nature immunology

ARTICLES

https://doi.org/10.1038/s41590-018-0046-x

Transcription factor Hoxb5 reprograms B cells into functional T lymphocytes

Mengyun Zhang[1,2,10], Yong Dong[1,2,10], Fangxiao Hu[1,10], Dan Yang[1], Qianhao Zhao[1], Cui Lv[1], Ying Wang[1,2], Chengxiang Xia[1,2], Qitong Weng[1,2], Xiaofei Liu[1], Chen Li[3], Peiqing Zhou[1,2], Tongjie Wang[1], Yuxian Guan[1], Rongqun Guo[1,2], Lijuan Liu[1], Yang Geng[1], Hongling Wu[1], Juan Du[1], Zheng Hu[4], Sheng Xu[5], Jiekai Chen[1,2], Aibin He[3], Bing Liu[6], Demin Wang[7,8], Yong-Guang Yang[4,9] and Jinyong Wang[1,2*]

Deletion of master regulators of the B cell lineage reprograms B cells into T cells. Here we found that the transcription factor Hoxb5, which is expressed in uncommitted hematopoietic progenitor cells but is not present in cells committed to the B cell or T cell lineage, was able to reprogram pro-pre-B cells into functional early T cell lineage progenitors. This reprogramming started in the bone marrow and was completed in the thymus and gave rise to T lymphocytes with transcriptomes, hierarchical differentiation, tissue distribution and immunological functions that closely resembled those of their natural counterparts. Hoxb5 repressed B cell 'master genes', activated regulators of T cells and regulated crucial chromatin modifiers in pro-pre-B cells and ultimately drove the B cell fate–to–T cell fate conversion. Our results provide a de novo paradigm for the generation of functional T cells through reprogramming in vivo.

Cell fate is strictly controlled during development, which ensures that each cell contributes to the function of the organism in a harmonious fashion. Multi- or pluripotent cells can be directed to differentiate into a specific cell type, and somatic cells can trans-differentiate from one lineage to another by the expression of specific transcription factors. Master regulators that can mediate hematopoietic-lineage conversions have been identified. Thus, Gata1 converts monocytic precursor cells into erythroid–megakaryocytic cells and eosinophils[1–3], and Cebpα converts B cells into macrophages[4]; deletion of Pax5 converts B cells into uncommitted hematopoietic progenitor cells[5,6]; Gata3 converts T lymphocytes into mast cells[7]; Cebpα and Spi1 convert T lymphocytes into macrophages and dendritic cells[8]; and deletion of Bcl11b converts T lymphocytes into natural killer–like cells[9]. Attempts to convert B cells into T cells by silencing 'master genes' of the B cell lineage have had limited success, in that it has not been possible to reconstitute the entire T cell lineage functionally and, in some instances, the manipulations have increased cancer risk[5,6,10,11]. In aggregate, these studies indicate that the hematopoietic cell fate can be manipulated genetically.

Hematopoietic stem cells (HSCs) and multipotent progenitors (MPPs) differentiate into various hematopoietic cell types through the activation of specific gene-regulatory networks[12,13]. The transcription factor Hoxb5 is expressed specifically in HSCs[14], although the entire gene cluster encoding the Hoxb family of proteins seems to be dispensable for hematopoiesis[15]. Here we found that expression of Hoxb5 alone in pro-pre-B cells, followed by transplantation of the pro-pre-B cells into sublethally irradiated recipient mice, produced early T cell lineage progenitors (ETPs) in the bone marrow (BM) and ultimately regenerated a full complement of functional T lymphocytes, whose transcriptomes, hierarchical differentiation, tissue distribution and immunological functions closely resembled those of natural T lymphocytes. To our knowledge, this is the first report of a procedure for generating fully functional T lymphocytes in vivo by lineage conversion.

Results

Ectopic expression of 15 transcription factors reprograms B cells into T cells. First, we investigated whether hematopoietic cells could be converted from one lineage to another (trans-differentiation) or could be converted back to uncommitted multipotent cells (de-differentiation) by transcription factors expressed 'preferentially' in HSCs and MPPs, not in lineage-committed cells. To identify transcription factors expressed 'preferentially' in HSCs and MPPs rather than in cells commited to the hematopoietic lineage, we sorted Lin$^-$CD48$^-$c-kit$^+$Sca-1$^+$CD150$^+$ HSCs, Lin$^-$CD48$^-$c-kit$^+$Sca-1$^+$CD150$^-$ MPPs, Ter119$^-$CD3$^-$CD19$^-$Mac1$^+$ myeloid cells, Ter119$^-$CD19$^-$Mac1$^-$CD3$^+$ T lymphoid cells and Ter119$^-$Mac1$^-$CD3$^-$CD19$^+$ B lymphoid cells from nucleated BM cells of 8-week-old female C57BL/6 mice and conducted gene-expression analysis by RNA-based next-generation sequencing (RNA-Seq). Genes were designated as being expressed 'preferentially' in HSCs and MPPs if they demonstrated over twofold higher expression in HSCs and MPPs than in lineage-committed cells ($P<0.05$ (two-sided-independent Student's t-test)). The genes that met these criteria were screened for a match in an online database of transcription factors (http://genome.gsc.riken.jp/TFdb/tf_list.html); this screen identified 15 candidate

[1]CAS Key Laboratory of Regenerative Biology and Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Institute of Molecular Medicine, Beijing Key Laboratory of Cardiometabolic Molecular Medicine, Peking University, Beijing, China. [4]The First Hospital and Institute of Immunology, Jilin University, Changchun, China. [5]National Key Laboratory of Medical Immunology & Institute of Immunology, Second Military Medical University, Shanghai, China. [6]State Key Laboratory of Proteomics, Translational Medicine Center of Stem Cells, 307-Ivy Translational Medicine Center, Laboratory of Oncology, Affiliated Hospital, Academy of Military Medical Sciences, Beijing, China. [7]Biomedical Research Center of South China, College of Life Sciences, Fujian Normal University, Fuzhou, Fujian, China. [8]Blood Research Institute, Blood Center of Wisconsin, Milwaukee, WI, USA. [9]Columbia Center for Translational Immunology, Columbia University College of Physicians and Surgeons, New York, NY, USA. [10]These authors contributed equally: Mengyun Zhang, Yong Dong and Fangxiao Hu. *e-mail: wang_jinyong@gibh.ac.cn

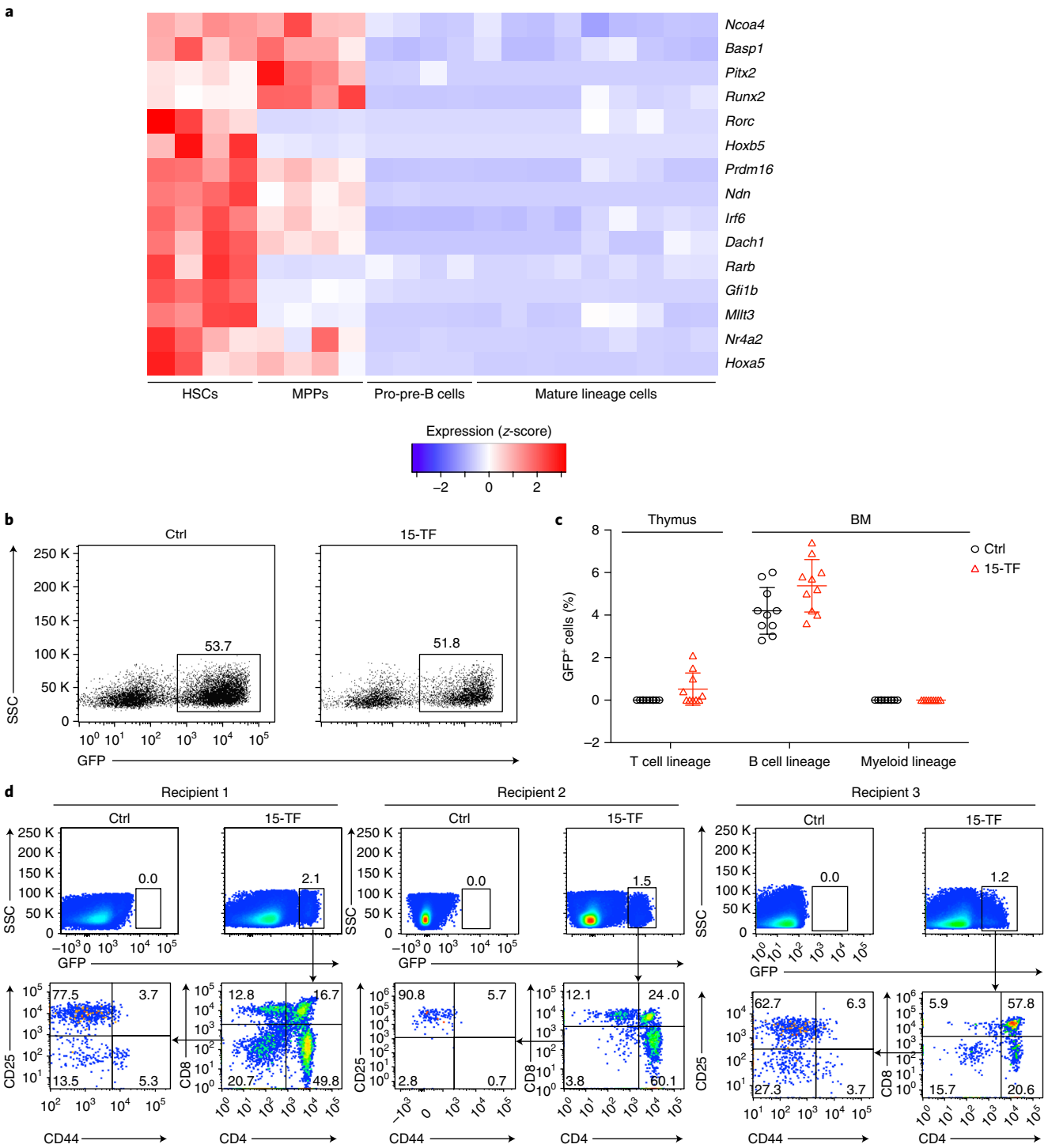

Fig. 1 | Screening for transcription factors involved in the B cell–to–T cell conversion. a, RNA-Seq analysis of genes encoding 15 transcription factors (right margin) in HSCs, MPPs, pro-pre-B cells, and mature T cells or B cells or myeloid cells (Mature lineage cells) (below plot; 1,000 cells per type; one biological replicate per column); results were screened by the principle of pairwise comparison (change in expression, over twofold; $P < 0.05$ (two-sided-independent Student's *t* test)) and are presented as z-score values (obtained by conversion of FPKM (fragments per kilobase million) values; key). **b**, Flow cytometry of Ter119$^-$Mac1$^-$CD3$^-$CD4$^-$CD8$^-$B220$^+$CD19$^+$CD93$^+$IgM$^-$ pro-pre-B cells transduced with retrovirus containing an empty cassette (control (Ctrl)) or a cassette encoding 15-TF (above plots). Numbers above outlined areas indicate percent GFP$^+$ (retrovirus-expressing) cells. **c**, Frequency of GFP$^+$ cells in the T cell lineage (thymus) or the B cell lineage and myeloid lineage (BM) (above plot and horizontal axis) of recipient mice ($n = 10$ per group) 4 weeks after transplantation of pro-pre-B cells transduced as in **b** (key). **d**, Flow cytometry of GFP$^+$ lymphocytes obtained from the thymus of recipient mice ($n = 3$ per group; at top above plots) 4 weeks after transplantation of pro-pre-B cells transduced as in **b** (above plots). Numbers above outlined areas indicate percent GFP$^+$ cells (top row); numbers in quadrants indicate percent cells in each (bottom row). SSC, side scatter. Each symbol (**c**) represents an individual host mouse; small horizontal lines indicate the mean (±s.d.). Data are representative of one experiment with $n = 4$ biological replicates (HSCs, MPPs and pro-pre-B cells) or $n = 9$ biological replicates (mature lineage cells) (**a**), four independent experiments (**b**) or two independent experiments (**d**) or are pooled from two independent experiments (**c**).

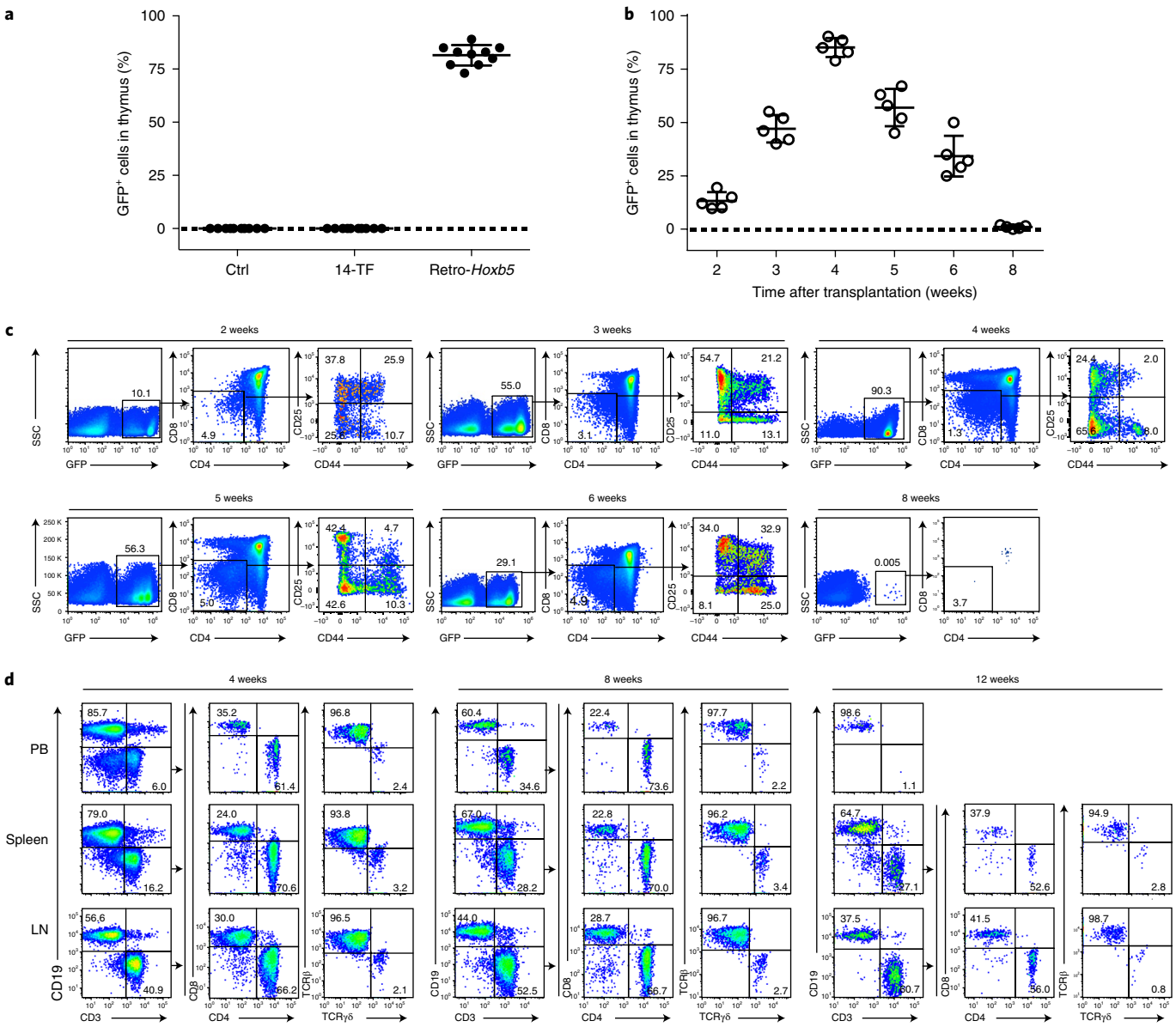

Fig. 2 | Expression of retro-*Hoxb5* in pro-pre-B cells converts B cells into T lymphocytes in vivo. a, Frequency of Ter119$^-$Mac1$^-$CD19$^-$ GFP$^+$ cells obtained from the thymus of sublethally irradiated congenic recipient mice ($n=10$ per group) 4 weeks after transplantation of pro-pre-B cells transduced with retrovirus expressing GFP only (Ctrl), 14-TF or retro-*Hoxb5* (horizontal axis), assessed by flow cytometry. **b**, Frequency of GFP$^+$ cells obtained from the thymus of recipient mice ($n=5$) at various times (horizontal axis) after transplantation of pro-pre-B cells transduced with retro-*Hoxb5*. **c**, Flow cytometry cells in the thymus of recipient mice at 2, 3, 4, 5, 6 and 8 weeks (horizontal axis) after transplantation of pro-pre-B cells transduced with retro-*Hoxb5*. Numbers above outlined areas indicate percent GFP$^+$ cells (left plot of each trio); numbers in outlined areas indicate percent CD8$^-$CD4$^-$ cells (middle plot of each trio); and numbers in quadrants indicate percent cells in each (right plot of each trio). **d**, Flow cytometry of GFP$^+$ T lymphocytes (gated from the GFP$^+$Ter119$^-$Mac1$^-$ population) obtained from the PB, spleen and LNs (left margin) of recipient mice ($n=5$) at 4, 8 and 12 weeks (above plots) after transplantation of pro-pre-B cells transduced with retro-*Hoxb5*. Numbers in quadrants indicate percent cells in each. Each symbol (**a**,**b**) represents an individual host mouse; small horizontal lines indicate the mean (±s.d.). Data are pooled from three independent experiments (**a**) or two independent experiments (**b**) or are representative of two independent experiments (**c**,**d**).

transcription factors expressed 'preferentially' in HSCs and MPPs but not in lineage-committed cells (Fig. 1a).

The genes encoding each of the 15 transcription factors identified above were cloned into a retroviral expression cassette containing sequence encoding green fluorescent protein (GFP), and a retroviral mixture containing clones for all 15 transcription factors (15-TF) was transduced into sorted Ter119$^-$Mac1$^-$CD3$^-$CD4$^-$CD8$^-$B220$^+$CD19$^+$CD93$^+$IgM$^-$ pro-pre-B cells (Supplementary Fig. 1a). We used pro-pre-B cells as targets for reprogramming because they carry genomic rearrangements of genes encoding immunoglobulin heavy-chain variable, diversity and joining regions (V(D)J rearrangements) that serve as natural 'genetic barcodes', and they have weak epigenetic barriers for reprogramming[16]. Viral titers in the mixture were adjusted to the same multiplicity of infection (0.69) to yield a transduction rate for each clone of ~50% (Fig. 1b). 5×10^6 GFP$^+$ (retrovirus-expressing) pro-pre-B cells (from 4-week-old female mice on the C57BL/6 background) transduced with the 15-TF retroviral mixture were transplanted into sublethally irradiated (6.5 Gy)

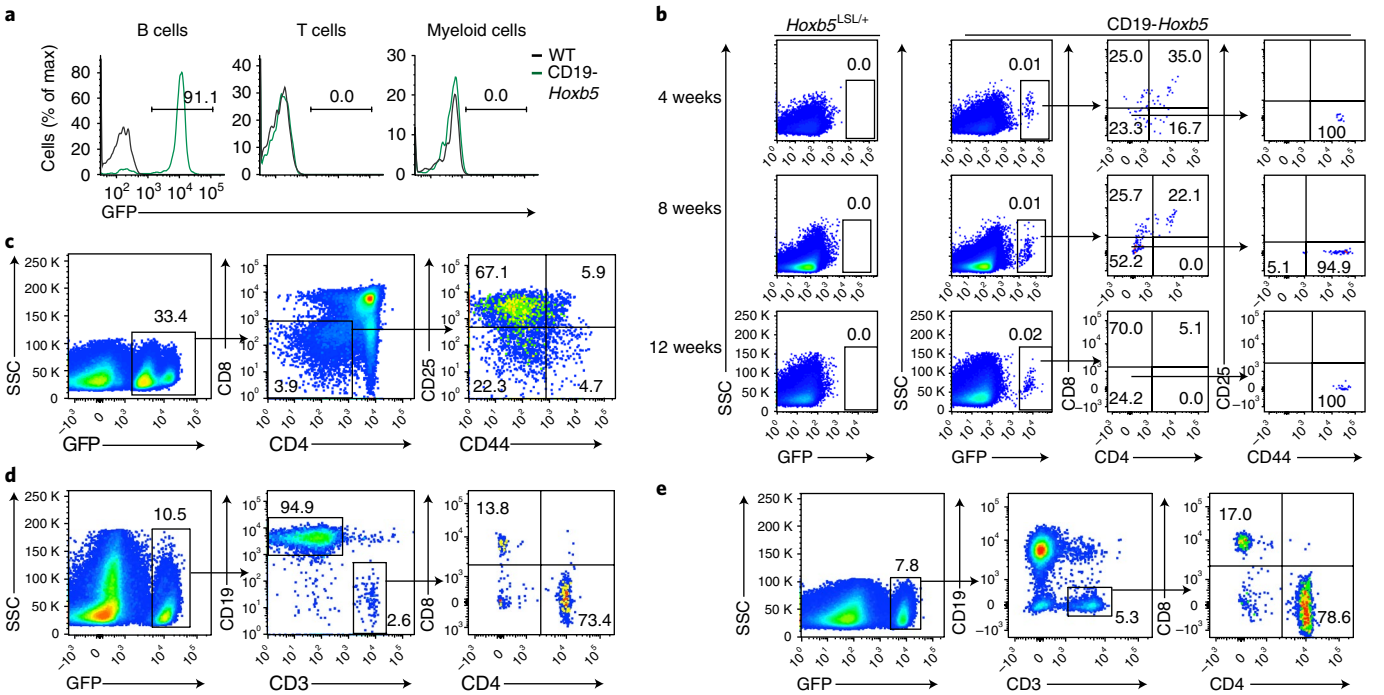

Fig. 3 | Conversion of B lymphocytes into T lymphocytes in the CD19-*Hoxb5* model. a, Ectopic expression of Hoxb5 (reported by GFP) in CD19$^+$ B cells, CD3$^+$ T cells and Mac1$^+$ myeloid cells (above plots) from the PB of wild-type (WT) or CD19-*Hoxb5* mice (key), analyzed by flow cytometry. Numbers above bracketed lines indicate percent GFP$^+$ cells. **b**, Flow cytometry of Ter119$^-$Mac1$^-$CD19$^-$ lymphocytes from the thymus of 4-, 8- and 12-week-old (left margin) CD19-*Hoxb5* mice and their *Hoxb5*$^{LSL/+}$ (control) littermates, under homeostasis. Numbers above outlined areas indicate percent GFP$^+$ cells (far and middle left) analyzed further at right; numbers in quadrants indicate percent cells in each (far and middle right). **c**–**e**, Flow cytometry of iDN cells (gated from the Ter119$^-$Mac1$^-$CD19$^-$ population) obtained from the thymus (**c**) or iT cells (gated from the Ter119$^-$Mac1$^-$ population) obtained from the spleen (**d**) and LNs (**e**) of recipient mice 4 weeks after transplantation of 3×10^6 CD19-*Hoxb5* pro-pre-B cells. Numbers in or adjacent to outlined areas or quadrants indicate percent cells in each. Data are representative of two independent experiments (**a**,**b**) or three independent experiments (**c**–**e**).

CD45.2$^+$ C57BL/6 mice (15-TF mice), and the frequency of GFP$^+$ cells in thymus and BM of recipients was analyzed by flow cytometry 4 weeks after transplantation (Supplementary Fig. 1b). A substantial proportion of GFP$^+$ thymocytes was observed in 50% of the 15-TF mice, but no GFP$^+$ thymocytes were observed in recipient mice given transplantation of pro-pre-B cells transduced with control virus containing empty vector (Fig. 1c). Flow cytometry of the GFP$^+$ thymocytes revealed the presence of CD44$^+$CD25$^-$ CD4$^-$CD8$^-$ (double-negative stage 1 (DN1)) thymocytes, CD44$^+$CD25$^+$ DN2 thymocytes, CD44$^-$CD25$^+$ DN3 thymocytes, CD44$^-$CD25$^-$ DN4 thymocytes, CD4$^+$CD8$^+$ (double-positive (DP)) thymocytes, CD4$^+$CD8$^-$ (CD4$^+$ single-positive (CD4SP)) T lymphocytes and CD4$^-$CD8$^+$ (CD8$^+$ single-positive (CD8SP)) T lymphocytes (Fig. 1d).

Flow cytometry of the B cell compartment of host mice (CD45.1$^+$ C57BL/6) that received pro-pre-B cells transduced with the 15-TF mixture showed that the frequency of GFP$^+$ CD45.2$^+$ B220$^+$CD19$^+$IgD^{lo}IgM^{hi} immature B cells in the spleen and peripheral blood (PB) (Supplementary Fig. 1c,d) and GFP$^+$ CD45.2$^+$ B220$^+$CD19$^+$CD93$^+$IgM$^-$ pro-pre-B cells in the BM (Supplementary Fig. 1e,f) was much lower, while the frequency of GFP$^+$ CD45.2$^+$ B220$^+$CD19$^+$IgD^{hi}IgM^{lo} mature B cells in the BM, spleen and PB was much higher (Supplementary Fig. 1c,d), than that of their GFP$^-$ CD45.2$^+$ counterparts; this suggested that expression of the 15-TF viral mixture altered B cell development. These results indicated that forced expression of these transcription factors was able to reprogram committed progenitors of B cells into distinct subsets of T lymphocytes.

Retroviral expression of Hoxb5 reprograms B cells into T cells. To identify which of the 15 candidate transcription factors induced the B cell–to–T cell conversion noted above, we sorted GFP$^+$ thymocytes from 15-TF mice and used single-cell PCR to identify retroviral sequences integrated into the genomic DNA of these cells. To avoid amplification of endogenous genes, we ensured that the PCR primer pairs flanked the introns of the genes encoding the 15 candidate transcription factors (Supplementary Table 1). All 70 GFP$^+$ thymocytes analyzed contained retrovirus-derived ectopic copies of *Hoxb5* (Supplementary Table 2). Furthermore, *Hoxb5*-expressing retrovirus (retro-*Hoxb5*), a mixture of retroviral clones expressing the genes encoding the 14 other transcription factors but lacking *Hoxb5* (14-TF), or control virus expressing GFP alone was transduced into pro-pre-B cells that were retro-orbitally transferred into sublethally irradiated congenic mice (CD45.2$^+$ mice on the C57BL/6 background). In all recipients of pro-pre-B cells expressing retro-*Hoxb5* (retro-*Hoxb5* mice), over 65% of single thymic nucleated cells were Ter119$^-$Mac1$^-$CD19$^-$GFP$^+$, while Ter119$^-$Mac1$^-$CD19$^-$GFP$^+$ single thymic nucleated cells were not detected in mice given transfer of pro-pre-B cells expressing 14-TF or control virus (Fig. 2a). The proportion of GFP$^+$Ter119$^-$Mac1$^-$CD19$^-$ cells increased from about 10% at 2 weeks after transplantation to > 40% at 3–4 weeks after transplantation and then gradually decreased from 5 weeks to 8 weeks after transplantation, until they were no longer detected (Fig. 2b,c). GFP$^+$ CD4$^+$ or CD8$^+$ mature T cells were detected in the PB, spleen and lymph nodes (LNs) of retro-*Hoxb5* mice at 4 and 8 weeks after transplantation, followed by a decrease in and disappearance from the PB 12 weeks after transplantation (Fig. 2d). GFP$^+$ T cells included both TCRβ$^+$ cells and TCRγδ$^+$ cells (Fig. 2d).

To address the effect of constitutive expression of Hoxb5 on T cell development, we crossed ROSA26-*Hoxb5* transgenic mice

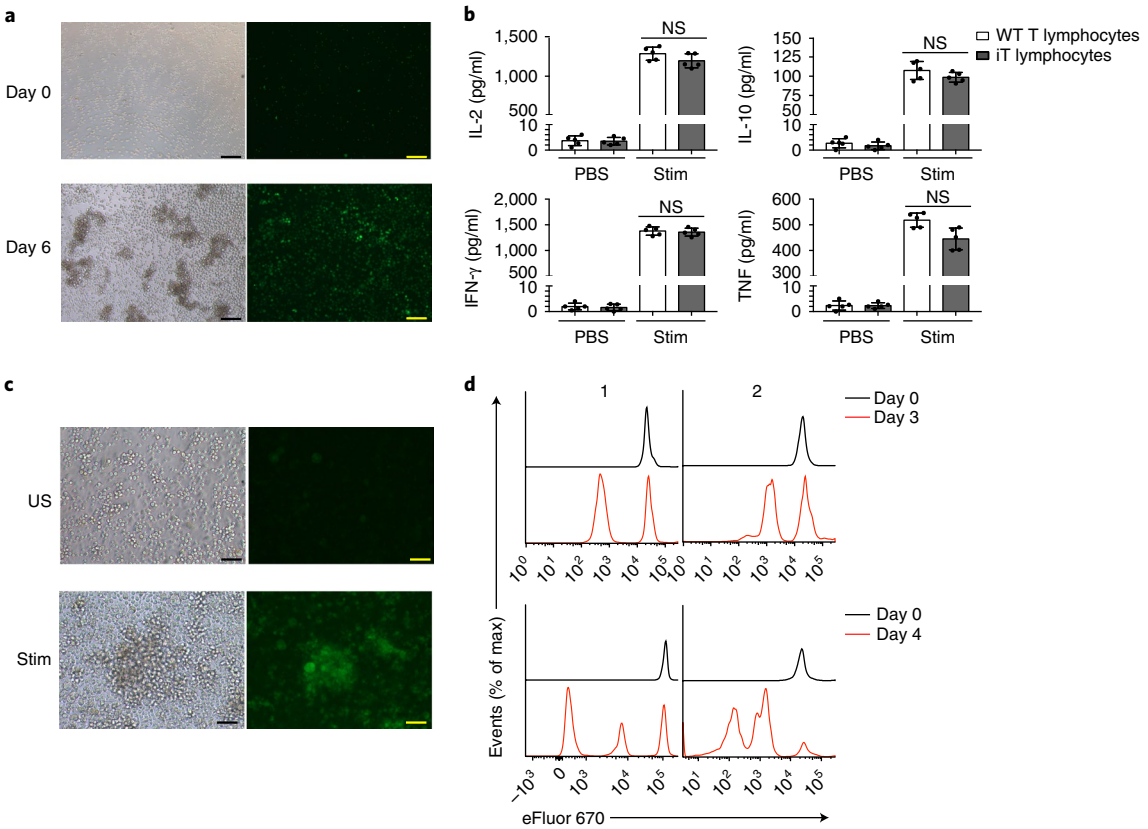

Fig. 4 | Immunological function of Hoxb5-induced iT lymphocytes. a, Microscopy assessing the morphology of iT cells before (day 0) and 6 d after (left margin) treatment with mAb to CD3 and mAb to CD28 (to stimulate cellular proliferation). Scale bars, 100 μM. **b**, ELISA of IL-2, IL-10, IFN-γ and TNF in the supernatant of wild-type splenic T lymphocytes or iT lymphocytes (key) treated with PBS or stimulated (Stim) for 6 d in vitro (0.5×10^6 cells per well; 96-well plate) with mAb to CD3 (50 μg/ml) and mAb to CD28 (5 μg/ml) (horizontal axis). NS, not significant ($P>0.05$) (two-sided-independent *t*-test). **c**, Microscopy assessing the morphology of iT cell blasts left unstimulated (US) or 7 d after stimulation (Stim) with inactivated allogeneic splenocytes (from a BALB/c mouse). Scale bars, 100 μm. **d**, Flow cytometry of proliferative CD4$^+$ iT lymphocytes in mixed-lymphocyte-reaction assay of splenic iT lymphocytes (0.5×10^6) sorted from *Hoxb5-Rag1*$^{-/-}$ mice with (1) or without (2) allogeneic skin transplantation (above plots), stained with the cell-proliferation dye eFluor 670 and then cultured for 3–4 d (key) with 0.5×10^6 inactivated allogeneic splenocytes (BALB/c) in individual wells of 96-well plates, or assessed without culture (Day 0). Each symbol (**b**) represents an individual host mouse. Data are representative of three independent experiments (**a**,**c**,**d**) or are pooled from two independent experiments (**b**; mean ± s.d. of $n=5$ biological replicates).

(*Hoxb5*$^{LSL/+}$) with Vav-Cre mice to generate *Hoxb5*$^{LSL/+}$Vav-Cre mice (Supplementary Fig. 2a), which constitutively express *Hoxb5* in all hematopoietic cells, including T lymphocytes. *Hoxb5*$^{LSL/+}$Vav-Cre mice had an abundance of ETPs in BM and thymus, DN and DP cells in thymus, and CD4$^+$ and CD8$^+$ T lymphocytes in the PB, spleen, LNs and BM comparable to that in *Hoxb5*$^{LSL/+}$ (control) mice (Supplementary Fig. 2d–f). To further assess the effect of *Hoxb5* expression on hematopoiesis, we retro-orbitally injected 0.5×10^6 total BM cells from either *Hoxb5*$^{LSL/+}$Vav-Cre (CD45.2$^+$) mice or *Hoxb5*$^{-/-}$ (CD45.2$^+$) mice, together with an equal number of total BM competitor cells from wild-type (CD45.1$^+$) mice, into lethally irradiated (9.0 Gy) CD45.1$^+$ recipient mice. *Hoxb5*$^{LSL/+}$Vav-Cre HSCs (Supplementary Fig. 3a–f) and *Hoxb5*$^{-/-}$ HSCs (Supplementary Fig. 3g–l) differentiated into multi-lineage cells with patterns similar to those of their wild-type competitors and over time frames similar to those of their wild-type competitors. Thus, overexpression of Hoxb5 had a minimal effect on hematopoiesis.

To confirm the B cell origin of the GFP$^+$ T cells, we assessed the immunoglobulin heavy-chain V(D)J rearrangements[17,18] in GFP$^+$ splenic T cells from retro-*Hoxb5* mice. Single-cell PCR analysis with primers that amplified gene segments encoding immunoglobulin heavy-chain V(D)J and T cell antigen receptor β-chain (TCRβ) V(D)J rearrangements (Supplementary Table 3) indicated that individual splenic GFP$^+$ T lymphocytes simultaneously contained B cell antigen receptor immunoglobulin heavy-chain V(D)J rearrangements and TCRβ V(D)J rearrangements (Supplementary Table 4 and Supplementary Fig. 4a), indicative of their B cell origin. In addition, the distinct TCRβ V(D)J sequences of individual GFP$^+$ T cells indicated they were polyclonal (Supplementary Table 4). These results indicated that the expression of Hoxb5 was sufficient to convert pro-pre-B cells into T lymphocytes (induced T cells (iT cells)) in vivo.

B cell–specific expression of Hoxb5 converts B cells into T cells. To exclude the possibility that retroviral transduction and integration had a role in the B cell–to–T cell conversion noted above, we crossed ROSA26-*Hoxb5* transgenic (*Hoxb5*$^{LSL/+}$) mice with CD19-Cre mice to generate *Hoxb5*$^{LSL/+}$CD19-Cre mice (CD19-*Hoxb5* mice, on a CD45.2$^+$ C57BL/6 background) (Supplementary Fig. 2a), in which a GFP reporter encoded in the *Hoxb5* cassette can be used to track B cell lineage–specific expression of Hoxb5. Approximately 90% of CD19$^+$ B lymphocytes in the blood of CD19-*Hoxb5* mice expressed GFP (Hoxb5), while CD3$^+$ T lymphocytes or Mac1$^+$ myeloid cells in the PB were GFP$^-$ (Fig. 3a). However, GFP$^+$ T lymphocytes were detected in the thymus of 4-, 8- and 12-week-old CD19-*Hoxb5* mice (Fig. 3b). Furthermore, B cells from CD19-*Hoxb5* mice showed

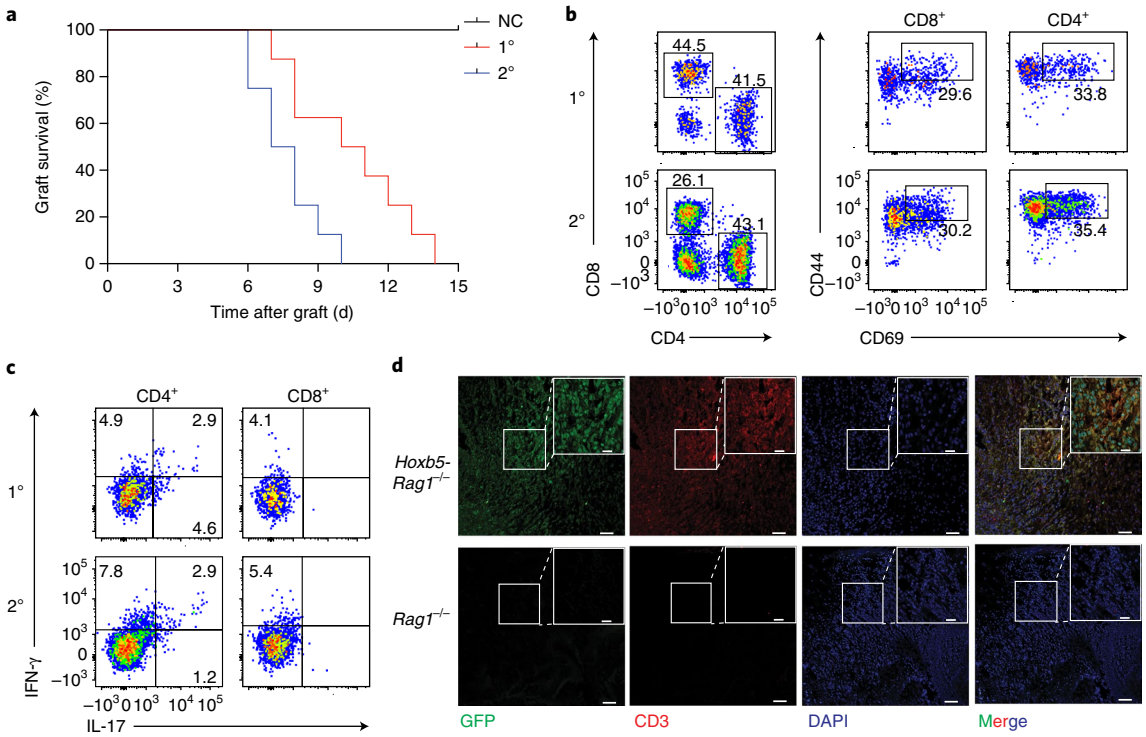

Fig. 5 | Hoxb5-induced iT lymphocytes show normal immunological function in vivo. a, Survival of allogeneic skin grafts from *Rag1*$^{-/-}$ mice (NC) and primary (1°) or secondary (2°) allografts from *Hoxb5-Rag1*$^{-/-}$ mice ($n = 8$ mice per group), presented as Kaplan-Meier curves. Median survival: 10.5 d (primary) or 7.5 d (secondary). $P < 0.001$, (log-rank test). **b**, Flow cytometry of Hoxb5-induced iT lymphocytes in rejected primary (day 10) and secondary (day 6) allogeneic skin grafts, assessed as single cells harvested by collagenase I digestion, identifying activated iT cells as CD45.2$^+$Ter119$^-$Mac1$^-$CD19$^-$CD4$^+$CD44^{hi}CD69$^+$ or CD45.2$^+$Ter119$^-$Mac1$^-$CD19$^-$CD8$^+$CD44^{hi}CD69$^+$. Numbers adjacent to outlined areas indicate percent cells in each. **c**, Intracellular staining of IFN-γ and IL-17 in activated CD45.2$^+$ CD4$^+$ or CD8$^+$ iT lymphocytes (above plots) from rejected primary (day 10) and secondary (day 6) allogeneic skin grafts. Numbers in quadrants indicate percent cells in each. **d**, Immunofluorescence staining of GFP$^+$ CD3$^+$ iT lymphocytes in tissue sections of allogeneic skin grafts from *Rag1*$^{-/-}$ and *Hoxb5-Rag1*$^{-/-}$ mice (left margin); areas outlined in main images are enlarged in insets. Scale bars, 20 μm (main images) or 50 μm (insets). Data are representative of two independent experiments (**a**) or three independent experiments (**b**–**d**).

a developmental pattern comparable to that of B cells from their *Hoxb5*$^{LSL/+}$ (control) littermates (Supplementary Fig. 2b,c), which indicated that induction of *Hoxb5* at the pro-pre-B cell stage had a minimal effect on B cell development. GFP$^+$ T lymphocytes were much less abundant in CD19-*Hoxb5* mice than in retro-*Hoxb5* mice (Fig. 3b); this might have reflected the competition of endogenous T lymphocytes generated from HSCs under homeostasis. To test that possibility, we isolated GFP$^+$ pro-pre-B cells from the BM of CD19-*Hoxb5* mice and transplanted the cells into sublethally irradiated wild-type congenic recipient mice. At 4 weeks after transplantation, >30% of T lymphocytes in thymus were GFP$^+$ (Fig. 3c). In addition, >10% of T lymphocytes in the spleen (Fig. 3d) and >7% of T lymphocytes in LNs (Fig. 3e) were GFP$^+$. Furthermore, splenic GFP$^+$ CD3$^+$ T cells had both immunoglobulin heavy-chain V(D)J rearrangements and TCRβ V(D)J rearrangements (Supplementary Table 4 and Supplementary Fig. 4b). Together these results indicated that transgenic expression of Hoxb5 reprogrammed B cells into T cells in vivo.

Hoxb5-induced T cells are functionally equivalent to wild-type T cells. To evaluate the function of the mature GFP$^+$ induced T lymphocytes (iT lymphocytes), we transplanted 3×10^6 retro-*Hoxb5* pro-pre-B cells into sublethally irradiated (3.5 Gy) *Rag1*$^{-/-}$ recipient mice, which lack B cells and T cells. GFP$^+$ iT lymphocytes were detected in the PB of *Rag1*$^{-/-}$ recipient mice at 3 weeks after transfer (Supplementary Fig. 5a). Splenic GFP$^+$ T cells isolated from *Rag1*$^{-/-}$ mice at 4 weeks after transplantation of retro-*Hoxb5* pro-pre-B cells (*Hoxb5-Rag1*$^{-/-}$ mice) and then incubated for 6 d in vitro with monoclonal antibody (mAb) to the invariant signaling protein CD3 and mAb to the co-receptor CD28 produced interleukin 2 (IL-2), IL-10, interferon-γ (IFN-γ) and tumor-necrosis factor (TNF) in amounts similar to those produced by wild-type splenic T cells (Fig. 4a,b). Furthermore, incubation of *Hoxb5-Rag1*$^{-/-}$ splenocytes with inactivated allogeneic splenocytes from BALB/c mice induced rapid proliferation of GFP$^+$ CD4$^+$ T cells from *Hoxb5-Rag1*$^{-/-}$ mice (Fig. 4c,d).

To investigate the function of Hoxb5-induced T cells in vivo, we grafted skin from BALB/c mice into C57BL/6 *Rag1*$^{-/-}$ mice at 3 weeks after the transfer of CD19-*Hoxb5* pro-pre-B cells. The allogeneic skin grafts were rapidly rejected by the resultant *Hoxb5-Rag1*$^{-/-}$ mice at about 10 d after transplantation, as indicated by bulged, ulcerative and necrotic lesions at the graft site (Supplementary Fig. 5b).GFP$^+$CD3$^+$ iT lymphocytes infiltrated the dermis of the allogeneic skin (Fig. 5d). To evaluate the memory response of the iT cells, we performed an assay of a secondary allogeneic skin graft 8 weeks after the primary transplantation of allogeneic skin (time interval between primary skin graft and secondary skin graft, 8 weeks). The median survival of the secondary allogeneic skin grafts (7.5 d) was significantly shorter than that of the primary allogeneic skin grafts (10.5 d) Fig. 5a). Flow cytometry indicated the presence of CD44^{hi}CD69$^+$ CD8$^+$ or CD4$^+$ T lymphocytes (Fig. 5b), as well as IL-17$^+$ and IFN-γ$^+$ CD4$^+$ T lymphocytes and IFN-γ$^+$ CD8$^+$ T lymphocytes, in the rejected primary and secondary grafts (Fig. 5c). These results indicated that the Hoxb5-induced T lymphocytes were activated in vitro and mediated rejection of allogeneic skin

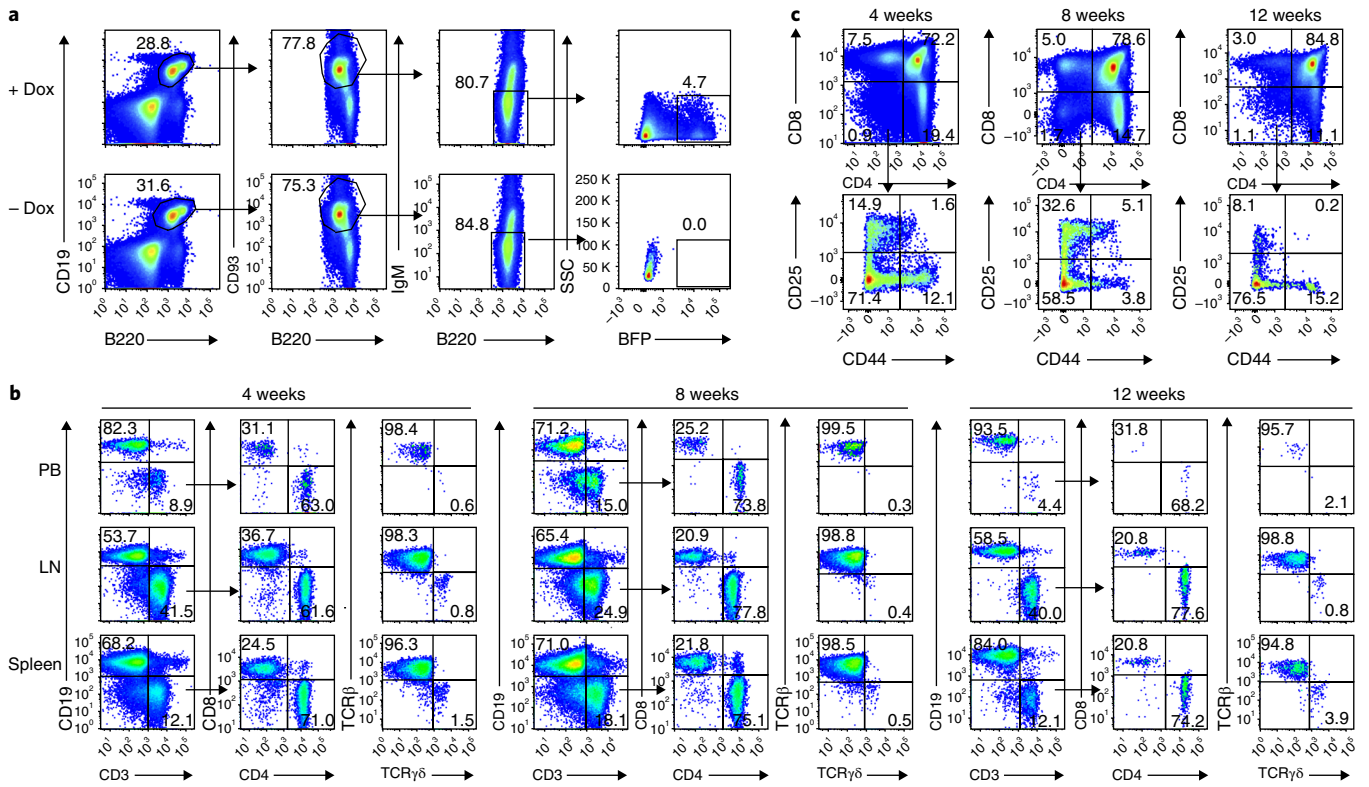

Fig. 6 | Transient expression of Hoxb5 reprograms B cells into T lymphocytes in the Tet-*Hoxb5* model. a, Flow cytometry of cells from the BM of Tet-*Hoxb5* mice maintained for 1 week on drinking water with (+Dox) or without (– Dox) doxycycline (1 mg/ml), identifying BFP$^+$ B220$^+$CD19$^+$CD93$^+$IgM$^-$ pro-pre-B cells. **b**,**c**, Flow cytometry of iT lymphocytes (gated from the CD45.2$^+$Ter119$^-$Mac1$^-$ population) obtained from the PB, spleen and LNs (**b**) or iDN cells (gated from CD45.2$^+$Ter119$^-$Mac1$^-$CD19$^-$ population) obtained from the thymus (**c**) of recipient mice at 4, 8 and 12 weeks (above plots) after transplantation of Tet-*Hoxb5* pro-pre-B cells. Numbers adjacent to outlined areas or in quadrants indicate percent cells in each throughout. Data are representative of three independent experiments (**a**) or two independent experiments (**b**,**c**).

grafts and sustained immunological memory, suggestive of a typical adaptive immune response.

Transient expression of Hoxb5 reprograms B lymphocytes into T lymphocytes. Next we constructed and characterized an inducible transgenic model in which ectopic expression of Hoxb5 can be induced by doxycycline (Tet-*Hoxb5* (CD45.2$^+$) mice, on the C57BL/6 background). The *Hoxb5* expression cassette in the Tet-*Hoxb5* mice encoded a blue fluorescent protein (BFP) reporter fused to the carboxyl terminus of Hoxb5 via a p2A splicing element (Supplementary Fig. 6a), which allowed tracking of cells with conditional expression of Hoxb5. Among mice maintained on 1 mg/ml doxycycline in the drinking water for 1 week, BFP$^+$ pro-pre-B cells were detected in the BM of 8-week-old Tet-*Hoxb5* mice (Fig. 6a). Next, BFP$^+$ pro-pre-B cells sorted from the BM of Tet-*Hoxb5* (CD45.2$^+$) mice were transferred into recipient (CD45.1$^+$) mice of the non-obese diabetic–severe combined immunodeficiency (NOD-SCID) strain (Tet-*Hoxb5*–NOD-SCID mice), which were treated with 1 mg/ml doxycycline in the drinking water from 1 d before to 4 weeks after transplantation. At 4 weeks after transfer, BFP$^+$CD45.2$^+$Ter119$^-$CD19$^-$Mac1$^-$CD3$^+$ T lymphocytes were detected in PB and spleen of recipient mice, although over 95% of the donor-derived CD45.2$^+$ T cells were BFP$^-$ (Supplementary Fig. 6c). When Tet-*Hoxb5*–NOD-SCID mice were maintained on doxycycline for 2 weeks after transplantation and then switched to water without doxycycline (Supplementary Fig. 6b), all T lymphocytes in PB and spleen were BFP$^-$ at 4 weeks after transplantation (Supplementary Fig. 6d). Furthermore, all BFP$^+$ T cells, but no BFP$^-$ T cells, in Tet-*Hoxb5*–NOD-SCID mice expressed *Hoxb5* (Supplementary Fig. 6e). Donor-derived CD45.2$^+$ T lymphocytes were still detected in the PB, LNs, spleen and thymus of Tet-*Hoxb5*–NOD-SCID mice at 4, 8 and 12 weeks after transplantation and then sharply decreased in the PB (Fig. 6b,c).

We next investigated whether Hoxb5 induced the de-differentiation of pro-pre-B cells into MPPs. We analyzed the upstream multi-lineage progenitor cells, including CD2$^-$CD3$^-$CD4$^-$CD8$^-$B220$^-$Mac1$^-$Gr1$^-$Ter119$^-$ (Lin$^-$) c-kit$^+$Sca-1$^+$ cells, which include HSC, MPP and lymphoid-primed MPP subsets, as well as Lin$^-$CD127$^+$c-kit^{int}Sca-1^{int} common lymphoid progenitors, from the BM of 15-TF mice, retro-*Hoxb5* mice, recipients of CD19-*Hoxb5* pro-pre-B cells, and recipients of Tet-*Hoxb5* pro-pre-B cells. We did not detect donor-derived Lin$^-$c-kit$^+$Sca-1$^+$ cells or common lymphoid progenitors in the BM of these mice (Supplementary Fig. 7a–d). In addition, the BFP$^-$ T cells in Tet-*Hoxb5*–NOD-SCID mice simultaneously carried B cell antigen receptor immunoglobulin heavy-chain V(D)J rearrangements and TCRβ V(D)J rearrangements (Supplementary Table 4 and Supplementary Fig. 4c), indicative of their B cell origin. Thus, transient expression of Hoxb5 in pro-pre-B cells was sufficient for stable conversion of B cells into T cells in vivo, and sustained expression of Hoxb5 was dispensable after the target cells committed to lineage conversion.

Hoxb5 converts B cells into ETPs. To determine whether Hoxb5 directly reprogrammed pro-pre-B cells into ETPs (induced ETPs (iETPs)), we investigated the dynamics of Lin$^-$CD44$^+$c-kit^{hi}CD25$^-$ iETPs (Fig. 7a) in the BM and thymus of retro-*Hoxb5* mice. GFP$^+$ iETPs were detected in the BM of retro-*Hoxb5* mice from week 2 to week 6 after transplantation (Fig. 7b), with maximal abundance at

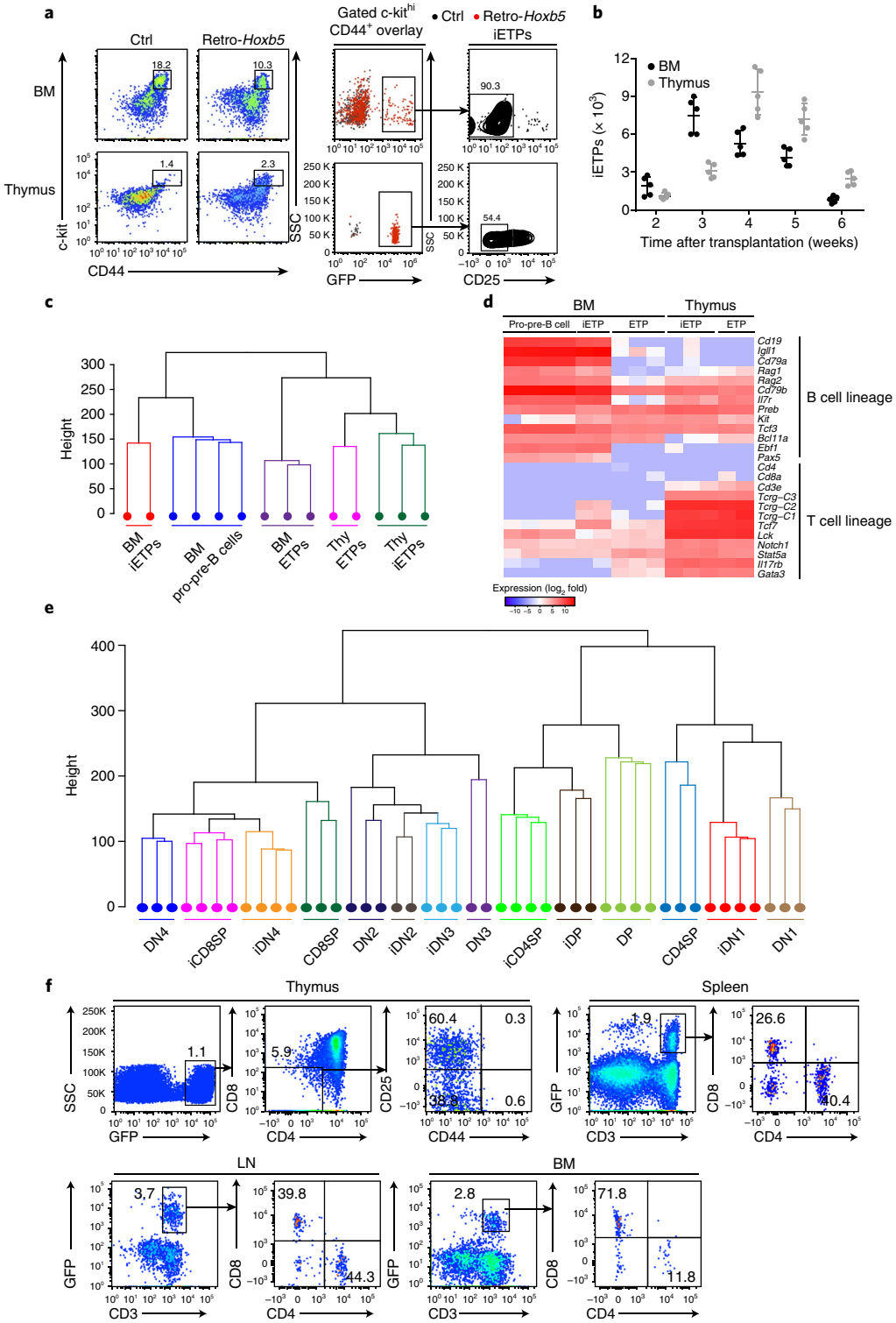

Fig. 7 | Hoxb5 directly converts B lymphocytes into ETP-like cells in the BM. a, Flow cytometry of iETPs obtained from the BM and thymus (left margin) of sublethally irradiated individual recipient mice (among $n=5$ per group) 4 weeks after transplantation of 5×10^6 sorted pro-pre-B cells transduced to express retro-*Hoxb5* or control virus expressing GFP alone (Ctrl) (above plots (left), or key (right plots)). **b**, Absolute number of iETPs cells obtained from the BM and thymus (key) of *Hoxb5* recipient mice ($n=5$ mice) at various times (horizontal axis) after transplantation of pro-pre-B cells transduced to express retro-*Hoxb5*. **c**, Unsupervised hierarchical clustering of RNA-Seq data of iETPs, natural ETPs ($Lin^-CD44^+c\text{-}kit^{hi}CD25^-$) from the BM (BM iETPs or BM ETPs) and thymus (Thy iETPs or Thy ETPs), and pro-pre-B cells (horizontal axis; 1×10^3 iETPs, pro-pre-B cells or ETPs) sorted from retro-*Hoxb5* or wild-type control mice. **d**, Expression of selected genes related to T cells or B cells (right margin) in pro-pre-B cells, iETPs and natural ETPs (above plot; one biological replicate per column), presented as $\log_2$ values (obtained from FPKM values; key). **e**, Unsupervised hierarchical clustering of RNA-Seq data of iDN1, iDN2, iDN3, iDN4, iDP, iCD4SP and iCD8SP cells and their wild type counterparts (horizontal axis). **f**, Flow cytometry of iT lymphocytes obtained from the thymus, spleen, LNs and BM (above plots) of a recipient mouse 3 weeks after secondary transplantation of iT thymocytes. Numbers in or adjacent to outlined areas or quadrants indicate percent cells in each. Each symbol (**b**) represents an individual host mouse; small horizontal lines indicate the mean (±s.d.). Data are representative of two independent experiments.

approximately 3 weeks after transplantation, and they were detected in thymus from week 2 to week 6 after transplantation (Fig. 7b), with maximal abundance at approximately 4 weeks after transplantation. RNA-Seq analysis was performed on GFP$^+$ Lin$^-$CD44$^+$c-kit^{hi}CD25$^-$ iETPs sorted from BM and thymus of retro-*Hoxb5* mice and GFP$^-$ Lin$^-$CD44$^+$c-kit^{hi}CD25$^-$ ETPs from the thymus of 8-week-old wild-type mice (CD45.2$^+$ mice on the C57BL/6 background). Unsupervised hierarchical clustering analysis revealed that GFP$^+$ iETPs in the thymus mapped closely to wild-type thymic ETPs (Fig. 7c). In contrast, GFP$^+$ iETPs in the BM clustered closely to wild-type pro-pre-B cells (Fig. 7c). Of note, some B cell–related genes, including *Cd19*, *Cd79a*, *Ebf1* and *Pax5*, were still expressed in BM GFP$^+$ iETPs (Fig. 7d), which suggested that the BM GFP$^+$ iETPs were in the process of lineage conversion. GFP$^+$ iETPs in the thymus had entirely repressed their expression of genes encoding the B cell lineage regulators Pax5 and Ebf1 (Fig. 7d), consistent with published reports that repression of *Pax5* and *Ebf1* is crucial for the B cell–to–T cell conversion[19,20]. These data supported the hypothesis that the Hoxb5-induced conversion of B cells into ETPs began in the BM, while the reprogramming was completed in thymus.

We further assessed the gene-expression pattern of the progeny of thymic GFP$^+$ iETPs in retro-*Hoxb5* mice at the DN1, DN2, DN3, DN4, DP, CD4SP and CD8SP stages of development (as defined above). Unsupervised hierarchical clustering analysis revealed that GFP$^+$ iDN1, iDN2 and iDN3 cells (where 'i' again indicates cells induced by Hoxb5) mapped closely to wild-type DN1, DN2 and DN3 thymocytes, respectively, and GFP$^+$ iDN4 cells mapped closely to iCD8SP cells and wild-type DN4 cells (Fig. 7e). GFP$^+$ iDP cells mapped closely to wild-type DP cells, whereas GFP$^+$ iCD8SP cells and iCD4SP cells were close to wild-type CD8SP cells and CD4SP cells, respectively (Fig. 7e). Therefore, the Hoxb5-induced thymocyte subsets resembled their wild-type counterparts at the transcriptome level.

To confirm the functionality of the iETPs in the thymus, we performed a secondary transplantation assay in which total thymocytes from retro-*Hoxb5* mice were transferred into the thymus of sublethally irradiated CD45.2$^+$ congenic mice (on the C57BL/6 background). At 3 weeks after secondary transplantation, GFP$^+$CD4$^-$CD8$^-$CD44$^-$CD25$^+$ iDN3 cells and GFP$^+$CD4$^-$CD8$^-$CD44$^-$CD25$^-$ iDN4 cells were detected in the thymus of secondary recipients, but GFP$^+$CD4$^-$CD8$^-$CD44$^+$CD25$^-$ DN1 cells and CD4$^-$CD8$^-$CD44$^+$CD25$^+$ iDN2 cells were not. We also detected mature GFP$^+$ CD3$^+$ iT cells in the spleen (1.9%), LNs (3.7%) and BM (2.8%) of the secondary recipients (Fig. 7f), with phenotypes and tissue distribution patterns similar to those of their counterparts from wild-type mice (Supplementary Fig. 7e–g); this suggested that iETPs normally differentiated into T cells in secondary recipient mice. Thus, Hoxb5 converted B cells into functional ETPs.

Hoxb5 targets B cell and T cell regulators and chromatin modifiers and remodelers. To investigate the genes targeted by Hoxb5 in pro-pre-B cells, we performed RNA-Seq analysis of pro-pre-B cells transduced with retro-*Hoxb5* or empty vector and cultured for 3 d with the cytokines IL-7, Flt3l and SCF. Differential gene-expression analysis indicated that 358 genes were significantly upregulated and 422 genes were significantly downregulated in pro-pre-B cells transduced with retro-*Hoxb5* relative to their expression in pro-pre-B cells transduced with empty vector (a difference in expression of over twofold; adjusted *P* value,<0.05 (DESeq2 R package); Fig. 8a and Supplementary Table 5). Using a lower threshold for this analysis (a difference in expression of over 1.2-fold; adjusted *P* value, < 0.01 (DESeq2 R package)), we identified 3,700 differentially expressed genes (Supplementary Table 6) as potential targets of Hoxb5, among which 232 (Supplementary Table 7) were identified previously as encoding transcription factors (http://genome.gsc.riken.jp/TFdb/tf_list.html). Gene-ontology analysis of these 232 transcription factor–encoding genes indicated enrichment for genes encoding proteins involved in chromatin modification and remodeling, myeloid and lymphoid cell differentiation and T cell differentiation in pro-pre-B cells transduced with retro-*Hoxb5* relative to their expression in pro-pre-B cells transduced with empty vector (Fig. 8b). Expression of Hoxb5 in pro-pre-B cells altered the expression of many genes encoding chromatin modifiers, including Hdac9 Ezh1, Ldb1, Cbx8 and Asxl1 (Fig. 8c), consistent with the alterations to the gene-expression pattern during B cell fate–to–T cell fate reprogramming. Notably, genes encoding transcription factors essential for early B cell development, such as Ebf1, Bcl11a, Foxp1 and Foxo1[21–24], were repressed by Hoxb5 (Fig. 8d), while genes encoding transcription factors critical for T cell development or function, such as Nfatc1, Tcf12, Lmo2 and Prdm1[25–28], were activated by Hoxb5 in pro-pre-B cells (Fig. 8d). In addition, expression of the gene encoding the B cell–specific marker CD19 was reduced, whereas expression of the gene encoding the T cell marker Il-2Rα was increased, by Hoxb5 expression in pro-pre-B cells (Supplementary Table 6). Gene set–enrichment analysis indicated that the target genes of the transcription factors Ikaros (encoded by *Ikzf1*) and Pax5, which are essential for B cell development[29–31], were significantly repressed in pro-pre-B cells transduced with retro-*Hoxb5* relative to their expression in pro-pre-B cells transduced with empty vector (Fig. 8e,f). Furthermore, target genes of the epigenetic modifier Mll (encoded by *Kmt2a*), a histone methyltransferase important for HSC self-renewal[32,33], were significantly repressed in pro-pre-B cells transduced with retro-*Hoxb5* relative to their expression in pro-pre-B cells transduced with empty vector (Fig. 8g). These results indicated that Hoxb5 expression in pro-pre-B cells repressed the expression of B cell lineage–specific transcription factors, enhanced the expression of T cell–related transcription factors and altered the expression of chromatin and epigenetic modifiers.

To identify the direct targets of Hoxb5 in pro-pre-B cells, we performed ChIP followed by deep sequencing (ChIP-Seq) of retro-*Hoxb5* pro-pre-B cells. Due to the lack of CHIP-Seq-grade antibodies to Hoxb5, we obtained pro-pre-B cells from Rosa26$^{BirA/BirA}$ transgenic (C57BL/6) mice, which constitutively express the enzyme BirA that can biotinylate biotin-tagged proteins, and transduced the cells with a Hoxb5–biotin acceptor sequence fusion protein (Bio-Hoxb5) (Supplementary Fig. 8a); this resulted in the biotinylation of Bio-Hoxb5 (Supplementary Fig. 8b). Of note, expression of Bio-Hoxb5 did not alter the ability of Hoxb5 to reprogram B lymphocytes into T lymphocytes in vivo (Supplementary Fig. 8c,d). B220$^+$ B cells isolated from the BM of 6-week-old Rosa26$^{BirA/BirA}$ mice were transduced with retrovirus encoding GFP–Bio-Hoxb5 and cultured for 3 d in vitro, and the viable GFP$^+$ Bio-Hoxb5–expressing pro-pre-B cells were sorted for streptavidin-mediated ChIP-Seq analysis of Hoxb5. 163 peak-correlated transcription factor–encoding genes were identified in pro-pre-B cells expressing Bio-Hoxb5 (Supplementary Table 8); these overlapped with the differentially expressed genes from the transcriptome analysis (Supplementary Table 6). Hoxb5 directly targeted *Ebf1*, *Pax5*, *Bcl11a*, *Foxp1* and *Foxo1*[21–24,34] (Fig. 8h), which encode transcription factors essential for B cell development, as well as *Lmo2*, *Nfatc1*, *Tcf12*, *Prdm1* and *Runx2* (Fig. 8h, Supplementary Fig. 8e and Supplementary Table 8), which encode transcription factors with important roles in the regulation of T cell function[25–28,35]. In addition, Hoxb5 directly targeted the chromatin modifier–encoding genes *Kmt2a* (which encodes Mll) and *Hdac9* and the chromatin remodeler–encoding genes *Ldb1* and *Smarca5*[36–39] (Fig. 8h and Supplementary Fig. 8e). Thus, Hoxb5 directly targeted genes encoding B cell master regulators, T cell regulators and crucial chromatin modifiers and remodelers in pro-pre-B cells, which ultimately drove the B cell fate–to–T cell fate conversion.

Discussion

Here we found that forced expression of Hoxb5 in pro-pre-B cells was sufficient to convert mouse B cells into T cells in vivo.

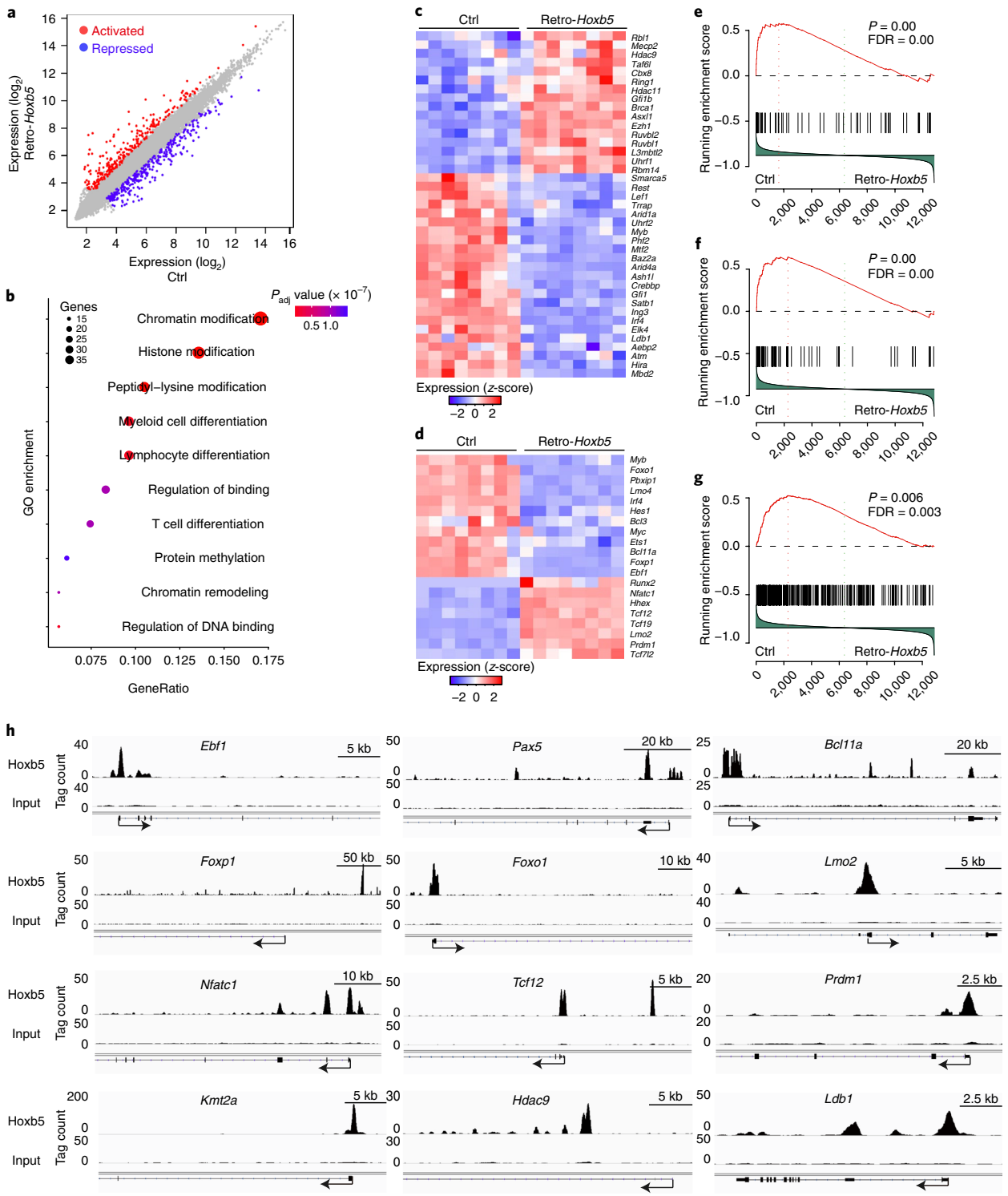

Fig. 8 | Hoxb5 targets in pro-pre-B cells. a, Expression of various genes in pro-pre-B cells transduced to express empty vector (horizontal axis) plotted against their expression in pro-pre-B cells transduced to express retro-*Hoxb5* (vertical axis) ($n = 8$ biological replicates), presented as normalized expression values (mean) and plotted as FPKM values; colors (key) indicate genes significantly upregulated (red; Activated) or downregulated (blue; Repressed) in retro-*Hoxb5*-expressing pro-pre-B cells (a difference in expression of over twofold; adjusted *P* value, < 0.05 (DESeq2 R package)). **b**, Gene ontology (GO)–enrichment analysis of the 232 differentially expressed transcription factor–encoding genes identified: each symbol represents a GO term (noted in plot); color indicates adjusted *P* value (P_{adj} (significance of the GO term); bottom key), and symbol size is proportional to the number of genes (top key). **c**,**d**, Expression of genes encoding chromatin-modification-related transcription factors (**c**) or lymphopoiesis-related transcription factors (**d**) (right margin), selected from the list of differentially expressed genes (a difference in expression of over 1.2-fold; adjusted *P* value, <0.01 (DESeq2 R package)), in pro-pre-B cells transduced to express empty vector (Ctrl) or retro-*Hoxb5* (above plots) (presented as in Fig. 1a; $n = 8$ biological replicates (one per column)). **e**–**g**, Gene set–enrichment analysis of gene targets activated by Ikaros (**e**), Pax5 (**f**) in pro-B cells or Mll (**g**) in hematopoietic progenitor cells ($CD48^-$ Lin^-c-$kit^+$$Sca\text{-}1^+$ cells), assessed in pro-pre-B cells transduced to express empty vector or retro-*Hoxb5* (below plots). **h**, CHIP-Seq analysis of the-binding of Hoxb5 to *Ebf1*, *Pax5*, *Bcl11a*, *Foxp1*, *Foxo1*, *Lmo2*, *Nfatc1*, *Tcf12*, *Prdm1*, *Kmt2a* (*Mll*), *Hdac9* and *Ldb1* (above plots) in pro-pre-B cells expressing Bio-Hoxb5. Data are representative of two independent experiments.

After transplantation into sublethally irradiated recipient mice, Hoxb5-expressing pro-pre-B cells gave rise to ETPs in the BM; those cells subsequently matured into fully functional polyclonal T lymphocytes in the thymus of the recipient mice. This cell-fate conversion was the consequence of Hoxb5-mediated repression of genes encoding B cell master regulators, activation of genes encoding T cell regulators and regulation of genes encoding chromatin and epigenetic modifiers and remodelers.

Under homeostasis, the efficiency of Hoxb5-induced B cell–to–T cell conversion was very low. HSC-derived natural T cell lymphopoiesis was able to outperform Hoxb5-induced T cell lymphopoiesis. However, in the adoptive-transfer experiments with irradiated mice, the irradiation at least partially eradicated HSCs and progenitor cells and transiently impaired endogenous T cell lymphopoiesis, which allowed pro-pre-B cell–derived iT cell lymphopoiesis to occur in the BM microenvironment. Hoxb5-mediated B cell–to–T cell conversion started in the BM and was completed in the thymus. Our detection of reprogrammed intermediate cells in the BM indicated an essential role for the BM niche and also indicated that the BM was required for the B cell–to–T cell reprogramming in the recipient mice. Complete silencing of the genes encoding the B cell master regulators Ebf1 and Pax5, was observed only in thymic iETPs, suggestive of a crucial role for the thymic niche as well in B cell–to–T cell reprogramming. Moreover, the thymic stromal niche guarantees the essential processes of negative and positive selection of iT cell lymphopoiesis[40] and thus reduces the risk of generating autoreactive iETPs derivatives.

We estimate that in the B cell–to–T cell conversion system induced by expression of retro-*Hoxb5*, a million retro-*Hoxb5*-expressing pro-pre-B cells gave rise to tens of thousands of ETPs in vivo, which would suggest that the reprogramming efficiency of Hoxb5 might have been as low as 1%. Thus, a minority of pro-pre-B cells overexpressing Hoxb5 were reprogrammed into iETPs cells, but the majority were able to differentiate normally into B cells. The transcription factor Cebpα can reprogram B into macrophages at an efficiency of 100% (ref. [4]), which might indicate the existence of co-factor(s) that enhance(s) the efficiency of the B cell–to–T cell reprogramming by Hoxb5. In addition, deletion of endogenous Hoxb5 did not affect hematopoiesis, including B cell and T cell lymphopoiesis. Consistent with that, mice lacking the entire gene cluster encoding Hoxb proteins have normal hematopoiesis[15]. One possible explanation for the dispensability of Hoxb5 during normal lymphopoiesis is the existence of potential compensatory molecules. The identification of such potential co-factors of Hoxb5 or molecules that could compensate for the absence of Hoxb5 might improve the understanding of Hoxb5-induced B cell–to–T cell conversion.

We emphasize that forced expression of Hoxb5 in pro-pre-B cells, by the procedure used here, yielded fully functional T lymphocytes in vivo whose transcriptomes, hierarchical differentiation, tissue distribution and immunological functions closely resembled those of natural mouse T cells. Thus, this approach has advantages over alternative approaches previously explored, including direct manipulation of the expression of Pax5 and Ebf1[10,11] This could reflect the fact that endogenous *Hoxb5* is expressed in HSCs and MPPs[14] but is not expressed in committed B cells or T cells. Thus, this research could lead to new insights into lineage conversion in the immune system.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41590-018-0046-x.

Received: 1 April 2017; Accepted: 8 January 2018;
Published online: 12 February 2018

References

1. Heyworth, C., Pearson, S., May, G. & Enver, T. Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. *EMBO J.* **21**, 3770–3781 (2002).
2. Kulessa, H., Frampton, J. & Graf, T. GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev.* **9**, 1250–1262 (1995).
3. Visvader, J. E., Elefanty, A. G., Strasser, A. & Adams, J. M. GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. *EMBO J.* **11**, 4557–4564 (1992).
4. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663–676 (2004).
5. Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556–562 (1999).
6. Rolink, A. G., Nutt, S. L., Melchers, F. & Busslinger, M. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* **401**, 603–606 (1999).
7. Taghon, T., Yui, M. A. & Rothenberg, E. V. Mast cell lineage diversion of T lineage precursors by the essential T cell transcription factor GATA-3. *Nat. Immunol.* **8**, 845–855 (2007).
8. Laiosa, C. V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L. & Graf, T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. *Immunity* **25**, 731–744 (2006).
9. Li, P. et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. *Science* **329**, 85–89 (2010).
10. Cobaleda, C., Jochum, W. & Busslinger, M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* **449**, 473–477 (2007).
11. Ungerbäck, J., Åhsberg, J., Strid, T., Somasundaram, R. & Sigvardsson, M. Combined heterozygous loss of Ebf1 and Pax5 allows for T-lineage conversion of B cell progenitors. *J. Exp. Med.* **212**, 1109–1123 (2015).
12. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
13. Månsson, R. et al. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* **26**, 407–419 (2007).
14. Chen, J. Y. et al. Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* **530**, 223–227 (2016).
15. Bijl, J. et al. Analysis of HSC activity and compensatory Hox gene expression profile in Hoxb cluster mutant fetal liver cells. *Blood* **108**, 116–122 (2006).
16. Riddell, J. et al. Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* **157**, 549–564 (2014).
17. Ramasamy, I., Brisco, M. & Morley, A. Improved PCR method for detecting monoclonal immunoglobulin heavy chain rearrangement in B cell neoplasms. *J. Clin. Pathol.* **45**, 770–775 (1992).
18. Ehlich, A., Martin, V., Müller, W. & Rajewsky, K. Analysis of the B-cell progenitor compartment at the level of single cells. *Curr. Biol.* **4**, 573–583 (1994).
19. Souabni, A., Cobaleda, C., Schebesta, M. & Busslinger, M. Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. *Immunity* **17**, 781–793 (2002).
20. Nechanitzky, R. et al. Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nat. Immunol.* **14**, 867–875 (2013).
21. Lin, H. & Grosschedl, R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* **376**, 263–267 (1995).
22. Liu, P. et al. Bcl11a is essential for normal lymphoid development. *Nat. Immunol.* **4**, 525–532 (2003).
23. Hu, H. et al. Foxp1 is an essential transcriptional regulator of B cell development. *Nat. Immunol.* **7**, 819–826 (2006).
24. Lin, Y. C. et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11**, 635–643 (2010).
25. Müller, M. R. et al. Requirement for balanced Ca/NFAT signaling in hematopoietic and embryonic development. . *Proc. Natl. Acad. Sci. USA* **106**, 7034–7039 (2009).
26. Braunstein, M. & Anderson, M. K. HEB in the spotlight: Transcriptional regulation of T-cell specification, commitment, and developmental plasticity. *Clin. Dev. Immunol.* **2012**, 678705 (2012).
27. Martins, G. & Calame, K. Regulation and functions of Blimp-1 in T and B lymphocytes. *Annu. Rev. Immunol.* **26**, 133–169 (2008).
28. Ferrando, A. A. et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell* **1**, 75–87 (2002).
29. Schwickert, T. A. et al. Stage-specific control of early B cell development by the transcription factor Ikaros. *Nat. Immunol.* **15**, 283–293 (2014).

30. Schebesta, A. et al. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* **27**, 49–63 (2007).
31. McManus, S. et al. The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J.* **30**, 2388–2404 (2011).
32. McMahon, K. A. et al. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**, 338–345 (2007).
33. Artinger, E. L. et al. An MLL-dependent network sustains hematopoiesis. *Proc. Natl. Acad. Sci. USA* **110**, 12000–12005 (2013).
34. Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nat. Immunol.* **8**, 463–470 (2007).
35. 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 β-selection. *J. Immunol.* **169**, 2866–2874 (2002).
36. Zhang, H. et al. MLL1 inhibition reprograms epiblast stem cells to naive pluripotency. *Cell Stem Cell* **18**, 481–494 (2016).
37. Zhou, X., Marks, P. A., Rifkind, R. A. & Richon, V. M. Cloning and characterization of a histone deacetylase, HDAC9. *Proc. Natl. Acad. Sci. USA* **98**, 10572–10577 (2001).
38. Krivega, I., Dale, R. K. & Dean, A. Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **28**, 1278–1290 (2014).
39. Kokavec, J. et al. The ISWI ATPase Smarca5 (Snf2h) is required for proliferation and differentiation of hematopoietic stem and progenitor cells. *Stem Cells* **35**, 1614–1623 (2017).
40. Starr, T. K., Jameson, S. C. & Hogquist, K. A. Positive and negative selection of T cells. *Annu. Rev. Immunol.* **21**, 139–176 (2003).

Acknowledgements

We thank T. Cheng, D. Pei and E. H. Bresnick for comments on the manuscript; Z. Liu (Institute of Biophysics, CAS, China) for *Rag1*$^{-/-}$ mice; and the animal center and instrument center of Guangzhou Institutes of Biomedicine and Health for the animal care, cell sorting and skin imaging. Supported by the Major National Research Project of China (2015CB964401 to J.W.; 2015CB964404 to Y.-G.Y. and Z.H.; 2015CB964902 to J.D.; and 2015CB964901 to H.W., CAS Key Research Program of Frontier Sciences (QYZDB-SSW-SMC057), the Major Scientific and Technological Project of Guangdong Province (2014B020225005), the Strategic Priority Research Program of the Chinese Academic of Sciences (XDA01020311), the co-operation program of the Guangdong Natural Science Foundation (2014A030312012), the National Natural Science Foundation of China (31471117 and 81470281 to J.W.; 31600948 to D.Y.; 91642208 to Y.-G.Y.; and 81770222 to D.W.), the National Key Research and Development Program of China (2017YFA0103401 to B.L.; and 2017YFA0103402 to A. H.) and the US National Institutes of Health (AI079087 and HL130724 to D.W.).

Author contributions

M.Z., Y.D. and F.H. performed the core experiments and contributed equally to this work; D.Y., Q.Z., C. Lv, Y.W., C.X., Q.W., X.L. and C. Li performed multiple experiments; P.Z., T.W., Y.G., R.G., L.L., Y.G. and H.W. performed certain in vitro experiments; J.D., Z.H., S.X., J.C., A.H., B.L., D.W., Y.-G.Y. and J.W. discussed the data and edited the manuscript; B.L., D.W. and J.W. wrote the manuscript; and J.W. designed the project and provided final approval of the manuscript.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41590-018-0046-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.W.

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

Methods

Mice. C57BL/6 (CD45.1$^+$), and CD19-Cre (C57BL/6, CD45.2$^+$) mice were purchased from the Jackson Laboratory. C57BL/6 (CD45.2$^+$), and NOD-SCID and BALB/c mice were purchased from Bejing Vital River Laboratory Animal Technology. *Rag1*$^{-/-}$ mice (C57BL/6) were a gift from Z. Liu from Institute of Biophysics (CAS, China) and from Model Animal Research Center of Nanjing University. *Hoxb5*$^{LSL/+}$ and Tet-*Hoxb5* mice were generated by targeting a mouse ES line (C57BL/6 line, Beijing Biocytogen) through homologous recombination at the ROSA26 locus. *Hoxb5*$^{LSL/+}$ mice were subsequently bred with CD19-Cre or Vav-cre mice to generate *Hoxb5*$^{LSL/+}$CD19-cre and *Hoxb5*$^{LSL/+}$Vav-cre mice. *Hoxb5*$^{-/-}$ mice were generated by direct targeting the zygotes of C57BL/6 (CD45.2) mice via TALEN to delete 37-bp from exon 1 of *Hoxb5* on Chr11. Mice were housed in the SPF-grade animal facility of the Guangzhou Institution of Biomedicine and Health, Chinese Academy of Science (GIBH, CAS, China). All animal experiments were approved by the Institutional Animal Care and Use Committee of Guangzhou Institutes of Biomedicine and Health (IACUC-GIBH).

Recombinant vectors, viral packaging and transduction of pro-pre-B cells. The cDNA encoding each factor was inserted into pMYs-IRES-EGFP to generate a recombinant vector (RTV-021, Cell Biolabs). Each recombinant vector or empty vector control was transduced (Calcium Phosphate Transfection method) into Plat-E cells (containing retroviral packaging elements) to produce high-titer, replication-incompetent viruses. The titer of each virus was adjusted to a multiplicity of infection of 0.69 using a NIH/3T3 cell line as described[41]. Subsequently, the transduction efficiency of each virus can reach ~50% in pro-pre-B cells in our experimental setting.

4- to 6-week-old donor mice (C57BL/6, CD45.2) were killed and BM cells were collected. After lysis of red blood cells, the nucleated cells were blocked by Fc blocker, incubated with biotin-conjugated antibody to B220 (anti-B220) and enriched by Anti-Biotin MicroBeads by AutoMACS Pro (Miltenyi Biotec). Pro-pre-B cells (Ter119$^-$Mac1$^-$CD3$^-$CD4$^-$CD8$^-$B220$^+$CD19$^+$CD93$^+$IgM$^-$) were sorted from the enriched B220$^+$ cells by Moflo Astrios (Beckman Coulter) and then cultured in medium (15%FBS, 100 mM GlutaMAX, 10^{-4} M β-ME, 10 ng/ml mSCF, 10 ng/ml Flt3L and 10 ng/ml IL7) for 12–16 h before virus transduction. Pro-pre-B cells (1×10^6 per ml) were transduced with the mixed viruses having adjusted titers by two-round spin-infection (800 *g*,90 min,35 °C).

Transplantation. For pro-pre-B cell transplantation, 1×10^6 to 5×10^6 sorted pro-pre-B cells were injected into the retro-orbital veins of sublethally irradiated recipient mice (C57BL/6 recipients, 6.5 Gy; NOD-SCID recipients, 2.25 Gy; *Rag1*$^{-/-}$ recipients, 3.5 Gy; RS2000, Rad Source). Mice were fed trimethoprim-sulfamethoxazole-treated water for 2 weeks to prevent infection. For competitive BM transplantation assay, 0.5×10^6 total BM cells from either *Hoxb5*$^{LSL/+}$Vav-Cre mice or *Hoxb5*$^{-/-}$ mice (CD45.2$^+$) with equivalent number of competitor cells (CD45.1$^+$) were retro-orbitally transplanted into lethally irradiated (9.0 Gy) individual CD45.1$^+$ recipients. For intra-thymus transplantation, the surgery procedure was performed as described[42]. The amount of thymocytes equivalent to a quarter of a donor thymus were injected into sublethally irradiated congenic(CD45.2$^+$) mice.

Flow cytometry. Antibodies to CD2 (RM2-5), CD3 (145-2C11), CD4 (RM4-5), CD8a (53-6.7), Gr1 (RB6-8C5), Mac1 (M1/70), Ter119 (TER-119), IgM (II/41), Thy1.2 (53-2.1), B220 (6B2), c-kit (2B8), Sca-1 (E13-161.7), FcγRII/III (2.4G2), CD25 (PC61), CD93 (PB.493), IgD (1.19), CD28 (37.51), Foxp3 (Fjk-16s), TCRβ (H57-597), TCRγδ (GL3), CD21/35 (7G6), CD23 (B3B4), CD5 (53-7.3), CD69-PE (H1.2F3), IFNγ-APC (XMG1.2), IL-17-Percp-cy5.5 (TC11-18H10.1), CD44 (IM7), CD127 (SB/199), Ly6D (49-H4) CD45.2 (104) CD45.1(A20) were from eBioscience or BioLegend. DAPI, 7-AAD and PI were used to stain dead cells. Flow cytometry was performed on an LSR Fortessa (BD Biosciences) and data were processed by FlowJo software (Tree Star).

Analysis of immunoglobulin heavy-chain V(D)J rearrangement. Semi-nested PCR for detecting immunoglobulin heavy chain V(D)J rearrangement was performed as described[18]. For single cell analysis, individual cells were sorted into PBS, lysed and subsequently amplified by Discover-sc single cell kit according to manufacturer's protocols (N601-02, Vazyme Biotech). Amplification processes were carried out in two rounds: first, containing following 5′primers (VHJ558, 5′-ARGCCTGGGRCTTCAGTGAAG-3′; VHQ52, 5′-GCGAAGCTTCTCACAGAGCCTGTCCATCAC-3′.) and the JH4E primer (5′-AGGCTCTGAGATCCCTAGACAG-3′), 200 ng DNA template, and with a 60 °C annealing/30 s extension, and 20 cycle program; second, 0.25 µl first round product was diluted and used as template. Then 35 cycles of amplification were performed with either primer pairs VHJ558 and the nested JH4A (5′-GGGTCTAGACTCTCAGCCGGCTCCCTCAGGG-3′), or primer pairs VHJ558 and the nested JH4A as described[18].

TCR beta chain VD,J rearrangement analysis. Individual iT lymphocytes were sorted into PBS using AriaII sorter (BD Biosciences). The single cell genome was amplified by Discover-sc single cell kit according to manufacturer's protocols (N601-02, Vazyme Biotech). PCR for detecting TCRβ V(D)J rearrangements was performed as described[10,43] using the following primers: Vβ2 (upstream): GGGTCACTGATACGGAGCTG, Vβ4 (upstream): GGACAATCAGACTGCCTCAAGT, Vβ5.1 (upstream): GTCCAACAGTTTGATGACTATCAC, Vβ8 (upstream): GATGACATCATCAGGTTTTGTC, and Jβ2 (downstream): TGAGAGCTGTCTCCTACTATCGATT. After 35 cycles of amplification (15 s at 95 °C, 20 s at 60 °C, 1 min at 72 °C), PCR products were purified and cloned into pMD18-T vector (6011, TaKaRa). To confirm the identities of the PCR products, three clones of each recombinant were sequenced. The specific V(D)J rearrangements were further aligned by the Igblast tool (https://www.ncbi.nlm.nih.gov/igblast)[44].

RNA-Seq and data analysis. The cDNA of sorted 1,000-cell aliquots were generated and amplified as described previously[45]. The quality of the amplified cDNA was assessed by qPCR analysis of housekeeping genes (*B2m*, *Actb*, *Gapdh* and *Ecf1a1*). Samples that passed quality control were used for sequencing library preparation by illumina Nextera XT DNA Sample Preparation Kit (FC-131-1096). All libraries were sequenced by illumina sequencers NextSeq 500 (illumina). The fastq files of sequencing raw data of the total 92 RNA-Seq samples were generated using illumina bcl2fastq software and were uploaded to Gene Expression Omnibus public database (GSE105057). Raw reads alignment and differential gene expression analysis were performed by Bowite2, DEseq2, Tophat2 and Cufflinks2.2.1 as reported[46,47]. The processed data were uploaded to the server of Gene Expression Omnibus (GSE105057). Unsupervised clustering analysis was performed using factextra (https://cran.r-project.org/web/packages/factoextra/index.html). Heat maps were plotted using gplots (heatmap.2). Gene set–enrichment analysis (GSEA) and gene-ontology (GO)–enrichment analysis (clusterProfiler R package) were performed as described[48,49]. The gene sets for GSEA were from literature as follows: Ikaros-activated targets in pro-B cells[29], Pax5-activated targets in pro-B cells[30,31] and Mll-activated targets in CD48$^-$ Lin$^-$c-kit$^+$Sca-1$^+$ cells[33].

iT lymphocyte stimulation assay in vitro. The iT lymphocyte–stimulation assay was performed according to the protocol described[50]. In brief, mAb to CD3 (2C11) at a concentration of 50 µg/ml was coated onto 96-well plates (100 µl per well) overnight at 4 °C, followed by two washes with PBS. Sorted splenic iT lymphocytes (GFP$^+$CD3$^+$) were suspended at 1×10^6 cells per ml in RPMI 1640 medium (containing 10% FBS and 100 mM glutamine) and were then seeded into the coated 96-well plates (2×10^5 cells per well). RPMI 1640 culture medium containing 5 µg/ml of mAb to CD28 was used for cell culture. The cells were incubated for 6 d, and the supernatants were collected and analyzed for IL-2 (BMS601), IL-10 (BMS614/2), IFN-γ (BMS606) and TNF (BMS607/3) by ELISA according to the manufacturer's guide (eBioscience).

One way mixed lymphocyte reaction assay. In brief, individual wells containing 5×10^5 responder iT cells (C57BL/6 background) from iT-*Rag1*$^{-/-}$ mice that had rejected the allogeneic skin graft and 5×10^5 irradiated (30 Gy) allogeneic splenic cells as stimulators (BABL/c background) were incubated in 1640 medium containing 10% FBS and 5 µM cell proliferation dye Fluor670 (eBioscience, 65-0840) at 37 °C in 5% CO2. The same responder cells incubated with medium without stimulators were used as a control. After 3–4 d, the cells were analyzed by flow cytometry.

Allogeneic skin graft and immunofluorescence staining. Individual *Rag1*$^{-/-}$ mice were given transplantation of 3×10^6 pro-pre-B cells from *Hoxb5*$^{LSL/+}$CD19-Cre or *Hoxb5*$^{LSL/+}$ control mice. After 3 weeks, the allogeneic skin (BALB/c background) was transplanted using a procedure as described[51]. Grafts were considered rejected if there was a loss of distinct border, visible signs of ulceration and necrosis to 80% of the graft area. After 6–10 d, the rejected skin tissues were removed for analysis. For activated T cell analysis, the cell suspensions were isolated as described[52]. The activated alloreactive iT lymphocytes were analyzed by LSR Fortessa (BD Biosciences) by gating on CD45.2$^+$Mac1$^-$CD19$^-$CD69$^+$CD44$^+$CD4$^+$CD8$^+$. Data were further processed by FlowJo software (Tree Star). For analysis of cytokines released by the alloreactive iT lymphocytes, we used a intracellular staining protocol (ebioscience) using anti-IL-17 (TC11-18H10.1, 1:100) and anti-IFN-γ (XMG1.2, 1:100). For immunofluorescence staining, skin tissues were fixed in 4% buffered formalin (alfa aesar, X07A031) at 4 °C overnight and cryosectioned (5 µM). The infiltration of iT lymphocytes in the rejected skin tissues (BALB/c) were detected by immunofluorescence staining as described[51]. The primary antibody mAb to CD3 (1:500, ab5690, Abcam) and secondary antibody donkey anti-rabbit IgG conjugated with Alexa fluor-555 (1:1000, ab150062, abcam) were used. All slides were observed using a microscope (LSM800, Zeiss) and the images were processed by Adobe Photoshop Element software (version 4.0, Adobe Systems).

CHIP-Seq and data analysis. 160×10^6 (1.6×10^8) enriched B220$^+$ cells from ten Rosa26$^{BirA/BirA}$ transgenic mice were transduced with Bio-Hoxb5 retro-virus (Bio-Hoxb5 pro-pre-B cells). The Bio-Hoxb5 pro-pre-B cells were cultured in vitro for 3 d before collection for streptavidin-mediated chromatin precipitation

using a protocol described previously[53,54]. In brief, 8×10^6 sorted GFP$^+$ Bio-Hoxb5 pro-pre-B cells were fixed with 1% formaldehyde in culture medium for 12 min at room temperature, followed by quenching with 0.125 M glycine for 5 min. The cells were washed twice with ice-cold PBS, and lysed in hypotonic buffer (20 mM HEPES, pH 7.5, 10 mM KCl, 10% glycerol, 1 mM EDTA and 0.2% NP-40, plus protease inhibitor cocktail) for at least 30 min on ice to break the cell membrane. Then the cells were lysed in lysis buffer (1% SDS, 10 mM EDTA and 50 mM Tris, pH 8.0, plus protease inhibitor cocktail) for at least 30 min on 4 °C to break the cell nucleus. The cross-linked chromatin were sonicated to an average fragment size of 300 bp by using a Bioruptor (UCD-300TO) Sonicator. After preclearing with protein G beads, 1% of the sonicated DNA fragments were used as CHIP-Seq input control, and the 99% sonicated DNA fragments were incubated with streptavidin beads (Dynabeads Streptavidin M280) at 4 °C overnight. Beads were washed with the following buffers: 2% SDS (twice); 0.1% deoxycholate, 1% Triton X-100, 2 mM EDTA, 50 mM HEPES and 500 mM NaCl (once); 0.1% deoxycholate, 1% Triton X-100, 2 mM EDTA, 50 mM HEPES and 150 mM NaCl (thrice); and TE buffer (20 mM Tris–HCl, pH 8.0, and 2 mM EDTA) (twice). To reverse cross-links, 100 µl elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA and 1% SDS) was added to the pelleted beads and placed in 70 °C water bath overnight. At the same time, the input chromatin sample was also added to 100 µl elution buffer for reverse cross-link along with precipitation samples. Proteinase K was added to the chromatin samples, followed by incubation for 6 h at 55 °C. Genomic DNA was isolated from the precipitated material by phenol extraction and ethanol precipitation. DNA segments from ChIP or sonication input control were end-repaired and ligated to indexed Illumina adaptors, followed by low-cycle PCR (VAHTS Universal DNA Library Prep Kit ND604). The resulting libraries were sequenced with a Illumina HiSeq X Ten platform. The sequencing raw data were converted to fastq files using illumina bcl2fastq software and were uploaded to Gene Expression Omnibus public database (GSE105057). The raw reads (Fastq files) were aligned to mouse genome (mm10) using Bowtie2 package. Peak calling was performed using homer package and visualized using IGV software.

Statistical analysis. Kaplan-Meier survival analysis used Log-rank test. Statistical analysis of RNA-Seq data used the DESeq2, GSEA and GO packages. Other significant analyses were performed using unpaired Student's *t*-test (GraphPad Prism, GraphPad Software).

Life Sciences Reporting Summary. Further information on experimental design and reagents is available in the Life Sciences Reporting Summary.

Data availability. The data that support the findings of this study are available from the corresponding author upon request. All RNA-Seq and CHIP-Seq data are in the GEO database with accession code GSE105057.

References

41. Limón, A. et al. High-titer retroviral vectors containing the enhanced green fluorescent protein gene for efficient expression in hematopoietic cells. *Blood* **90**, 3316–3321 (1997).
42. Goldschneider, I., Komschlies, K. L. & Greiner, D. L. Studies of thymocytopoiesis in rats and mice. I. Kinetics of appearance of thymocytes using a direct intrathymic adoptive transfer assay for thymocyte precursors. *J. Exp. Med.* **163**, 1–17 (1986).
43. Wolfer, A., Wilson, A., Nemir, M., MacDonald, H. R. & Radtke, F. Inactivation of Notch1 impairs VDJβ rearrangement and allows pre-TCR-independent survival of early αβ lineage thymocytes. *Immunity* **16**, 869–879 (2002).
44. Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* **41**, W34–40 (2013).
45. Tang, F. et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5**, 516–535 (2010).
46. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
48. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
49. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
50. Trickett, A. & Kwan, Y. L. T cell stimulation and expansion using anti-CD3/CD28 beads. *J. Immunol. Methods* **275**, 251–255 (2003).
51. Lan, P., Tonomura, N., Shimizu, A., Wang, S. & Yang, Y. G. Reconstitution of a functional human immune system in immunodeficient mice through combined human fetal thymus/liver and CD34$^+$ cell transplantation. *Blood* **108**, 487–492 (2006).
52. Jiang, X. et al. Skin infection generates non-migratory memory CD8+ T(RM) cells providing global skin immunity. *Nature* **483**, 227–231 (2012).
53. Ai, S. et al. EED orchestration of heart maturation through interaction with HDACs is H3K27me3-independent. *eLife* **6**, e24570 (2017).
54. He, A. et al. Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. *Nat. Commun.* **5**, 4907 (2014).

nature research

Corresponding author(s): Jinyong Wang

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Pilot studies were used for estimation of the sample size to ensure adequate power. In most of the experiments, 3 to 10 mice/samples was sufficient to identify differences between groups with at least 80% power and a 5% significance level.

2. Data exclusions

Describe any data exclusions.

No data were excluded.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Allocation into experimental groups was done randomly

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Yes, all of us were blinded to group allocation during collection and analysis.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Flowjo for FACS data analysis, Bowite2, Tophat2, Cufflinks2.2.1, gplots, clusterprofiler, DESeq2, GSEA and GO R packages for RNA-Seq data analysis, Homer and IGV software for CHIP-Seq data analysis, GraphPad Prism for survival analysis.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	All materials used are readily available from the authors.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	All the antibodies are from comercial sources and were purchased from eBioscience, BD Bioscience, Biolegend or Abcam. eBioscience. anti-mouse CD2 (clone RM2-5,1/200 dilution, cat #13-0021-85) anti-mouse CD3 (clone 145-2C11,1/200 dilution, cat #11-0031-82) anti-mouse CD4 (clone RM4-5, 1/200 dilution, cat # 13-0041-85) anti-mouse CD8a (clone 53-6.7, 1/200 dilution, cat # 13-0081-85) anti-mouse Gr1 (clone RB6-8C5, 1/200 dilution, cat # 13-5931-85) anti-mouse Mac1 (clone M1/70, 1/200 dilution, cat # 13-0112-85) anti-mouse Ter119 (clone TER-119, 1/200 dilution, cat # 13-5921-85) anti-mouse IgM (clone II/41, 1/200 dilution, cat # 17-5790-82) anti-mouse B220 (clone 6B2, 1/200 dilution, cat #13-0452-85) anti-mouse c-Kit (clone 2B8, 1/200 dilution, cat # 47-1171-82) anti-mouse Sca1 (clone E13-161.7, 1/200 dilution, cat # 45-5981-82) anti-mouse CD16/32(clone 93, 1/200 dilution, cat # 14-0161-86) anti-mouse CD25 (PC61, 1/200 dilution, cat # 48-0253-80) anti-mouse CD93 (clone PB.493, 1/200 dilution, cat # 25-5892) anti-mouse IgD (clone 11-26, 1/200 dilution, cat # 17-5993-82) anti-mouse TCRβ (clone H57-597,1/200 dilution, cat # 17-5961-81) anti-mouse CD44 (clone IM7, 1/200dilution, cat # 25-0441-82) anti-mouse CD127 (clone SB/199, 1/200dilution, cat # 47-1271-80) anti-mouse CD45.2 (clone 104, 1/200dilution, cat # 47-0454-82) anti-mouse CD45.1(clone A20, 1/200dilution, cat # 13-0453-85) Biolegend anti-mouse TCRγδ (clone GL3, 1/200 dilution, cat # 118103) anti-mouse CD23 (clone B3B4, 1/200 dilution, cat # 101614) anti-mouse CD5 (clone 53-7.3, 1/200 dilution, cat # 100623) anti-mouse CD69 (clone H1.2F3, 1/200 dilution, cat #104507) anti-mouse IFNγ (clone XMG1.2, 1/200dilution, cat # 505810) anti-mouse IL-17 (clone TC11-18H10.1, 1/200dilution, cat # 506903) BD Bioscience anti-mouse CD21/35 (clone 7G6, 1/200 dilution, cat #561770) Abcam anti-mouse CD3 (1/500 dilution, cat # ab5690) Donkey Anti-Rabbit IgG H&L (1/1000 dilution, cat # ab150062)

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

Plate E cell line (CELL BIOLABS,INC. RV-101); NIH/3T3 cell line (ATCC CRL-1658)

b. Describe the method of cell line authentication used.

The cell lines were not authenticated.

c. Report whether the cell lines were tested for mycoplasma contamination.

Yes, the cell line was tested and is mycoplasma free.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified cell line was used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

C57BL/6 (CD45.1), and CD19-Cre (C57BL/6, CD45.2) mice were purchased from the Jackson Laboratory. C57BL/6 (CD45.2), NOD-SCID and BALB/c mice were purchased from Bejing Vital River Laboratory Animal Technology Ltd. Rag1-/- mice (C57BL/6) were gifted from Dr. Zhihua Liu from Institute of Biophysics (CAS, China) and from Model Animal Research Center of Nanjing University. Hoxb5LSL/+ and Tet-Hoxb5 mice were generated by targeting a mouse ES line (C57BL/6 line, Beijing Biocytogen Co., Ltd.) through homologous recombination at the ROSA26 locus. Hoxb5LSL/+ mice were subsequently bred with CD19-Cre or Vav-cre mice to generate Hoxb5LSL/+CD19-cre and Hoxb5LSL/+Vav-cre mice. Hoxb5-/- mice were generated by direct targeting the zygotes of C57BL/6 (CD45.2) mice via TALEN to delete 37-bp from Exon 1 of Hoxb5 on Chr11. All mice were used at 4-8 weeks of age. Mice were housed in the SPF grade animal facility of the Guangzhou Institution of Biomedicine and Health, Chinese Academy of Science (GIBH, CAS, China). All animal experiments were approved by the Institutional Animal Care and Use Committee of Guangzhou Institutes of Biomedicine and Health (IACUC-GIBH).

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This study did not involve human research participants.

nature research

Corresponding author(s): Jinyong Wang

☐ Initial submission ☐ Revised version ☒ Final submission

ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data deposition

1. For all ChIP-seq data:

☒ a. Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ b. Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

2. Provide all necessary reviewer access links. *The entry may remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE105007
3. Provide a list of all files available in the database submission.	Hoxb5-input_clean1_R1.fastq.gz Hoxb5-input_clean1_R2.fastq.gz Hoxb5_clean1_R1.fastq.gz Hoxb5_clean1_R2.fastq.gz
4. If available, provide a link to an anonymized genome browser session (e.g. UCSC).	IGV software (http://software.broadinstitute.org/software/igv/).

Methodological details

5. Describe the experimental replicates.	One hundred and sixty million (1.6x10e8) enriched B220+ cells from ten Rosa26BirA/BirA transgenic mice were transduced with Bio-Hoxb5 retro-virus . The Bio-Hoxb5 virus transduced pro/pre-B cells were cultured in vitro for three days. Eight million sorted GFP+ Bio-Hoxb5 pro/pre-B cells were used as cell input for streptavidin-mediated Bio-Hoxb5 CHIP-Seq.
6. Describe the sequencing depth for each experiment.	For the amplification of the library after adaptor ligation, 18 low-cycle PCR was used according to the VAHTSTM Universal DNA Llibrary Prep Kit for Illumina (ND604). Twenty-six million pair-end reads for input control sample, 34 million pair-end reads for Hoxb5 sample. Reads length for all these files are 100-150bp after adapter cutoff (cutadapt parameter -m 100 -q 10,10 for the pair-end file adapters cutoff).
7. Describe the antibodies used for the ChIP-seq experiments.	Chromatin Immunprecipitation (CHIP) by dynabeads M-280 Streptavidin antibody (Thermo Fisher, catalog is 11205D,2 mL, lot:00381796).
8. Describe the peak calling parameters.	mm10='/bt2_index/mm10'(Bowtie2 index files) bowtie2 -x $mm10 -q -1 Hoxb5_clean1_R1.fastq.gz -2 Hoxb5_clean1_R2.fastq.gz -S Hoxb5.sam bowtie2 -x $mm10 -q -1 Hoxb5-input_clean1_R1.fastq.gz -2 Hoxb5-input_clean1_R2.fastq.gz -S Hoxb5_Input.sam makeTagDirectory Hoxb5_tag Hoxb5.sam -format sam makeTagDirectory Hoxb5_Input_tag Hoxb5_Input.sam-format sam makeUCSCfile Hoxb5_tag/ -o auto makeUCSCfile Hoxb5_Input_tag/ -o auto makeUCSCfile Hoxb5_tag/ -i Hoxb5_Input_tag/ -norm -o auto findPeaks Hoxb5_tag -style factor -o auto -i Hoxb5_Input_tag/ annotatePeaks.pl Hoxb5_tag/peaks.txt mm10 >Hoxb5_annotated.txt
9. Describe the methods used to ensure data quality.	Firstly, we used fastqc to check the fastq files. And used cutadapt software

	to cutoff the adapters. Secondary, default parameters were used for peak calling (FDR:0.001, 4-fold), and 40693 peaks were identified by homer.
10. Describe the software used to collect and analyze the ChIP-seq data.	Homer and IGV

nature research

Corresponding author(s): Jinyong Wang

☐ Initial submission ☐ Revised version ☒ Final submission

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ 3. All plots are contour plots with outliers or pseudocolor plots.

☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

Methodological details

5. Describe the sample preparation.	Bone marrow cells were flushed out through the tibias and femurs by 3ml 2% FBS/DPBS using injector. Spleen and thymus were ground in 2%FBS/ DPBS with pestle. All suspended cell samples were filtered through 70 um strainer to make single cell suspension before antibody staining. Rejected skin graft was chopped into small fragments and incubated in HBSS supplemented with 1 mg/ml collagenase A and 40 µg/ml DNase I at 37°C for 30 min. After filtered through a 70 µm strainer, cells were collected and washed thoroughly with cold PBS before antibody staining.
6. Identify the instrument used for data collection.	BD LSR Fortessa was used for FACS analysis. Moflo Astrios (Beckman Coulter) and AriaII were used for cell sorting.
7. Describe the software used to collect and analyze the flow cytometry data.	We used BD FACS Diva software to collect flow cytometry data, and used FlowJo vX.07 to analyse the flow cytometry data.
8. Describe the abundance of the relevant cell populations within post-sort fractions.	Sorted cells were further analysed by FACS. The purity of the sorted cells reached >99.9%.
9. Describe the gating strategy used.	Cell debris were first excluded by SSC-A and FSC-A. Positive and negative populations were defined in comparison with the unstain sample control of the same type of cells.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☒