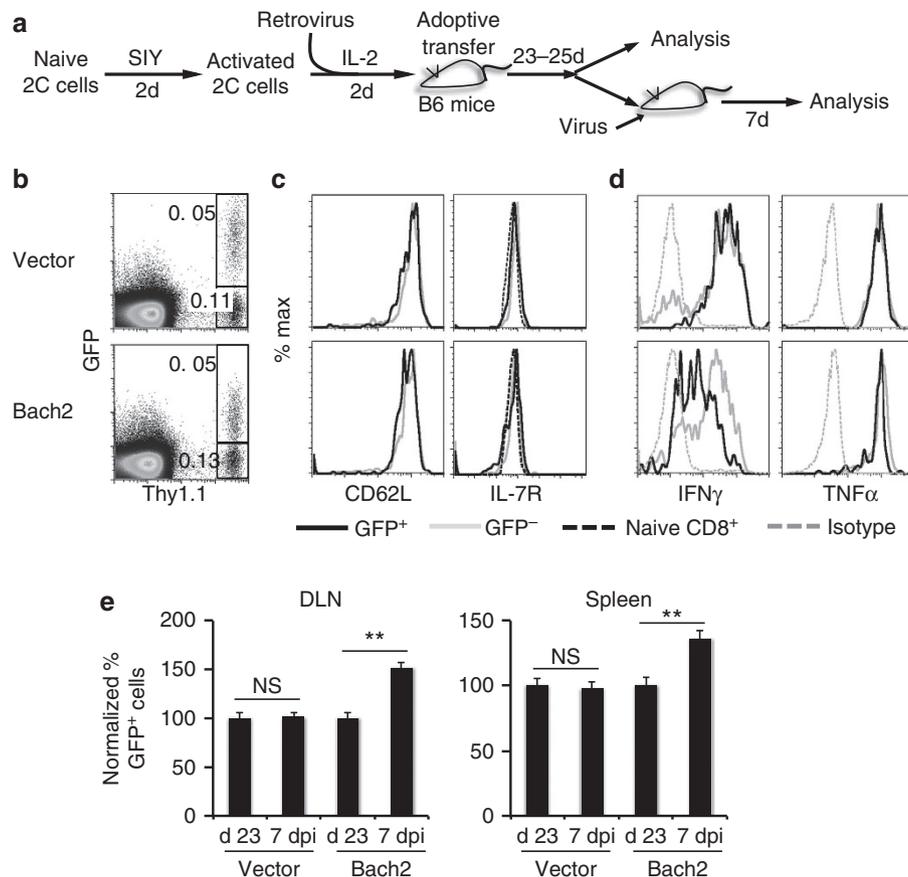
**Figure 3 | Effect of overexpression and knockdown of TFs on memory CD8<sup>+</sup> T-cell recall proliferation.** Naive 2C T cells were activated *in vitro* with SIY peptide and then transduced with retroviruses expressing *Sox4*, *Bhlhe40*, *Bach2* or *Runx2* or expressing shRNA specific for one of 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<sup>+</sup> (transduced) versus GFP<sup>-</sup> (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 proportions of GFP<sup>+</sup> 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 *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  $\pm$  s.e.m. Pairwise two-tailed *t*-tests were used for statistical analyses. \**P* < 0.05; \*\**P* < 0.01. NS, not significant.

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 results suggest that overexpression of *Bach2* likely promotes the generation of memory CD8<sup>+</sup> T cells *in vivo*.

We further investigated the effect of *Bach2* overexpression in naive 2C T cells on memory CD8<sup>+</sup> 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*<sup>-/-</sup> mice to generate naive 2C T cells that express GFP or GFP plus *Bach2* (Fig. 6a). The resulting transduced and non-transduced naive 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 analysed on 7 dpi. Vector-transduced and non-transduced 2C T cells from the same recipient mice had the same expression profile of CD62L, IL-7R, Klrp1, IFN $\gamma$  and TNF $\alpha$

(Fig. 6b). While *Bach2*-transduced 2C T cells expressed similar levels of IFN $\gamma$  and TNF $\alpha$  as non-transduced and vector-transduced 2C T cells, more cells expressed CD62L but fewer cells expressed Klrp1, resembling to CD62L<sup>hi</sup>Klrp1<sup>low</sup> 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, Klrp1, IFN $\gamma$  and TNF $\alpha$  (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<sup>+</sup> 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 in Fig. 5a, except the retrovirus expressed shRNA specific for *Bach2*. Briefly, activated 2C T cells were transduced with shRNA-expressing retrovirus and adoptively transferred into



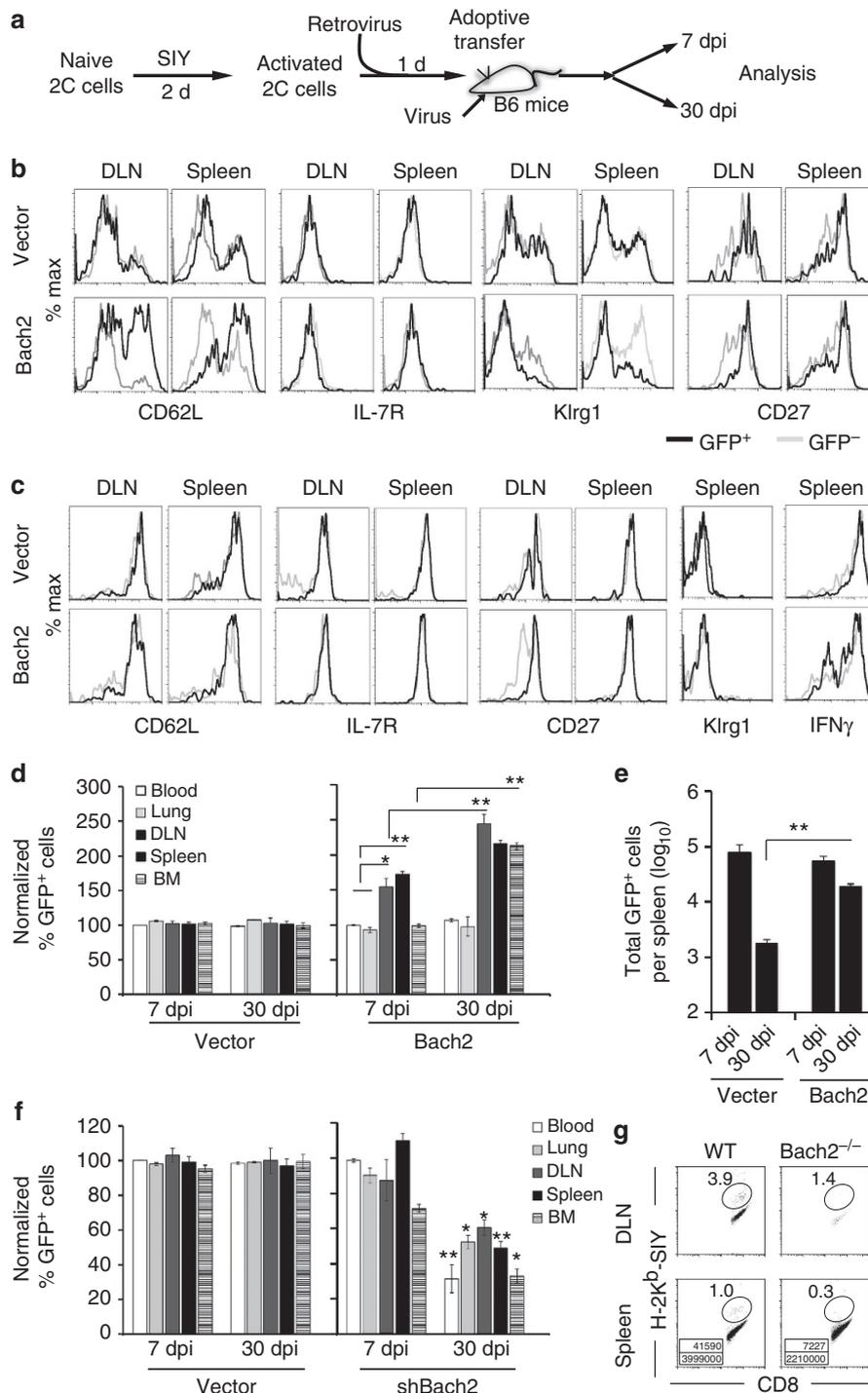
**Figure 4 | *Bach2* promotes recall proliferation of memory CD8<sup>+</sup> T cells.** (a) Scheme of experimental protocol. (b–d) Phenotype and function of persisting memory 2C cells. Twenty-two days post transfer, single-cell suspension was prepared from spleen and analysed for CD8, Thy1.1, GFP plus CD62L or IL-7R directly or stimulated *in vitro* with SIY peptide for 5 h before staining for CD8, Thy1.1, GFP plus intracellular IFN $\gamma$  or TNF $\alpha$ . 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 IFN $\gamma$  and TNF $\alpha$  expression (d) between GFP<sup>+</sup> and GFP<sup>-</sup> 2C T cells. Grey trace, non-transduced (GFP<sup>-</sup>) 2C T cells; black trace, transduced (GFP<sup>+</sup>) 2C T cells; dash grey trace, isotype control for intracellular staining; dash black trace, IL-7R staining of naive 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<sup>+</sup> and GFP<sup>-</sup> 2C T cells in the DLN and spleen was quantified by flow cytometry 7 dpi. Comparison of proportions of GFP<sup>+</sup> 2C T cells in the DLN and spleen before (d 23) and after antigen restimulation (7 dpi). Representative data from three independent experiments with 2–3 mice per group per experiment are shown as mean  $\pm$  s.e.m. Two-tailed Student's *t*-tests were used for statistical analyses. \*\**P* < 0.01; NS, not significant.

B6 mice followed by infection with WSN-SIY influenza virus. The number and frequency of GFP<sup>+</sup> and GFP<sup>-</sup> 2C T cells in various organs were analysed on 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, Klr1, CD27 and IFN $\gamma$  expression (Supplementary 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 with the numbers of non-transduced and vector-transduced 2C T cells 30 dpi (Supplementary Fig. S12c).

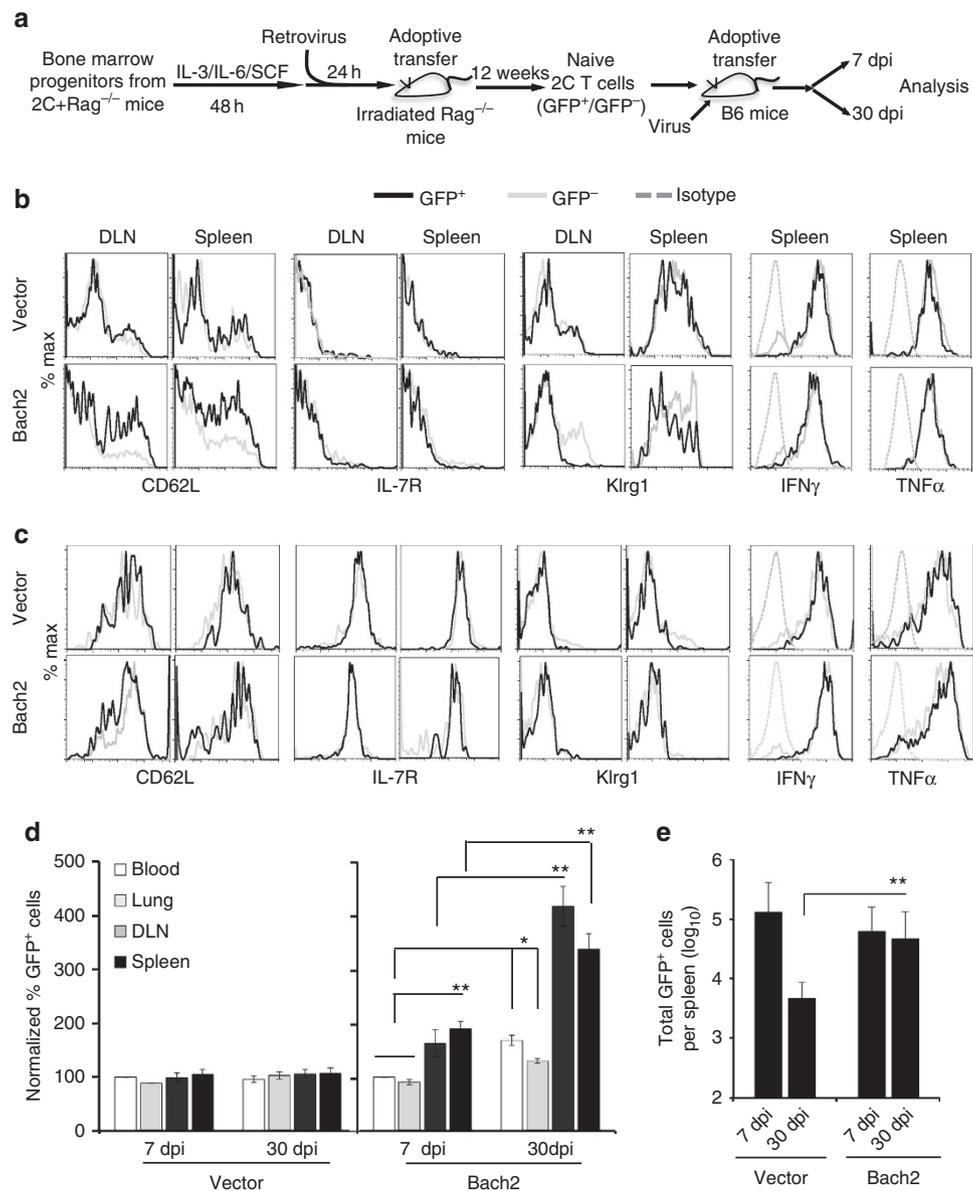
To investigate the effect of *Bach2* knockout on memory T-cell development, we constructed chimeric mice where T and B cells were deficient in *Bach2* by adoptively transferring bone marrow cells from *Bach2* knockout mice into sublethally irradiated Rag2<sup>-/-</sup> mice. Three months after reconstitution, mice were infected with WSN-SIY influenza virus and analysed for the presence of SIY-specific memory CD8<sup>+</sup> T cells 30 days later. The percentage of SIY-specific CD8<sup>+</sup> T cells was lower in the DLN and spleen of chimeric mice that were reconstituted with *Bach2*<sup>-/-</sup> than *Bach2*<sup>+/+</sup> bone marrow cells (Fig. 5g).

Consistently, the number of SIY-specific CD8<sup>+</sup> T cells in the spleen was lower in mice reconstituted with *Bach2*<sup>-/-</sup> than *Bach2*<sup>+/+</sup> bone marrow cells. Considering that *Bach2* expression was downregulated in effector and then upregulated in memory T cells during naive to effector to memory cell transition (Table 1), together these results show that *Bach2* promotes memory CD8<sup>+</sup> 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<sup>+</sup>) versus non-transduced (GFP<sup>-</sup>) 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 of 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



**Figure 5 | *Bach2* promotes memory CD8<sup>+</sup> 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 analysed for CD62L, IL-7R, Klr1, CD27 and IFN $\gamma$  as in Fig. 4. Shown are histograms of CD62L, IL-7R, Klr1 and CD27 expression of Thy1.1<sup>+</sup> CD8<sup>+</sup> 2C T cells 7 (b) and 30 dpi (c). Grey trace, non-transduced (GFP<sup>-</sup>) 2C T cells; black trace, transduced (GFP<sup>+</sup>) 2C T cells. Representative data from six mice in three independent experiments are shown. (d) Proportion of GFP<sup>+</sup> transduced 2C cells in different organs normalized to the average of the blood at 7 dpi. (e) Total *Bach2*-transduced 2C T cells (GFP<sup>+</sup>) in the spleen 7 and 30 dpi. (f) Effect of *Bach2* knockdown on memory T-cell development. Proportion of *Bach2*-knocked down 2C T cells (GFP<sup>+</sup>) 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*<sup>-/-</sup> and *Bach2*<sup>+/+</sup> bone marrow cells into sublethally irradiated Rag2<sup>-/-</sup> 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<sup>b</sup>-SIY and anti-CD8. Shown are staining profiles of H-2K<sup>b</sup>-SIY versus CD8. The numbers in the plots indicate percentage of SIY-specific memory CD8<sup>+</sup> T cells. The numbers in the boxes indicate the number of SIY-specific memory CD8<sup>+</sup> T cells (top) and total CD8<sup>+</sup> 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  $\pm$  s.e.m. Two-tailed Student's *t*-tests were used for statistical analyses. \**P* < 0.05; \*\**P* < 0.01.

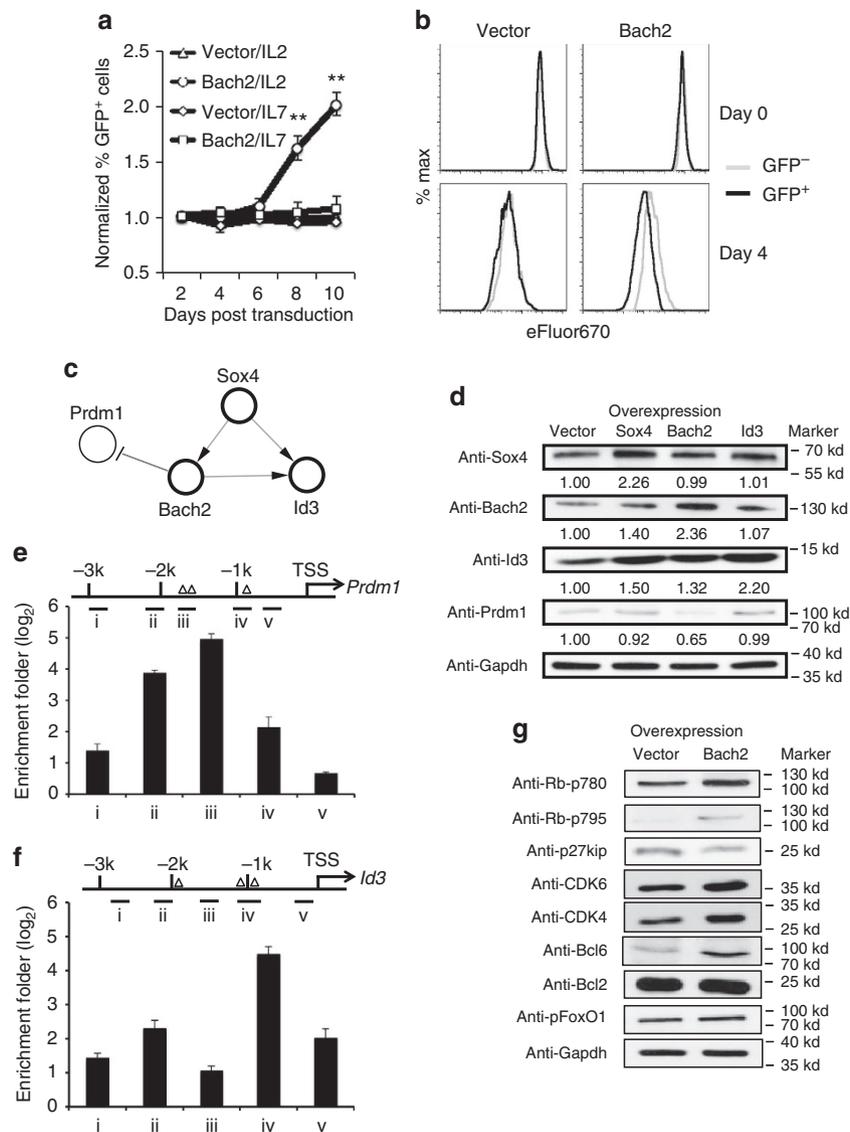


**Figure 6 | *Bach2* promotes memory CD8<sup>+</sup> T-cell development.** (a) Scheme of experimental protocol. (b–e) Persistence and phenotype of transferred CD8<sup>+</sup> T cells over time. Seven and 30 dpi, 2C T cells in various organs were analysed as in Fig. 4. Shown are histograms of CD62L, IL-7R, Klr1, IFN $\gamma$  and TNF $\alpha$  expression of Thy1.1<sup>+</sup> CD8<sup>+</sup> 2C T cells 7 (b) and 30 dpi (c). Grey trace, non-transduced (GFP<sup>-</sup>) 2C T cells; black trace, transduced (GFP<sup>+</sup>) 2C T cells; dash grey trace, isotype control for intracellular staining. Representative data from 9–12 mice in three independent experiments are shown. (d) Proportion of GFP<sup>+</sup> transduced 2C cells in different organs normalized to the average of the blood at 7 dpi. (e) Total *Bach2*-transduced 2C T cells (GFP<sup>+</sup>) in the spleen 7 and 30 dpi. Data from three independent experiments with 3–4 mice per group per experiment are shown as mean  $\pm$  s.e.m. Two-tailed Student's *t*-tests were used for statistical analyses. \**P* < 0.05; \*\**P* < 0.01.

of *Bach2*-transduced 2C T cells increased significantly over time in the IL-2 cultures (Fig. 7a). When the cells were labeled with eFluor and followed over time, *Bach2*-transduced 2C T cells diluted the fluorescent dye more extensively than non-transduced 2C T cells (Fig. 7b). These data suggest that *Bach2* promotes proliferation of activated CD8<sup>+</sup> 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; Supplementary Fig. S16). 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<sup>15,16</sup>, 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; Supplementary Fig. S16). These results suggest that *Bach2* regulates memory T-cell development and recall proliferation by regulating cell cycle control possibly through *Prdm1* and *Id3*.



**Figure 7 | *Bach2* promotes proliferation of CD8<sup>+</sup> 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<sup>+</sup> 2C T cells was quantified. Shown (mean  $\pm$  s.e.m.) are changes in proportion of GFP<sup>+</sup> 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 **a**) were labelled with eFluor670 and cultured in the presence of IL-2 for 4 days. Shown are representative histograms of eFluor670 from one of the two experiments. Grey trace, non-transduced (GFP<sup>-</sup>) 2C T cells; black trace, transduced (GFP<sup>+</sup>) 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 h 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 indicated as triangle. Data are from two independent experiments, error bar: s.e.m. **(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 (**a**). The levels of phosphorylated Rb and FoxO1, CDK4, CDK6, Bcl6, Bcl2 and p27kip in the transduced 2C T cells were assayed by western blotting.

## Discussion

At the molecular level, development of memory CD8<sup>+</sup> T cells is the establishment of MSG expression programme, which ultimately is controlled by TFs. Although several TFs have been described to regulate memory CD8<sup>+</sup> T-cell development, for a comprehensive understanding of transcriptional regulation of memory CD8<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T cells of the mouse. Applying MRA to this network, we have identified 21 key TFs, down-narrowed from 1,038 putative TFs, which regulate the expression of 70% of MSGs. Our approach is valid based on the following considerations. First, our method identified 8 of the 12 TFs known to be involved in memory CD8<sup>+</sup> 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<sup>+</sup> 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- $\kappa$ b* has important roles in both effector and memory T-cell development<sup>27,28</sup>. We found that *NF- $\kappa$ b* was a key TF that regulates effector signature genes when MRA was applied to the network and effector signature genes. Second, our method identified several TFs that are not known to function in the memory CD8<sup>+</sup> 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 have important roles in memory CD8<sup>+</sup> T-cell development and function. Overexpression of *Sox4* and *Bach2* promoted a significantly higher recall proliferation of memory CD8<sup>+</sup> 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<sup>+</sup> T-cell development (Figs 4–6). Third, compared with the traditional method of differential gene expression analysis<sup>2,3</sup>, 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<sup>+</sup> T-cell development is coordinately regulated by several TFs (reviewed in refs 22,24), including *Eomes* and *Tbx21* (ref. 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 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<sup>29</sup>. As two classes of TFs have been proposed to control the developmental potentials of effector and memory fates in a quantitative manner<sup>22,24,30</sup>, the dense overlapping regulation might contribute to the quantitative regulations during effector to memory CD8<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T-cell development. However, it stabilizes  $\beta$ -catenin to modulate Wnt-*Tcf7* signalling<sup>31,32</sup>, which promotes memory CD8<sup>+</sup> T-cell development. *Sox4* also regulates ‘stemness’ of cancer cells<sup>33</sup>, a property shared by memory T cells. Furthermore, evidence suggests that *Sox4* might be a direct regulatory target of TGF $\beta$  signalling<sup>34</sup>, which is essential for the differentiation of CD8<sup>+</sup> T cells<sup>35</sup>. 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<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T cells from *Tcf7*-deficient mice<sup>9</sup>. 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<sup>+</sup> T cells that are in transition to memory T cells, whereas the previous study used memory CD8<sup>+</sup> T cells directly from mice. Second, in our study, we overexpressed *Tcf7* for a short period (72 h) 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* versus *in vivo*, overexpression versus 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<sup>17,36,37</sup>, 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<sup>+</sup> T-cell development. This regulatory motif includes two known TFs (*Id3* and *Prdm1*) and two unknown TFs (*Sox4* and *Bach2*) in memory CD8<sup>+</sup> T-cell development. Through both overexpression and knockdown/knockout in CD8<sup>+</sup> T cells both *in vitro* and in mice, we provide extensive evidence showing that *Bach2* promotes memory CD8<sup>+</sup> T-cell development (Figs 4–6). One mechanism appears to be stimulating the induction of memory T-cell precursors as *Bach2* expression in effector T cells leads to a phenotype of CD62L<sup>hi</sup>Klrg1<sup>lo</sup>CD27<sup>+</sup> (Fig. 5b,c and Supplementary Fig. S10), which are considered as central memory T-cell precursors with high proliferative potential<sup>38,39</sup>. Recently, two studies report that *Bach2* regulates CD4<sup>+</sup> T-cell development and function by suppressing effector gene expression<sup>40,41</sup>. Our observation that *Klrg1* is suppressed by *Bach2* suggests that suppression of effector function may also be important for the development of memory CD8<sup>+</sup> 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 labelled with eFluro dye and adoptively transferred into Rag<sup>-/-</sup> mice, *Bach2*-transduced T cells diluted the fluorescent dye more extensively than the non-transduced and vector-transduced T cells (Supplementary 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<sup>37</sup>. *Id3* is known to promote cell cycle and recall proliferation of memory CD8<sup>+</sup> T cells by binding to and inhibiting E proteins<sup>15,42</sup>. 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<sup>43</sup>. *Bcl6*<sup>-/-</sup> mice exhibit a profound deficiency of memory T cells<sup>44,45</sup>, whereas in the absence of p27kip, memory CD8<sup>+</sup> T cells exhibit enhanced homeostatic and recall proliferation<sup>46</sup>. 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<sup>+</sup> T-cell development and recall proliferation through Id3- and Prdm1-mediated cell cycle control.

Development of memory CD8<sup>+</sup> T cells requires integration of multiple external and internal signals to establish a new transcriptional programme of MSGs that endows memory CD8<sup>+</sup> 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<sup>+</sup> 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<sup>+</sup> T-cell development and function.

## Methods

**Regulatory network and MRA.** A total of 386 public microarrays related to CD8<sup>+</sup> T cells from 35 independent GEO data sets (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 programme with a gcRMA method. Gene expression data was used to construct the regulatory network with the putative TFs using a reverse-engineering algorithm CLR<sup>19</sup>. Among the 1,445 putative TFs identified according to the domain predictions of protein sequences<sup>18</sup>, 1,038 were manually mapped to the latest mouse genome. To compare gene expression in different CD8<sup>+</sup> T cells, samples were grouped into naive, effector and memory based on the cell types from which the microarray analysis were done (Supplementary Table S1). Gene was considered to be expressed in CD8<sup>+</sup> T cells if the average gene expression level in one of the three groups was >8 (gcRMA values).

To identify TFs associated with memory CD8<sup>+</sup> T-cell development, we used 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<sup>+</sup> T cells and naive/effector CD8<sup>+</sup> T cells identified previously<sup>3</sup>. 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<sup>+</sup> 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<sup>+</sup> 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<sup>20</sup>. The positive candidates were analysed for enrichment of DNA-binding motifs in the promoter regions (-2,000 to -1) of the MSGs using the programme MatInspector<sup>47</sup> or differential expression among naive, effector and memory CD8<sup>+</sup> 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<sup>-/-</sup> and C57BL/6 (B6 Thy1.1<sup>+</sup>) background (2C<sup>+</sup>Rag<sup>-/-</sup>) were maintained in the animal facility at the Massachusetts Institute of Technology (MIT). These mice express the 2C TCR on CD8<sup>+</sup> T cells specific for SIYRYVGL peptide (SIY) in association with MHC class I K<sup>b</sup> molecule<sup>48</sup>. B6 and Rag2<sup>-/-</sup> 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<sup>26</sup>. For infection, mice were anesthetized and given 100 p.f.u. (sublethal dose) intranasally.

**Flow cytometry and cell sorting.** Antibodies specific for CD8 $\alpha$ , Thy1.1, Klrp1, CD62L, CD127 (IL-7R), CD27, IFN $\gamma$ , TNF $\alpha$ , 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- $\mu$ m nylon filters, respectively. Lungs were collected and ground through a cell strainer, followed by incubation with 2 ml of digestion buffer (RPMI 1640 medium containing 3 mg ml<sup>-1</sup> of collagenase A (Roche), 5% FBS and 10 mM HEPES) at 37 °C for 1 h. Red blood cells in the spleen, blood and lung were lysed with red blood cell lysis buffer (Gibco) and the cells were washed with complete RPMI. The cells were counted and 1–3  $\times$  10<sup>6</sup> cells were used for surface staining. Cells were washed twice with PBS plus 2% FBS before cytometry analysis. For intracellular staining, splenocytes were stimulated with SIY peptide for 5 h 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 labelled antibodies using an intracellular staining kit (Cytofix/Cytoperm kit; BD Biosciences) according to the manufacturer's instructions. Stained cells were analysed on either a FACSCalibur or Accuri C6 flow cytometer (BD Biosciences). About 0.5–2  $\times$  10<sup>6</sup> events were collected and analysed 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 (10 mM TrisCl (pH7.5), 50 mM NaCl and 1 mM 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 six-well plates. Cells were co-transfected with retroviral vector plasmid (4  $\mu$ g) and packing plasmids pCL-Eco (1  $\mu$ g) with 150  $\mu$ l DMEM and 15  $\mu$ l TransIT-LT1 (Mirus) according to the manufacturer'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- $\mu$ m low-protein binding membrane (Pall Life Science). Fresh viral supernatants were used for spin infection of CD8<sup>+</sup> T cells in all experiments.

For infection, cells from spleens and lymph nodes were harvested from 2C<sup>+</sup> RAG<sup>-/-</sup> mice, pooled and cultured in six-well plates in the presence of SIY peptide (1  $\mu$ g ml<sup>-1</sup>) in the complete RPMI medium (RPMI 1640 supplemented with 10% FBS, 5 mM HEPES, 2 mM glutamine, 100 U ml<sup>-1</sup> penicillin, 100  $\mu$ g ml<sup>-1</sup> streptomycin and 50  $\mu$ M  $\beta$ -mercaptoethanol (Invitrogen)). Two days later, activated 2C T cells were collected, washed and resuspended at 2  $\times$  10<sup>6</sup> cells per ml in the complete RPMI medium. Fresh retrovirus supernatants (1 ml) and 0.25 ml 2C cells with a final concentration of 5  $\mu$ g ml<sup>-1</sup> polybrene (American Bioanalytical) were added to one well of a 24-well plate and spun for 90 min at 2500 r.p.m. at 32 °C to infect T cells. Twenty four hours later, cells were collected for direct adoptive transfer, or resuspended and cultured in 3 ml fresh RPMI medium with 100 U ml<sup>-1</sup> IL-2 (eBioscience). After being cultured for 24 h, 2C T cells were analysed for GFP expression by flow cytometry and prepared for injection into mice to generate *in vivo* memory T cells or further cultured 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 5 ng ml<sup>-1</sup> IL-7 (Peprotech) for 7 days with change of fresh IL-7-supplemented medium every 2 days. Cells were analysed for memory phenotype by flow cytometry on day 7. To test the recall proliferation, *in vitro* memory 2C T cells (1  $\times$  10<sup>5</sup>) were cultured with B6 splenocytes (5  $\times$  10<sup>5</sup>) in a 12-well plate in complete RPMI medium supplemented with 1  $\mu$ g ml<sup>-1</sup> SIY peptide and 100 units per ml IL-2. The numbers and phenotype of 2C T cells were analysed by flow cytometry 4 and 6 days later. Alternatively, *in vitro* memory 2C T cells (2  $\times$  10<sup>5</sup>) were transferred to B6 recipients and challenged with WSN-SIY virus. 2C cells were analysed by flow cytometry 5 dpi.

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

**Bone marrow chimera mice.** Bone marrow cells were collected from the tibia and femur of 2C<sup>+</sup> Rag<sup>-/-</sup> mice. Stem and progenitor cells were enriched using a progenitor enrichment kit (Stemcell Technologies) according to the manufacturer's instructions. The enriched cells were cultured for 48 h in complete RPMI medium supplemented with IL-3 (30 ng ml<sup>-1</sup>), IL-6 (10 ng ml<sup>-1</sup>) and SCF (15 ng ml<sup>-1</sup>). The cells were resuspended at 2  $\times$  10<sup>6</sup> cells per ml in complete RPMI. Fresh retrovirus supernatants (600  $\mu$ l) and 400  $\mu$ l cells plus a final concentration of 6  $\mu$ g ml<sup>-1</sup> polybrene were added to one well of a 24-well plate and spun for 90 min at 2,500 r.p.m. at 32 °C. On the second day, cells were collected and washed and injected into Rag<sup>-/-</sup> mice that had been irradiated for 500 rads 4 h earlier. Eight weeks later, mice were bled to determine the reconstitution of CD8<sup>+</sup> T cells and GFP proportion by flow cytometry. Twelve weeks later, cells were collected from spleen and analysed for 2C T-cell percentage and phenotype. Splenocytes containing 5  $\times$  10<sup>4</sup> 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 analysed on 7 and 30 dpi. To generate *Bach2*<sup>-/-</sup> chimeric mice, bone marrow cells from *Bach2*<sup>-/-</sup> mice<sup>49</sup> (kind gift of Dr Kazuhiko Igarashi of Tohoku University, Japan) were directly injected into sublethally irradiated Rag2<sup>-/-</sup> mice. Eight weeks later, reconstitution of CD8<sup>+</sup> T cells were verified by

flow cytometry of peripheral blood mononuclear cells. Twelve weeks later, mice were infected with 50 p.f.u. WSN-SIY virus and 30 dpi SIY-specific CD8<sup>+</sup> T cells in various tissues were identified by H-2K<sup>b</sup> 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<sup>+</sup> 2C T cells by retrovirus transduction as described above. Transduced 2C T cells were cultured in the presence of IL-7 for 24 h and GFP<sup>+</sup> 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 manufacturer's instructions. First strand cDNA was synthesized from 1 µg total RNA using the TaqMan Reverse Transcription Reagents (ABI). Diluted cDNA (2 µl; total 200 µl) were used as template for the quantitative PCR with LightCycler480 SYBR Green and LightCycler480 machine (Roche). For each TF-transduced CD8<sup>+</sup> 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 naive, effector and memory CD8<sup>+</sup> T cells, naive 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 naive, 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 crosslinked with formaldehyde at a final concentration of 1%, sheared by sonication to 800–1,000 bp, 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), crosslinked DNA-protein complexes were digested to 400–600 bp by micrococcal nuclease (Cell Signaling Technology).

**Protein extraction and western blotting.** Proteins were extracted from transduced 2C T cells with the CellLytic Lysis Reagent (Sigma). Samples containing 20 µg 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, 1:2,000), anti-Sox4 (C-20, Santa Cruz Biotechnology, 1:500), anti-Bach2 (AP10133b, Abgent, 1:500), anti-Id3 (6-1, CalBioReagents, 1:2,500), anti-Blimp-1 (6D3, eBioscience, 1:1,000), anti-Bcl2 (BioLegend, 1:500), anti-Bcl6 (BioLegend, 1:2,000), anti-pFoxO1 (Cell Signaling Technology, 1:1,000), anti-CDK4 (Cell Signaling Technology, 1:1,000), anti-CDK6 (Cell Signaling Technology, 1:1,000), anti-Rb-p780 (Cell Signaling Technology, 1:1,000), anti-Rb-p795 (Cell Signaling Technology, 1:1,000) and anti-p27kip (Cell Signaling Technology, 1:1,000) 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, 1:2,000; anti-mouse, 1:3,000; Santa Cruz Biotechnology: anti-Rat, 1:2,000; 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  $q = p \times n/i$ , ( $p = P$  value,  $n =$  total number of tests,  $i =$  sorted rank of *P*-value).

## References

- Goldrath, A. W. & Bevan, M. J. Selecting and maintaining a diverse T-cell repertoire. *Nature* **402**, 255–261 (1999).
- Kaech, S. M., Hemby, S., Kersh, E. & Ahmed, R. Molecular and functional profiling of memory CD8 T cell differentiation. *Cell* **111**, 837–851 (2002).
- Wherry, E. J. *et al.* Molecular signature of CD8<sup>+</sup> T cell exhaustion during chronic viral infection. *Immunity* **27**, 670–684 (2007).
- Intlekofer, A. M. *et al.* Effector and memory CD8<sup>+</sup> T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* **6**, 1236–1244 (2005).
- Pearce, E. L. *et al.* Control of effector CD8<sup>+</sup> T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041–1043 (2003).
- Sullivan, B. M., Juedes, A., Szabo, S. J., von Herrath, M. & Glimcher, L. H. Antigen-driven effector CD8 T cell function regulated by T-bet. *Proc. Natl Acad. Sci. USA* **100**, 15818–15823 (2003).
- Gattinoni, L. *et al.* Wnt signaling arrests effector T cell differentiation and generates CD8<sup>+</sup> memory stem cells. *Nat. Med.* **15**, 808–813 (2009).
- Zhao, D. M. *et al.* Constitutive activation of Wnt signaling favors generation of memory CD8 T cells. *J. Immunol.* **184**, 1191–1199 (2010).
- Zhou, X. *et al.* Differentiation and persistence of memory CD8(+) T cells depend on T cell factor 1. *Immunity* **33**, 229–240 (2010).
- Schober, S. L. *et al.* 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).
- 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).
- Carlson, C. M. *et al.* Kruppel-like factor 2 regulates thymocyte and T-cell migration. *Nature* **442**, 299–302 (2006).
- Yang, C. Y. *et al.* The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8<sup>+</sup> T cell subsets. *Nat. Immunol.* **12**, 1221–1229 (2011).
- Cannarile, M. A. *et al.* Transcriptional regulator Id2 mediates CD8<sup>+</sup> T cell immunity. *Nat. Immunol.* **7**, 1317–1325 (2006).
- Ji, Y. *et al.* Repression of the DNA-binding inhibitor Id3 by Blimp-1 limits the formation of memory CD8<sup>+</sup> T cells. *Nat. Immunol.* **12**, 1230–1237 (2011).
- Kallies, A., Xin, A., Belz, G. T. & Nutt, S. L. Blimp-1 transcription factor is required for the differentiation of effector CD8(+) T cells and memory responses. *Immunity* **31**, 283–295 (2009).
- Rutishauser, R. L. *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).
- Gray, P. A. *et al.* Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* **306**, 2255–2257 (2004).
- Faith, J. J. *et al.* Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5**, e8 (2007).
- Eppig, J. T., Blake, J. A., Bult, C. J., Kadin, J. A. & Richardson, J. E. Group MGD. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Res.* **40**, D881–D886 (2012).
- Angelosanto, J. M. & Wherry, E. J. Transcription factor regulation of CD8<sup>+</sup> T-cell memory and exhaustion. *Immunol. Rev.* **236**, 167–175 (2010).
- Rutishauser, R. L. & Kaech, S. M. Generating diversity: transcriptional regulation of effector and memory CD8 T-cell differentiation. *Immunol. Rev.* **235**, 219–233 (2010).
- D'Cruz, L. M., Rubinstein, M. P. & Goldrath, A. W. Surviving the crash: transitioning from effector to memory CD8<sup>+</sup> T cell. *Semin. Immunol.* **21**, 92–98 (2009).
- Kaech, S. M. & Cui, W. Transcriptional control of effector and memory CD8(+) T cell differentiation. *Nat. Rev. Immunol.* **12**, 749–761 (2012).
- Joshi, N. S. *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).
- Shen, C. H. *et al.* 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).
- Hettmann, T., Opferman, J. T., Leiden, J. M. & Ashton-Rickardt, P. G. A critical role for NF-kappaB transcription factors in the development of CD8<sup>+</sup> memory-phenotype T cells. *Immunol. Lett.* **85**, 297–300 (2003).
- Teixeiro, E. *et al.* Different T cell receptor signals determine CD8<sup>+</sup> memory versus effector development. *Science* **323**, 502–505 (2009).
- Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461 (2007).
- Kaech, S. M. & Wherry, E. J. Heterogeneity and cell-fate decisions in effector and memory CD8<sup>+</sup> T cell differentiation during viral infection. *Immunity* **27**, 393–405 (2007).
- Bernard, P. & Harley, V. R. 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).
- 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).
- Ikushima, H. *et al.* Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell Stem Cell* **5**, 504–514 (2009).
- 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).
- Sanjabi, S., Mosaheb, M. M. & Flavell, R. A. Opposing effects of TGF-beta and IL-15 cytokines control the number of short-lived effector CD8<sup>+</sup> 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, V. R. *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. USA* **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, A. L. *et al.* BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* **13**, 199–212 (2000).
44. Crotty, S., Johnston, R. J. & Schoenberger, S. P. 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. *et al.* 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, H. N. & Kranz, D. M. 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, R. Q., Flavell, R. A. & 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, J. C. & Cameron, E. R. 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, P. A. IL-12-programmed long-term CD8 + T cell responses require STAT4. *J. Immunol.* **177**, 7618–7625 (2006).
54. Rao, R. R., Li, Q., Odunsi, K. & Shrikant, P. A. 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, R. A. 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, K. J., Yoon, H., Ahmed, R. & Boss, J. M. 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, H. J. & Nel, A. E. 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).

### Acknowledgements

We 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<sup>-/-</sup> 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 J.C. conceived and designed the study. G.H. carried out all computations and experimentations. G.H. and J.C. performed data analysis and drafted the paper. All authors read and approved the final manuscript.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Hu, G. and Chen, J. A genome-wide regulatory network identifies key transcription factors for memory CD8<sup>+</sup> T-cell development. *Nat. Commun.* 4:2830 doi: 10.1038/ncomms3830 (2013).