

An *Nfil3-Zeb2-Id2* pathway imposes *Irf8* enhancer switching during cDC1 development

Prachi Bagadia^{1,13}, Xiao Huang^{1,13}, Tian-Tian Liu^{1,2,13}, Vivek Durai¹, Gary E. Grajales-Reyes¹, Maximilian Nitschké³, Zora Modrusan⁴, Jeffrey M. Granja^{5,6,7}, Ansuman T. Satpathy¹^{ID}^{5,8}, Carlos G. Briseño⁹, Marco Gargaro^{ID}¹⁰, Arifumi Iwata¹¹, Sunkyoung Kim¹, Howard Y. Chang^{ID}^{5,12}, Andrey S. Shaw³, Theresa L. Murphy¹ and Kenneth M. Murphy^{ID}^{1,2,*}

Classical type 1 dendritic cells (cDC1s) are required for antiviral and antitumor immunity, which necessitates an understanding of their development. Development of the cDC1 progenitor requires an E-protein-dependent enhancer located 41 kilobases downstream of the transcription start site of the transcription factor *Irf8* (+41-kb *Irf8* enhancer), but its maturation instead requires the *Batf3*-dependent +32-kb *Irf8* enhancer. To understand this switch, we performed single-cell RNA sequencing of the common dendritic cell progenitor (CDP) and identified a cluster of cells that expressed transcription factors that influence cDC1 development, such as *Nfil3*, *Id2* and *Zeb2*. Genetic epistasis among these factors revealed that *Nfil3* expression is required for the transition from *Zeb2*^{hi} and *Id2*^{lo} CDPs to *Zeb2*^{lo} and *Id2*^{hi} CDPs, which represent the earliest committed cDC1 progenitors. This genetic circuit blocks E-protein activity to exclude plasmacytoid dendritic cell potential and explains the switch in *Irf8* enhancer usage during cDC1 development.

Development of cDC1s has become a topic of interest because of the critical role this lineage plays in antitumor immunity and checkpoint blockade therapy¹. Dendritic cells (DCs) are an immune lineage encompassing classical DCs (cDCs) and plasmacytoid DCs (pDCs)^{2,3}. The cDCs comprise two branches, cDC1 and cDC2, which exert distinct functions in vivo and rely on different transcriptional programs⁴. pDCs and cDCs can both arise from the CDP^{5–7}. cDC progenitors (pre-cDCs) include clonogenic populations separately committed to cDC1 or cDC2 lineages^{8,9}. Similar progenitors have been confirmed in human DC development^{10–12}. However, the precise transcriptional programs underlying DC specification and commitment remain unclear.

The transcription factors interferon regulatory factor 8 (*Irf8*) and basic leucine zipper transcription factor, ATF-like 3 (*Batf3*) are required for cDC1 development^{9,13,14}, but cDC1s develop from CDPs that express *Irf8* independently of *Batf3*, yet later become dependent on *Batf3* to maintain *Irf8* expression. The basis for this switch from *Batf3*-independent to *Batf3*-dependent *Irf8* expression is unclear. A clonogenic cDC1 progenitor, the pre-cDC1, develops normally in *Batf3*^{–/–} bone marrow (BM) but fails to maintain *Irf8* expression⁹, causing it to divert into cells that are transcriptionally similar to cDC2¹⁵. An enhancer located at +32 kilobases (kb) of the *Irf8* transcriptional start site contained several AP1-IRF composite elements that bind IRF8 and BATF3 in cDC1s in vivo⁹. CRISPR-mediated deletion of the +32-kb *Irf8* enhancer in mice (*Irf8*+32^{–/–}) suggests that *Batf3* supports *Irf8* autoactivation using this enhancer. Similar to *Batf3*^{–/–} mice, *Irf8*+32^{–/–} mice lack mature cDC1s but maintain pre-cDC1 development in vivo. Instead, the development

of this progenitor depends on a +41-kb *Irf8* enhancer, which binds E-proteins and is active in mature pDCs and cDC1 progenitors, but not mature cDC1s. In vivo deletion of this enhancer eliminated *Irf8* expression in pDCs and also completely eliminated development of the specified pre-cDC1. This enhancer activity requires E-proteins to induce sufficient levels of IRF8 during specification of the pre-cDC1, but it is still unclear why mature cDC1s require BATF3 and the +32-kb *Irf8* enhancer to maintain *Irf8* expression.

Other transcription factors are known to influence cDC1 development, such as nuclear factor, interleukin 3, regulated (*Nfil3*), inhibitor of DNA binding 2 (*Id2*) and zinc finger E-box binding homeobox 2 (*Zeb2*) (refs. ^{16–20}). *NFIL3*, a basic leucine zipper transcriptional repressor²¹, is expressed in cDC1s and is required for cDC1 development^{16,22}, but how it functions is unknown^{4,16}. *ID2* is a known inhibitor of E-proteins, is expressed in both cDC1 and cDC2, but is required only for cDC1 development^{17,18}. *ID2* may exclude pDC fate by blocking the activity of E-proteins, particularly E2-2 (*Tcf4*), required for pDCs^{23–25}. However, this model predicts that *Id2*^{–/–} mice should lack both cDC1 and cDC2 lineages, because both lineages must exclude pDC fate. Finally, the transcriptional repressor *ZEB2* is required for pDC development and suppresses cDC1 development, perhaps through inhibition of *Id2* transcription^{19,20}. How these factors precisely interact and at what stage they influence cDC1 specification is unknown.

In the present study, single-cell RNA-seq (scRNA-seq) and genetic epistasis were used to determine the functional hierarchy of transcription factors involved in cDC1 specification. A transcriptional circuit was organized to explain the switch in *Irf8*

¹Department of Pathology and Immunology, Washington University in St. Louis, School of Medicine, St. Louis, MO, USA. ²Howard Hughes Medical Institute, Washington University in St. Louis, School of Medicine, St. Louis, MO, USA. ³Research Biology, Genentech, South San Francisco, CA, USA.

⁴Department of Molecular Biology, Genentech, South San Francisco, CA, USA. ⁵Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, CA, USA. ⁶Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. ⁷Biophysics Program, Stanford University School of Medicine, Stanford, CA, USA. ⁸Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. ⁹Department of Oncology, Amgen Inc., South San Francisco, CA, USA. ¹⁰Department of Experimental Medicine, University of Perugia, Perugia, Italy. ¹¹Department of Allergy and Clinical Immunology, Graduate School of Medicine, Chiba University, Chiba, Japan. ¹²Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. ¹³These authors contributed equally: Prachi Bagadia, Xiao Huang, Tian-Tian Liu. *e-mail: kmurphy@wustl.edu

expression from being *Batf3* independent to being *Batf3* dependent. The CDP originates in a *Zeb2^{hi}* and *Id2^{lo}* state in which *Irf8* expression is maintained by the +41-kb *Irf8* enhancer. The scRNA-seq identified a fraction of the CDP that exclusively possesses cDC1 fate potential. This fraction's development arises when *Nfil3* induces a transition into a *Zeb2^{lo}* and *Id2^{hi}* state. A circuit of mutual *Zeb2*-*Id2* repression serves to stabilize states before and after this transition. *Id2* expression in the specified pre-cDC1 inhibits E-proteins, blocking activity of the +41-kb *Irf8* enhancer, thereby imposing a new requirement for *Batf3* to maintain *Irf8* expression via the +32-kb *Irf8* enhancer.

Results

The earliest committed cDC1 progenitor arises within the CDP. The CDP was originally defined as a Lin⁻CD117^{int}CD135⁺CD115⁺ BM population and was observed to be, although not defined as, largely negative for major histocompatibility complex II (MHC-II) and CD11c expression⁶. Subsequently, pre-cDC1 and pre-cDC2 progenitors were identified as arising from the CDP but were not contained within the CDP^{8,9}. Pre-cDC1s were defined as Lin⁻CD117^{int}CD135⁺CD11c⁺MHC-II^{lo-int} cells and were largely CD115⁻. They can be defined using two methods, relying either on *Zbtb46*-GFP expression in *Zbtb46^{gfp/+}* reporter mice, or on conventional surface markers (Fig. 1a)^{9,26}. In each case, it was noticed that approximately 10% of pre-cDC1s expressed CD115. The expression of CD115 in the pre-cDC1 suggested that cDC1 specification could occur at an earlier developmental stage in the CDP. In agreement, 5–10% of CDPs, defined on the strict exclusion of CD11c- and MHC-II-expressing cells, are *Zbtb46*-GFP^{pos} (Fig. 1b). These *Zbtb46*-GFP^{pos} CDPs had almost exclusive cDC1 potential in an in vitro Fms-related tyrosine kinase 3 ligand (Flt3L) culture, comparable to pre-cDC1s, and completely lacked pDC and cDC2 potential. This was in contrast to the *Zbtb46*-GFP^{neg} CDPs, which produced cells from all three DC lineages (Fig. 1c and see Supplementary Fig. 1a).

The transcriptional profile of these *Zbtb46*-GFP^{pos} CDPs suggests that they represent an intermediate population between a non-specified CDP, the *Zbtb46*-GFP^{neg} CDP and the pre-cDC1 (Fig. 1d,e). For example, genes that had their expression changed more than eightfold between the *Zbtb46*-GFP^{neg} CDPs and the pre-cDC1s were considered. For such genes, their expression in *Zbtb46*-GFP^{pos} CDPs was consistently intermediate between their expression in *Zbtb46*-GFP^{neg} CDPs and pre-cDC1s (Fig. 1d,e and see Supplementary Tables 1 and 2). *Id2* expression in *Zbtb46*-GFP^{neg} CDPs was increased by 34-fold in pre-cDC1s, but only by 15-fold in *Zbtb46*-GFP^{pos} CDPs. Likewise, *Zeb2* expression in *Zbtb46*-GFP^{neg} CDPs was reduced by 9-fold in pre-cDC1s, but only by 3.6-fold in *Zbtb46*-GFP^{pos} CDPs. As expected, the *Zbtb46*-GFP^{pos} CDPs were segregated away from the pre-cDC2s (Fig. 1e). Thus, these results indicate that *Zbtb46*-GFP^{pos} CDPs are an earlier and distinct stage of cDC1 specification compared with the more abundant pre-cDC1 described previously.

scRNA-seq of the CDP identifies factors associated with cDC1 specification. The identification of *Zbtb46*-GFP-expressing cells in the CDP that had almost exclusive cDC1 potential suggested that the CDP might contain cells that have already specified to cDC1 fate. scRNA-seq was performed on 9,554 CDPs defined as Lin⁻CD127⁻CD117^{int}CD115⁺CD135⁺MHC-II⁻CD11c⁻ cells (Fig. 2a) on the 10x Genomics platform to assay for unrecognized heterogeneity within this population. Uniform Manifold Approximation and Projection (UMAP) analysis^{27–29} identified eight connected clusters (Fig. 2b,c). Although it was possible to identify genes that were specifically enriched in certain clusters, others such as *Klf4* and *Ly6d* were not enriched in any one cluster (see Supplementary Fig. 1b). However, scRNA-seq could identify a cluster that was enriched in

Zbtb46 expression, corroborating the data above with the *Zbtb46*-GFP reporter mice. *Zbtb46* was expressed in cluster 3, which also showed restricted expression of *Id2* and *Batf3*, but excluded expression of *Tcf4* and *Zeb2* (Fig. 2d,e). Cluster 3 also showed reduced *Csf1r* expression (Fig. 2d), consistent with lower CD115 surface protein levels in pre-cDC1 and incongruent with the higher CD115 surface protein levels in the bulk CDP (Fig. 1a). As expected, *Flt3* and *Irf8* were uniformly and highly expressed (Fig. 2d,e). Cluster 7, the only other *Tcf4*-negative cluster, probably contained macrophage or neutrophil contamination because this cluster expressed *Ccl6* and did not contain many cells (Fig. 2c,d). Other factors impacting DC development such as *Bcl11a*, *Spi1*, *Klf4* and *Notch2* (refs. 30–33) were not differentially expressed across the CDP, perhaps suggesting that specification of cDC2s and pDCs occurs after the CDP (Fig. 2d and see Supplementary Fig. 1b). In addition, the CDP appeared to be homogeneous with respect to markers of proliferation (Fig. 2f). Thus, scRNA-seq identifies a cluster of cells within the CDP that coordinately induces *Nfil3*, *Id2*, *Batf3* and *Zbtb46* and reduces *Tcf4* and *Zeb2*, suggesting that these genes may regulate cDC1 specification at an earlier stage than previously recognized.

Specification of cDC1s is functionally characterized by low *Zeb2* and high *Id2* expression. To test the functional importance of these genes for cDC1 specification, two reporter mouse lines expressing a ZEB2-EGFP fusion protein (*Zeb2^{gfp}*)³⁴ or an *Id2*-IRES-GFP cassette (*Id2^{gfp}*)³⁵ were analyzed. Both reporters exhibit green fluorescent protein (GFP) expression consistent with the level of *Zeb2* or *Id2* gene expression across many immune lineages (see Supplementary Fig. 2a,b). In *Zeb2^{gfp}* mice, 90% of CDPs expressed high levels of ZEB2-EGFP, but 10% expressed low levels of ZEB2-EGFP, similar to the low levels of ZEB2-EGFP expressed by pre-cDC1s (Fig. 3a). In *Id2^{gfp}* mice, 94% of CDPs expressed low *Id2*-GFP, but 6% expressed high levels of *Id2*-GFP similar to the high levels of *Id2*-GFP expressed by pre-cDC1s (Fig. 3b). Thus, both *Zeb2^{gfp}* and *Id2^{gfp}* reporter lines confirm the existence of ZEB2-EGFP^{lo} and *Id2*-GFP^{hi} cells within the CDP, as predicted by scRNA-seq.

Next, the developmental potential of CDPs expressing high or low levels of ZEB2-EGFP, *Id2*-GFP and *Zbtb46*-GFP was analyzed in an in vitro Flt3L culture system. CDPs expressing low levels of ZEB2-EGFP showed significantly increased cDC1 potential (66%) compared with CDPs expressing high levels of ZEB2-EGFP (26%) (Fig. 3c,e). Likewise, CDPs expressing high levels of *Id2*-GFP showed significantly increased cDC1 potential (77%) compared with CDPs expressing low levels of *Id2*-GFP (30%) at both days 5 and 7 of an in vitro Flt3L culture (Fig. 3d,e and see Supplementary Fig. 2c, d). Finally, CDPs expressing *Zbtb46*-GFP developed almost exclusively into cDC1s (96%), whereas CDPs lacking *Zbtb46*-GFP developed into both cDC1s (30%) and cDC2s (70%) (Figs. 1c and 3e). In all three cases, pDCs developed exclusively from CDPs that were *Zbtb46*-GFP^{neg}, ZEB2-EGFP^{hi} or *Id2*-GFP^{lo} (see Supplementary Fig. 2e–j). These results suggest that CDPs expressing low levels of ZEB2-EGFP or high levels of *Id2*-GFP are biased toward cDC1 development, but not as completely as CDPs expressing *Zbtb46*-GFP.

The transcriptional profile of CDPs expressing low levels of ZEB2-EGFP or high levels of *Id2*-GFP suggests that these cells are a population intermediate between non-specified CDPs and pre-cDC1s (Fig. 3f–i). Genes with expression that differed more than fivefold between the pre-cDC1s and either ZEB2-EGFP^{hi} CDPs (Fig. 3f,g) or *Id2*-GFP^{lo} CDPs (Fig. 3h,i) were considered. The expression of such genes in ZEB2-EGFP^{lo} CDPs was consistently intermediate between the expression in ZEB2-EGFP^{hi} CDPs and pre-cDC1s (Fig. 3f,g and see Supplementary Table 3). Likewise, the expression of such genes in *Id2*-GFP^{hi} CDPs was consistently intermediate between the expression in *Id2*-GFP^{lo} CDPs and that in pre-cDC1s (Fig. 3h,i and see Supplementary Table 4). In addition, the

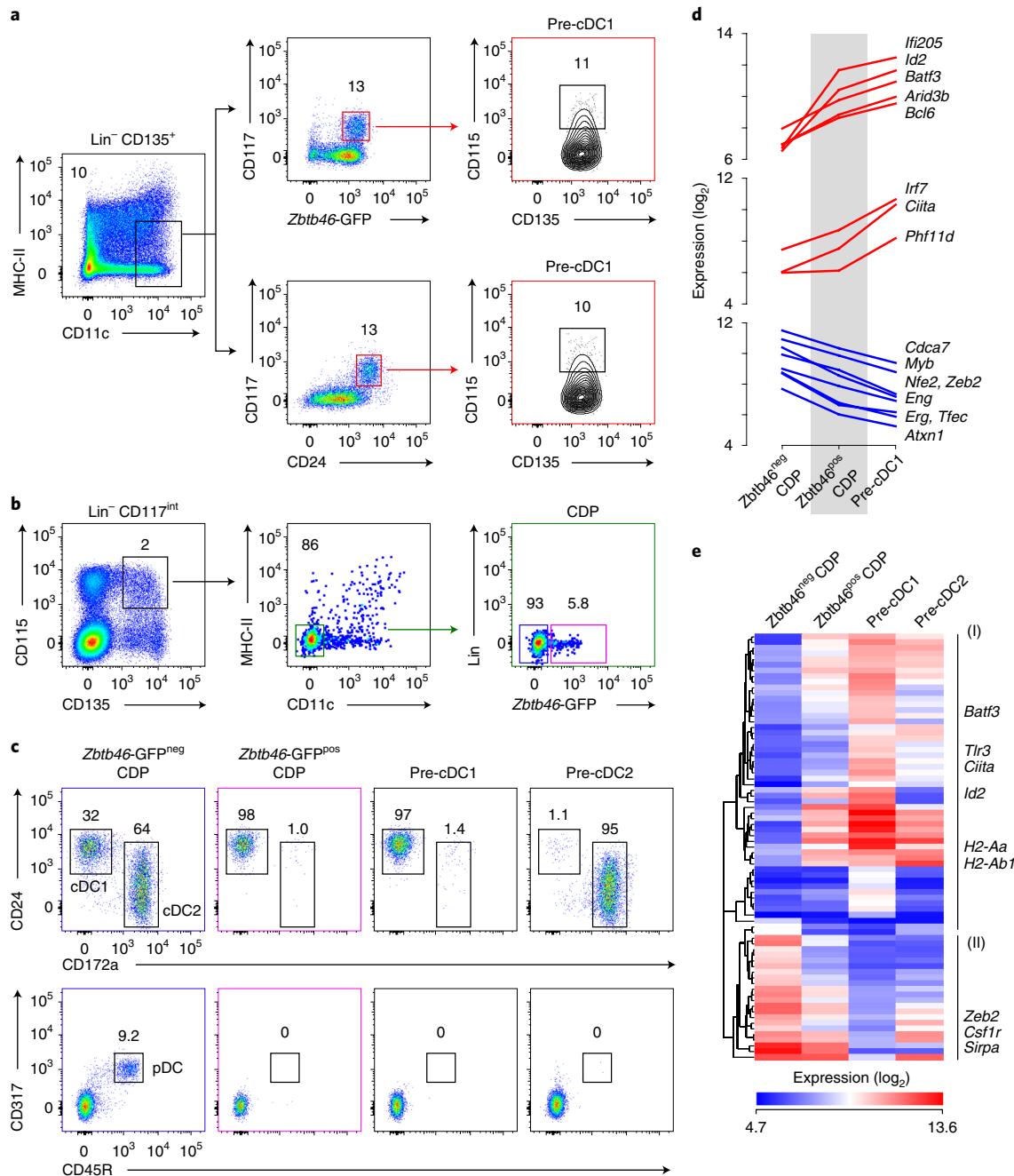


Fig. 1 | Zbtb46-GFP expression in CDPs identifies the earliest committed cDC1 progenitor. **a**, BM from *Zbtb46*^{gfp/+} mice was analyzed by flow cytometry to identify pre-cDC1s as defined by *Zbtb46*-GFP or CD24 expression. Lineage (Lin) included CD3, CD19, NK1.1, Ly6G, TER-119, CD105, CD127 and Siglec-H. Numbers are the percentage of cells in the indicated gates (representing three independent experiments, $n=3$ mice). **b**, BM from *Zbtb46*^{gfp/+} mice was analyzed by flow cytometry to identify the percentage of *Zbtb46*-GFP expression within the CDP. Lineage was defined as in **a** (representing three independent experiments, $n=3$ mice). **c**, *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1s and pre-cDC2s were sorted purified from *Zbtb46*^{gfp/+} mice, cultured for 5 d in Flt3L, and analyzed by flow cytometry for development of pDCs and cDC1s (representing three independent experiments, $n=4$ for *Zbtb46*-GFP^{pos}, *Zbtb46*-GFP^{neg} CDPs and pre-cDC1s and $n=3$ for pre-cDC2s). **d,e**, *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1s and pre-cDC2s were purified as in **c** and analyzed using gene expression microarrays. Expression of transcription factors, with at least fourfold differences between *Zbtb46*-GFP^{neg} CDPs and pre-cDC1s (**d**) or hierarchical clustering for genes, with at least eightfold differences between *Zbtb46*-GFP^{neg} CDP and pre-cDC1s (**e**), are shown (results averaged from biological triplicates for *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs and pre-cDC1 or biological replicates for pre-cDC2).

cells that are ZEB2-EGFP^{lo} within the CDP have induced *Id2*, and those that are *Id2*-GFP^{hi} within the CDP have downregulated *Zeb2* (Fig. 3f-i). Both of these populations also show increasing *Zbtb46* expression compared with the non-specified CDPs. Although these three cDC1-specified CDP populations differ in cDC1 potential, their transcriptional profiles suggest that they are highly

overlapping. In summary, CDPs that express low ZEB2-EGFP or high *Id2*-GFP represent an earlier stage of cDC1 specification compared with the previously identified pre-cDC1.

***Nfil3* is required for cDC1 specification within the CDP.** *Nfil3* is required for cDC1 development¹⁶, but its mechanism and

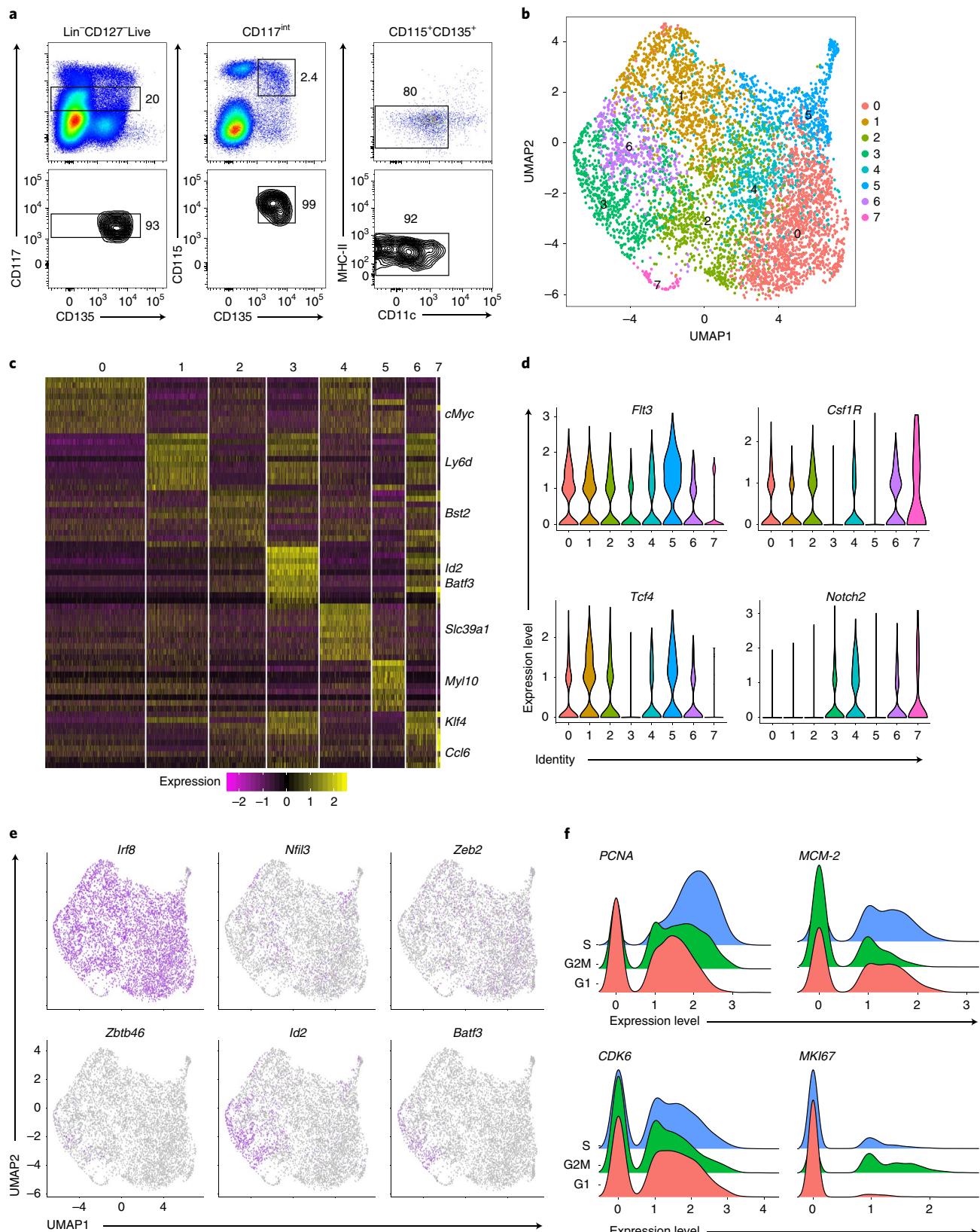


Fig. 2 | Single-cell RNA transcriptome analysis of CDPs. **a**, CDPs gated as Live, Lin⁻CD127⁻CD117^{int}CD115⁺CD135⁺MHC-II⁻CD11c⁻ cells were purified by sorting from C57BL/6J mice. The presort (top) and postsort (bottom) were shown for cells collected for scRNA-seq. Lineage (Lin) included CD105, CD3, CD19, Ly6G and Ter119. **b**, UMAP clustering of CDPs from Seurat analysis (data represent combined analysis of two independent sequencing runs). **c**, Heatmap of 9,954 cells for the top ten genes of each cluster from Seurat analysis. The names of the representative genes within each cluster are shown. **d**, Violin plots depicting cluster identity and expression level for the indicated genes expressed in each cluster as described in **b**. **e**, UMAP plots for the indicated genes as described in **b**. **f**, Joy plots depicting expression level and cell cycle stage for genes involved in the cell cycle.

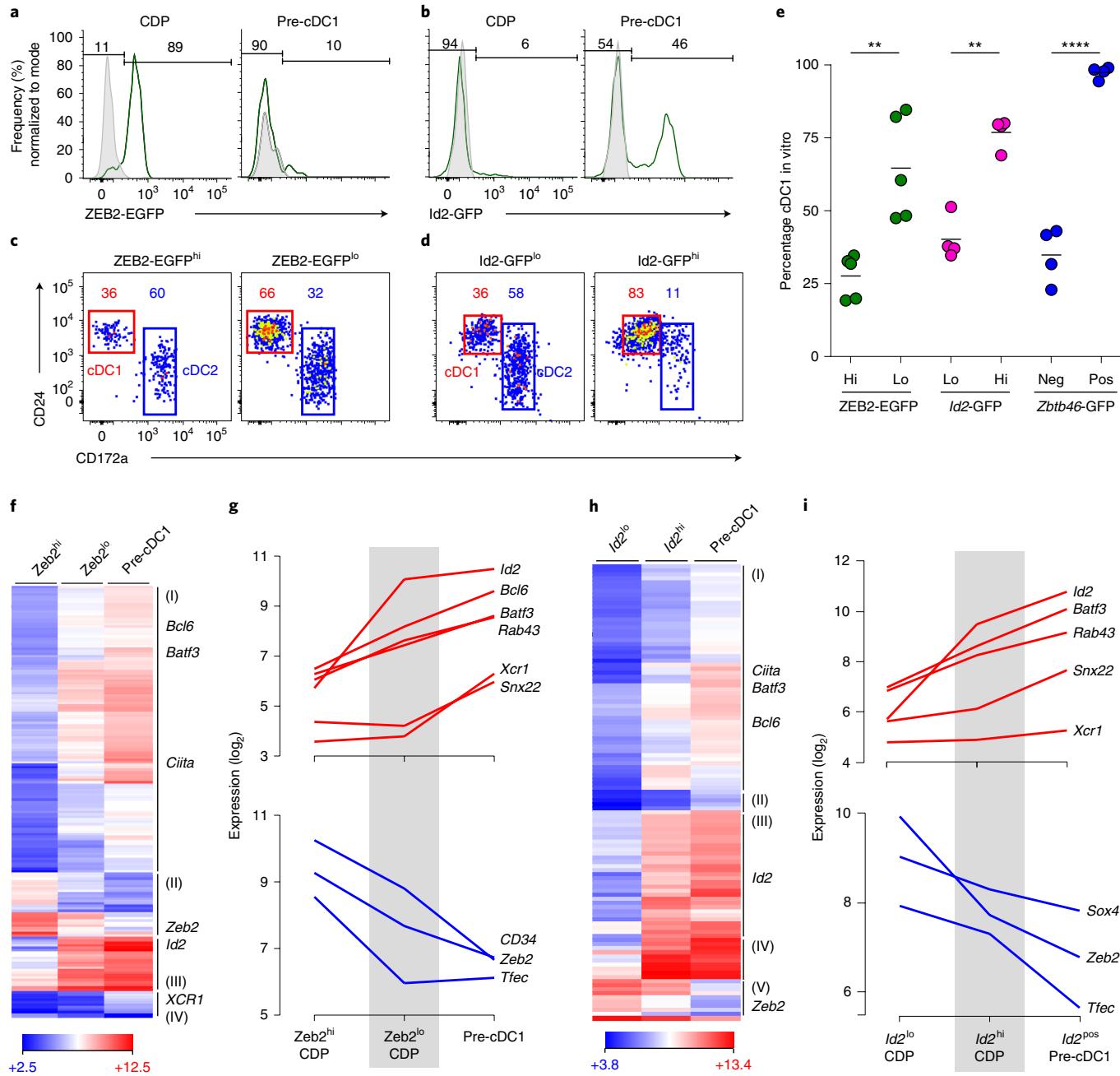


Fig. 3 | *Zeb2* and *Id2* heterogeneity identifies cDC1 specification in CDPs. **a,b**, BM from *Zeb2*^{egfp/egfp} (**a**) or *Id2*^{gfp/+} (**b**) mice were analyzed by flow cytometry to identify GFP expression in CDPs and pre-cDC1s. WT mice (*Zeb2*^{+/+} or *Id2*^{+/+}) are shown as gray histograms. Numbers indicate the percentage of cells in the indicated gates (representing three independent experiments, $n=3$ mice). **c,d**, *ZEB2*-EGFP^{lo} and *ZEB2*-EGFP^{hi} CDPs (**c**) or *Id2*-GFP^{hi} and *Id2*-GFP^{lo} CDPs (**d**) were purified by sorting, cultured for 5 d in Flt3L and analyzed by flow cytometry for development of cDC1 (red) and cDC2 (blue) (representing three independent experiments, $n=5$ for *ZEB2*-EGFP^{lo} and *ZEB2*-EGFP^{hi} CDPs and $n=4$ for *Id2*-GFP^{hi} and *Id2*-GFP^{lo} CDPs). **e**, The indicated cells purified as described in **c** and **d** or in Fig. 1c were cultured as in **c** and analyzed by flow cytometry for cDC1 development, shown as a percentage of total cDCs (CD45⁺CD31⁺MHC-II⁺CD11c⁺) (pooled from three independent experiments, $n=5$ for *ZEB2*-EGFP^{lo} and *ZEB2*-EGFP^{hi} CDPs, $n=4$ for *Id2*-GFP^{hi} and *Id2*-GFP^{lo} CDPs and *Zbtb46*-GFP^{pos} and *Zbtb46*-GFP^{neg} CDPs). Small horizontal lines indicate the mean. **f**, Hierarchical clustering of genes expressed at least fivefold differently between pre-cDC1s and *ZEB2*-EGFP^{hi} CDPs (results averaged from three independent experiments). **g**, Expression of the indicated genes described in **f**. **h**, Hierarchical clustering of genes expressed at least fivefold differently between pre-cDC1s and *Id2*-GFP^{lo} CDPs (results averaged from two independent experiments). **i**, Expression of the indicated genes described in **h**. Data are presented as the mean and the two-tailed, unpaired, Student's *t*-test was used to compare groups. ** $P < 0.01$, *** $P < 0.0001$.

timing of action remain obscure. To determine the stage where *Nfil3* acts in cDC1 development, *Nfil3*^{-/-} mice were crossed with *ZEB2*-EGFP, *Id2*-GFP and *Zbtb46*-GFP reporter mice, and it was assayed whether cDC1-specified progenitors developed in the BM. In *Nfil3*^{+/+}/*Zbtb46*^{gfp/+} reporter mice, cDC1-specified cells

can be identified as CD117^{int}*Zbtb46*-GFP^{pos} cells that include pre-cDC1s and *Zbtb46*-GFP^{pos} CDPs, and comprise approximately 5% of Lin⁻CD135⁺ BM (Fig. 4a,b). However, these cells are absent in *Nfil3*^{-/-}/*Zbtb46*^{gfp/+} mice, although they do develop normally in *Batf3*^{-/-}/*Zbtb46*^{gfp/+} mice as previously described (see Supplementary

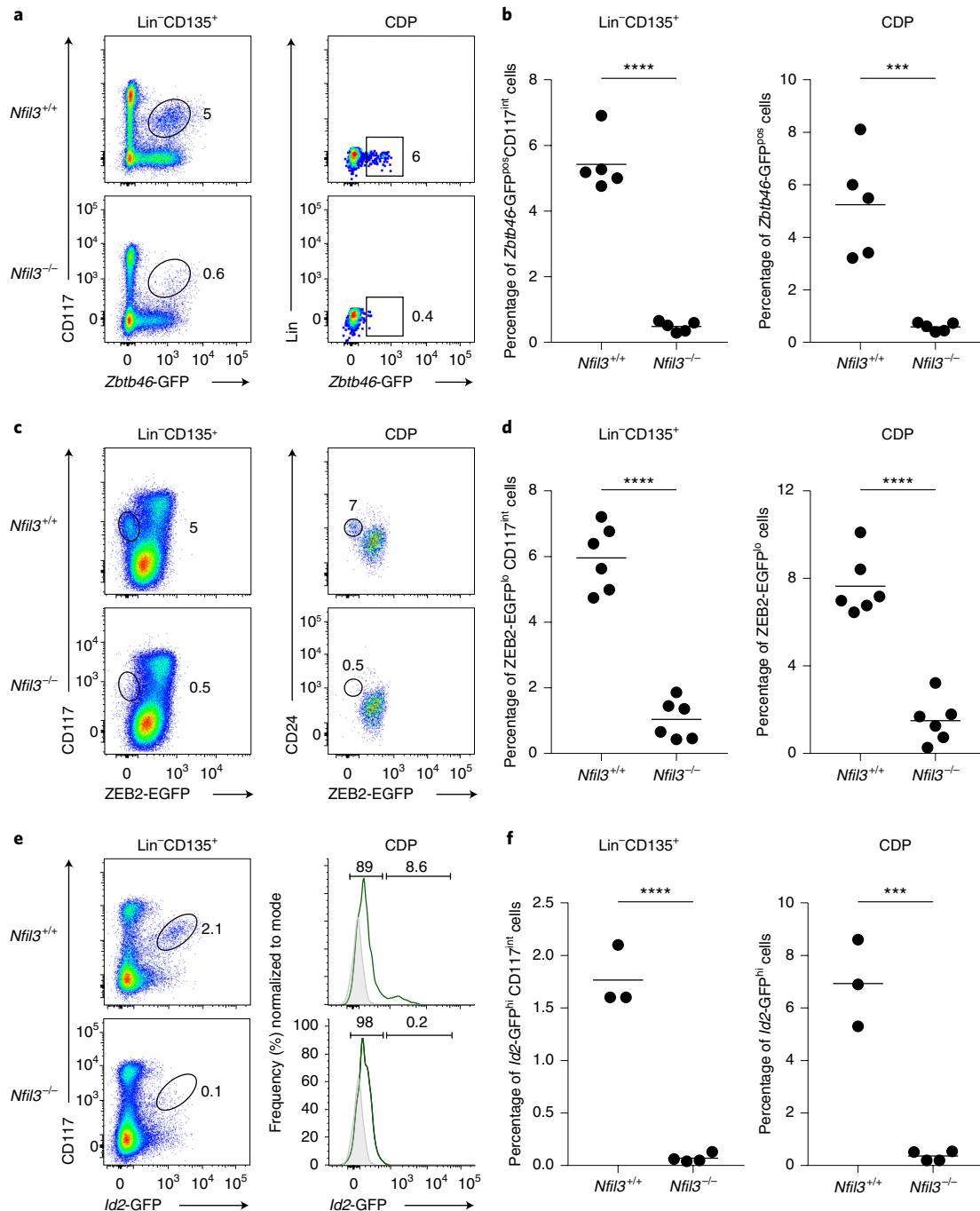


Fig. 4 | *Nfil3* is required for cDC1 specification. **a**, BM from *Nfil3*^{+/+}/*Zbtb46*^{gfp/+} and *Nfil3*^{-/-}/*Zbtb46*^{gfp/+} mice was analyzed by flow cytometry for Lin⁻CD135⁺CD117^{int}*Zbtb46*-GFP^{pos} cells (left) or *Zbtb46*-GFP^{pos} CDPs (right). Numbers indicate the percentage of cells in the indicated gates (representing five independent experiments, $n = 5$ mice). Lineage (Lin) included CD105, CD3, CD19, Ter119, Ly6G and CD127. **b**, Cells from **a** are shown as a percentage of Lin⁻CD135⁺ BM (left) or of CDPs (right). Small horizontal lines indicate the mean. **c**, BM from *Nfil3*^{+/+}/*Zeb2*^{gfp/+} and *Nfil3*^{-/-}/*Zeb2*^{gfp/+} mice was analyzed for Lin⁻CD135⁺CD117^{int}*ZEB2*-EGFP^{lo} cells (left) or *ZEB2*-EGFP^{lo} CDPs (right) (representing three independent experiments, $n = 6$ mice). **d**, Cells from **c** are shown as a percentage of Lin⁻CD135⁺ BM (left) or of CDPs (right). Small horizontal lines indicate the mean. **e**, BM from *Nfil3*^{+/+}/*Id2*^{gfp/+} and *Nfil3*^{-/-}/*Id2*^{gfp/+} mice was analyzed for Lin⁻CD135⁺CD117^{int}*Id2*-GFP^{hi} cells (left) or *Id2*-GFP^{hi} CDPs (right) (representing three independent experiments, $n = 3$ for *Nfil3*^{+/+}/*Id2*^{gfp/+} mice and $n = 4$ for *Nfil3*^{-/-}/*Id2*^{gfp/+} mice). **f**, Cells from **e** are shown as a percentage of Lin⁻CD135⁺ BM (left) or of CDPs (right). Small horizontal lines indicate the mean. Data in **b**, **d** and **f** are presented as the mean and the two-tailed, unpaired, Student's *t*-test was used to compare groups. *** $P < 0.001$; **** $P < 0.0001$.

Fig. 3a)⁹. Within the CDP, cDC1-specified cells can be identified as *Zbtb46*-GFP^{pos} cells that comprise 5% of the CDP (Fig. 4a,b). However, these cells are also absent in *Nfil3*^{-/-}/*Zbtb46*^{gfp/+} mice.

In *Nfil3*^{+/+}/*Zeb2*^{gfp/+} reporter mice, cDC1-specified cells are identified as CD117^{int}*ZEB2*-EGFP^{lo} cells, which include pre-cDC1s and ZEB2-EGFP^{lo} CDPs, and comprise approximately 6% of Lin⁻CD135⁺

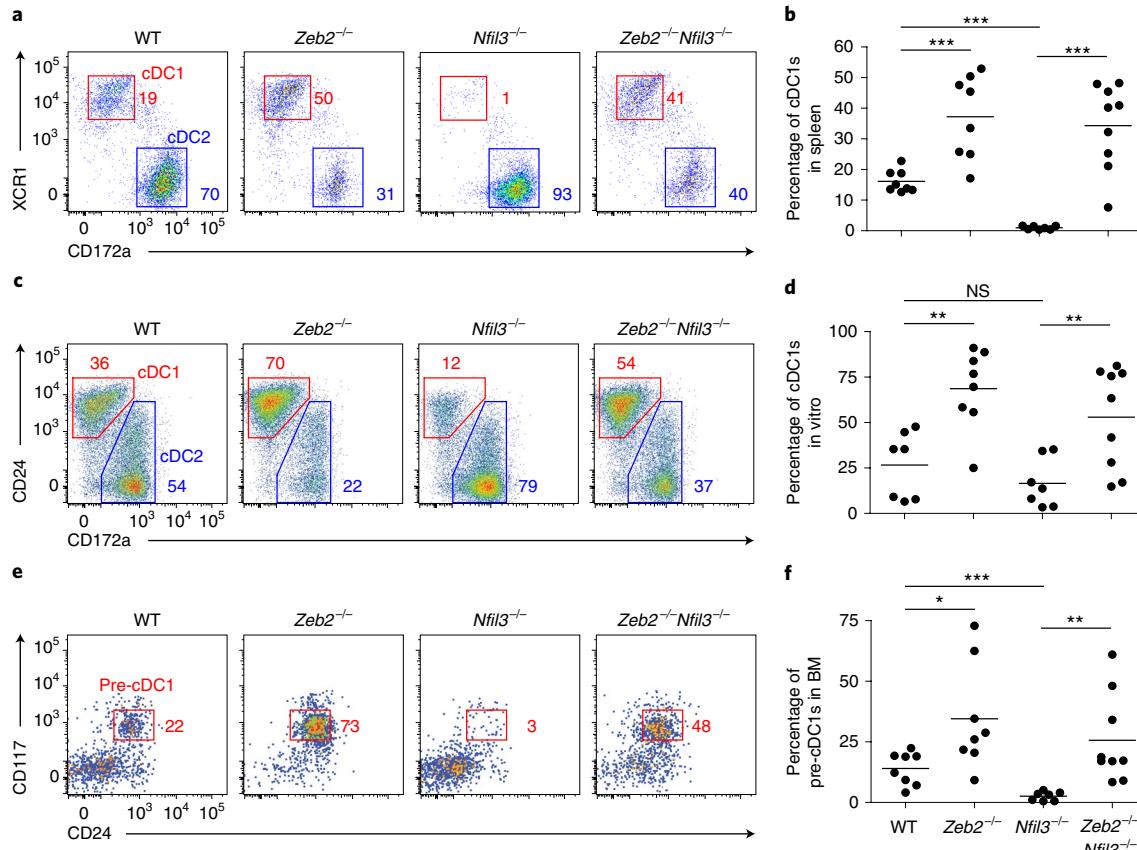


Fig. 5 | Zeb2 is downstream of Nfil3 in cDC1 development. **a**, Splenic cDCs from *Nfil3*^{+/+}*Zeb2*^{fl/fl}*Mx1-cre*⁻ (WT), *Zeb2*^{fl/fl}*Mx1-cre*⁺ (*Zeb2*^{-/-}), *Nfil3*^{-/-} (*Nfil3*^{-/-}) and *Nfil3*^{-/-}*Zeb2*^{fl/fl}*Mx1-cre*⁺ (*Nfil3*^{-/-}*Zeb2*^{-/-}) mice, gated as in Fig. 3e, were analyzed for cDC1 (red) and cDC2 (blue) frequency. Numbers are the percentage of cells in the indicated gates (data representing three independent experiments, $n=7$ for WT and *Zeb2*^{-/-} mice, $n=8$ for *Nfil3*^{-/-} mice and $n=9$ for *Nfil3*^{-/-}*Zeb2*^{-/-} mice). **b**, Analysis from **a** is presented as individual mice. Small horizontal lines indicate the mean. **c**, The cDCs derived from in vitro Flt3L cultures of BM from mice in **a** were analyzed for cDC1 (red) and cDC2 (blue) frequency as in **a** (data representing three independent experiments, $n=7$ for WT and *Zeb2*^{-/-} mice, $n=8$ for *Nfil3*^{-/-} mice and $n=9$ for *Nfil3*^{-/-}*Zeb2*^{-/-} mice). **d**, Analysis from **c** is presented for individual mice. Small horizontal lines indicate the mean. **e**, BM from mice in **a** was analyzed for the frequency of pre-cDC1s (red). BM cells are pre-gated as Lin⁻Siglec-H⁺CD135⁺ (data representing three independent experiments, $n=7$ for WT and *Zeb2*^{-/-} mice, $n=8$ for *Nfil3*^{-/-} mice and $n=9$ for *Nfil3*^{-/-}*Zeb2*^{-/-} mice). **f**, Analysis from **e** is presented for individual mice. Small horizontal lines indicate the mean. The mean and the two-tailed, unpaired, Student's *t*-test were used to compare groups. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant.

BM (Fig. 4c,d). However, these cells are absent in *Nfil3*^{-/-}*Zeb2*^{egfp/+} mice. In *Nfil3*^{+/+}*Zeb2*^{egfp/+} reporter mice, cDC1-specified CDPs can be identified as ZEB2-EGFP^{lo} cells, which comprise 7% of CDPs (Fig. 4c,d), which again are absent in *Nfil3*^{-/-}*Zeb2*^{egfp/+} mice. Finally, in *Nfil3*^{+/+}*Id2*^{gfp/+} reporter mice, cDC1-specified cells can be identified as CD117^{int}*Id2*-GFP^{hi} cells, which include pre-cDC1s and *Id2*-GFP^{hi} CDPs, and comprise approximately 2% of Lin⁻CD135⁺ BM (Fig. 4e,f). However, these cells are absent in *Nfil3*^{-/-}*Id2*^{gfp/+} mice. Furthermore, cDC1-specified CDPs can be identified as *Id2*-GFP^{hi} cells, which comprise 7% of CDPs (Fig. 4e,f), but are absent in *Nfil3*^{-/-}*Id2*^{gfp/+} mice. In summary, *Nfil3* is required for the appearance of all cDC1-specified progenitors identified by *Zbtb46*-GFP, ZEB2-EGFP or *Id2*-GFP.

Zeb2 functions downstream of Nfil3 in cDC1 specification. Next, the interactions between *Nfil3* and other factors were evaluated using genetic mutants rather than GFP reporters. First, interactions between *Nfil3* and *Zeb2* were examined. *Nfil3*^{-/-} mice were crossed to *Zeb2*^{fl/fl}*Mx1-Cre* mice in which ZEB2 can be inactivated by poly(I:C) treatment (*Zeb2*^{-/-}). The development of cDC1s and the presence of cDC1-specified progenitors in *Nfil3*^{+/+}*Zeb2*^{fl/fl}*Mx1-cre*⁻ (wild-type (WT)), *Nfil3*^{-/-}, *Zeb2*^{-/-} and *Nfil3*^{-/-}*Zeb2*^{-/-} mice was examined (Fig. 5). First, *Zeb2*^{-/-} mice had a more than twofold increase in

splenic cDC1s compared with WT mice (Fig. 5a,b), consistent with a previous study²⁰. Furthermore, *Nfil3*^{-/-} mice lacked cDC1s in the spleen, as previously reported¹⁶. However, *Nfil3*^{-/-}*Zeb2*^{-/-} mice had a splenic cDC1 population which, similar to *Zeb2*^{-/-} mice, was about twofold greater than in WT mice. Similarly, in vitro cDC1 development was increased in *Zeb2*^{-/-} BM and reduced in *Nfil3*^{-/-} BM (Fig. 5c,d). However, in vitro cDC1 development from *Nfil3*^{-/-}*Zeb2*^{-/-} BM was increased compared with *Nfil3*^{-/-} BM. Finally, direct examination of pre-cDC1 development was carried out in these mice. *Zeb2*^{-/-} mice had increased numbers of pre-cDC1s compared with WT mice, whereas *Nfil3*^{-/-} mice had greatly reduced numbers of pre-cDC1s (Fig. 5e,f). However, *Nfil3*^{-/-}*Zeb2*^{-/-} mice had markedly restored pre-cDC1 development compared with *Nfil3*^{-/-} mice. In summary, for both in vivo and in vitro cDC1 development and for in vivo cDC1 specification, the phenotype of *Zeb2* deficiency dominates over that of *Nfil3* deficiency, suggesting that *Zeb2* genetically functions downstream of *Nfil3*. The repression of *Zeb2* by *Nfil3* is required in the early stages of cDC1 specification.

Zeb2 functions downstream of Id2 with respect to cDC1 specification. Some evidence suggests that *Zeb2* may function genetically upstream of *Id2* in cDC1 development^{19,20}, but no mechanism has been established. To evaluate the genetic interaction between

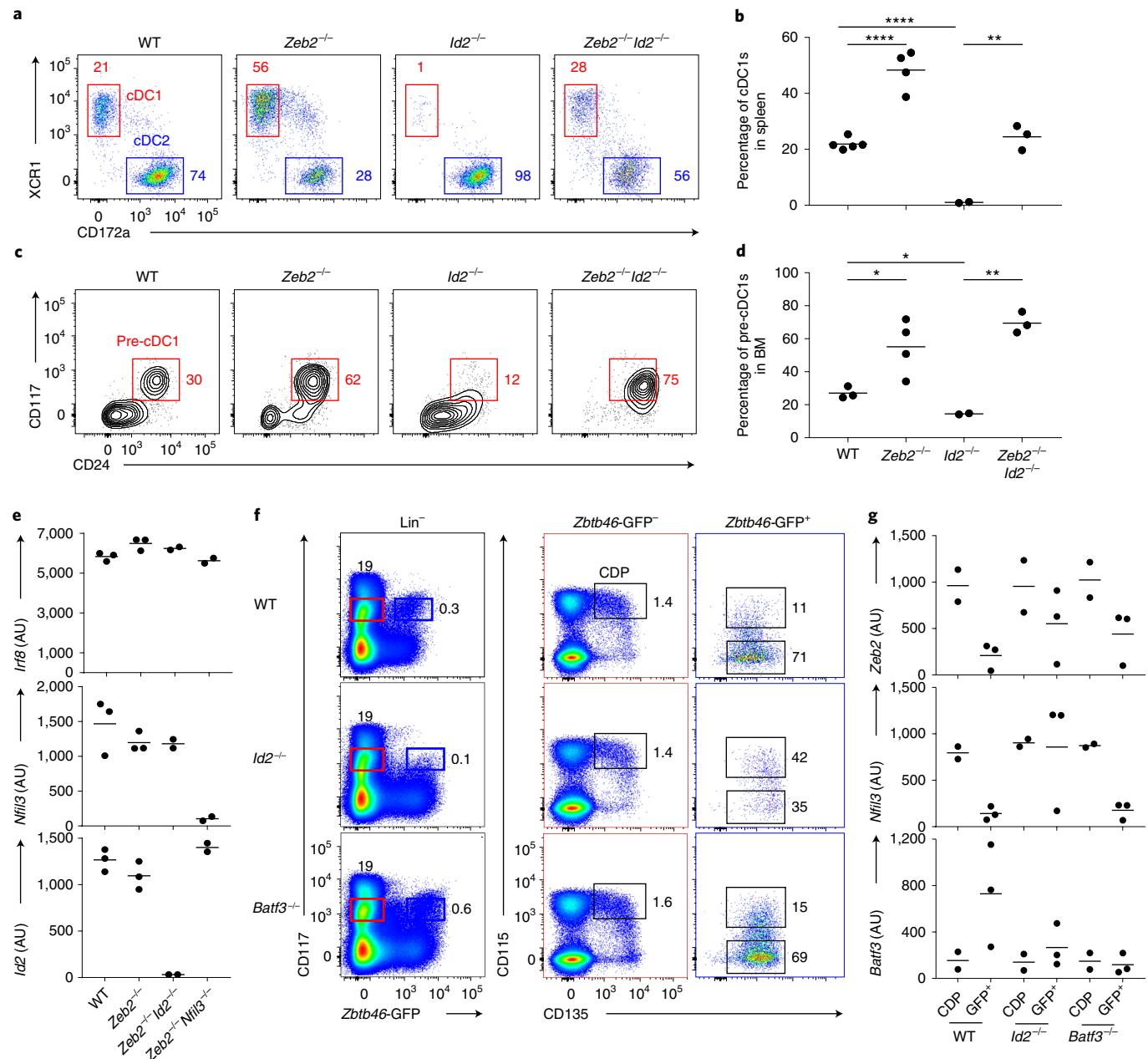


Fig. 6 | Expression of *Id2* and *Zeb2* is mutually repressive in the CDP. **a**, Splenic cDCs harvested from WT, *Zeb2*^{fl/fl}*Rosa26*^{cre-ERT2/+} (*Zeb2*^{-/-}), *Id2*^{fl/fl}*Rosa26*^{cre-ERT2/+} (*Id2*^{-/-}) and *Id2*^{fl/fl}*Zeb2*^{fl/fl}*Rosa26*^{cre-ERT2/cre-ERT2} (*Zeb2*^{-/-}*Id2*^{-/-}) mice were analyzed for cDC1 (red) and cDC2 (blue) frequency, gated as in Fig. 3e. Numbers are the percentage of cells in the indicated gates (data representing two independent experiments, $n=2$ for *Id2*^{-/-} mice, $n=3$ for *Zeb2*^{-/-}*Id2*^{-/-} mice, $n=4$ for *Zeb2*^{-/-} mice and $n=5$ for WT mice). **b**, Data from **a** are presented for individual mice. Small horizontal lines indicate the mean. **c**, BM from mice in **a** was analyzed for the frequency of pre-cDC1s (red). BM cells are pre-gated as Lin⁻Siglec-H⁻CD135⁺ (data representing two independent experiments, $n=2$ for *Id2*^{-/-} mice, $n=3$ for *Zeb2*^{-/-}*Id2*^{-/-} mice, $n=4$ for *Zeb2*^{-/-} mice and $n=5$ for WT mice). **d**, Data from **c** are presented for individual mice. Small horizontal lines indicate the mean. **e**, The expression of *Irf8*, *Nfil3* and *Id2* in splenic cDC1s from WT, *Zeb2*^{-/-}, *Zeb2*^{-/-}*Id2*^{-/-} and *Zeb2*^{-/-}*Nfil3*^{-/-} mice ($n=3$ for WT and *Zeb2*^{-/-} mice, $n=2$ for *Zeb2*^{-/-}*Id2*^{-/-} and *Zeb2*^{-/-}*Nfil3*^{-/-} mice). Small horizontal lines indicate the mean. **f**, BM from *Zbtb46*^{gfp/+} (WT), *Id2*^{-/-}*Zbtb46*^{gfp/gfp} (*Id2*^{-/-}) and *Batf3*^{-/-}*Zbtb46*^{gfp/gfp} (*Batf3*^{-/-}) mice was gated as Lin⁻ cells, and the CD117^{int}*Zbtb46*-GFP^{neg} (red) or CD117^{int}*Zbtb46*-GFP^{pos} (blue) cells were separately analyzed for CD115 and CD135 expression (data representing five independent experiments, $n=5$ mice). **g**, CDPs and *Zbtb46*-GFP^{pos} cells in **f** were sort purified and analyzed by gene expression microarray. The gene expression levels are shown for *Zeb2*, *Nfil3* and *Batf3* (data representing three independent experiments, $n=2$ for CDPs and $n=3$ for *Zbtb46*-GFP^{pos} cells). Small horizontal lines indicate the mean. Data are shown as the mean and the two-tailed, unpaired, Student's *t*-test was used to compare groups. * $P < 0.05$, ** $P < 0.01$; **** $P < 0.0001$; AU, arbitrary units.

Zeb2 and *Id2*, the *Rosa26*^{cre-ERT2} strain was crossed with *Zeb2*^{fl/fl}, *Id2*^{fl/fl} and *Zeb2*^{fl/fl}*Id2*^{fl/fl} mice to produce mice in which tamoxifen administration can conditionally inactivate ZEB2 (*Zeb2*^{-/-}), ID2 (*Id2*^{-/-}) or both (*Zeb2*^{-/-}*Id2*^{-/-}), respectively. First, pre-cDC1 specification and cDC1 development were evaluated in these mice (Fig. 6a–d).

Zeb2^{-/-} mice showed a twofold increase in cDC1s and pre-cDC1s compared with WT mice, similar to mice with *Zeb2* deficiency generated using poly(I:C) and *Mx1-Cre* (Fig. 5). *Id2*^{-/-} mice lacked splenic cDC1s, as expected¹⁸ and also lacked pre-cDC1s in BM. However, *Zeb2*^{-/-}*Id2*^{-/-} mice showed a restored development of

splenic cDC1s and BM pre-cDC1s (Fig. 6a–d). Moreover, similar results were obtained from in vitro Flt3L cultures of BM cells from these mice (see Supplementary Fig. 4a,b). In summary, for cDC1 development, *Zeb2* deficiency dominates over *Id2* deficiency in *Zeb2^{-/-}Id2^{-/-}* DKO mice, suggesting that, with respect to cDC1 specification, *Zeb2* genetically functions downstream of *Id2*.

***Zeb2* functions upstream of *Id2* with respect to *Id2* expression.** Next, a comparison was made of the transcriptional profiles of splenic cDC1s in WT, *Zeb2^{-/-}*, *Zeb2^{-/-Id2^{-/-}}* and *Nfil3^{-/-Zeb2^{-/-}}* mice using gene expression microarrays (Fig. 6e and see Supplementary Fig. 4c and Supplementary Table 5). As expected, the cDC1s from all genotypes expressed high *Irf8* and *Batf3* and low *Irf4* and *Tcf4* levels. *Nfil3* was highly expressed in cDC1s isolated from WT, *Zeb2^{-/-}* and *Zeb2^{-/-Id2^{-/-}}* mice and was absent in cDC1s isolated from *Nfil3^{-/-Zeb2^{-/-}}* mice, consistent with *Nfil3* genetically functioning upstream of both *Zeb2* and *Id2*. Furthermore, *Id2* was expressed at the expected high levels in cDC1s from WT and *Zeb2^{-/-}* mice, and absent in cDC1s from *Zeb2^{-/-Id2^{-/-}}* mice, in agreement with *Id2* genetically functioning upstream of *Zeb2*. Unexpectedly, *Id2* gene expression remained high in cDC1s from *Nfil3^{-/-Zeb2^{-/-}}* mice, despite the absence of *Nfil3* normally required for cDC1 specification. These results indicate that, in the absence of *Nfil3*, loss of *Zeb2* is sufficient for *Id2* induction, suggesting that *Zeb2* acts upstream of *Id2* with respect to *Id2* expression.

***Id2* and *Zeb2* expression are mutually repressive.** The above results indicate that *Zeb2* functions downstream of *Id2* with respect to cDC1 specification, because *Zeb2* deficiency can restore cDC1s in *Id2^{-/-}* mice, but acts upstream of *Id2* with respect to *Id2* gene expression. Thus, *Id2* appears to repress *Zeb2* expression, and *Zeb2* appears to repress *Id2* expression, to create a circuit of mutual repression in which *Nfil3* seems to initiate cDC1 specification by repressing *Zeb2*.

This model predicts that cDC1 specification in the CDP could occur in the absence of *Id2*, and that *Id2^{-/-}* pre-cDC1s would maintain *Zeb2* expression, unlike *Id2^{+/+}* pre-cDC1s. To test this, chimeric mice reconstituted with *Id2^{-/-Zbtb46^{gfp/gfp}}* BM (*Id2^{-/-Zbtb46^{gfp/gfp}}*

activating the +41-kb *Irf8* enhancer¹⁵. This enhancer is transiently active during cDC1 progenitor development, but is required for the development of both pre-cDC1s and cDC1s *in vivo*^{9,37}. This 454-basepair (bp) region contains six E-box motifs that are conserved between human and murine *Irf8* loci (Fig. 7a), and is known to bind E2-2 in human pDCs (see Supplementary Fig. 6)³⁸. Using the 454-bp region in a retroviral (RV) reporter system⁹, robust activity was found that was specific for pDCs, but not for cDC1s or cDC2s (Fig. 7b,c). The activity of three individual enhancer segments, each containing two E-box motifs, was also examined. Segments A and C showed reduced overall activity compared with the 454-bp enhancer, but retained pDC specificity, whereas the middle segment B retained overall activity, but reduced pDC specificity (Fig. 7b,c). Mutation of both E-boxes 1 and 2 in the 454-bp enhancer notably reduced enhancer activity in pDCs (see Supplementary Fig. 7a,b). Within segment A, mutation of either E-box alone reduced overall activity, whereas mutation of both E-boxes together completely extinguished activity (Fig. 7d and see Supplementary Fig. 7c). The most active segment, B, was also E-box dependent, showing reduced overall activity upon mutation of E-boxes 3 and 4 (Fig. 7e and see Supplementary Fig. 7d). These results indicate that the +41-kb *Irf8* enhancer activity relies on the redundant activity of the six E-box motifs contained within this 454-bp region. In agreement with the role of *Id2* in repressing E-proteins that act at E-box motifs, overexpression of retroviral *Id2* diminished +41-kb *Irf8* enhancer activity (Fig. 7f).

This suggests that *Id2* induction in the CDP can extinguish E-protein activity at the +41-kb *Irf8* enhancer, thereby imposing a requirement for a new enhancer in pre-cDC1s to maintain *Irf8* expression necessary for cDC1 development. To identify a potential enhancer, ATAC-seq was performed on the macrophage and DC precursor (MDP), CDP and pre-cDC1 progenitors, and a peak was found that indicated accessibility within the *Irf8* region only in the pre-cDC1s and the mature cDC1s, but not in the earlier MDPs, CDPs or mature cDC2s (Fig. 7g, red dashed line). This peak was located at +32 kb of the *Irf8* transcription start site and was shown to bind BATF3⁹. The induction of *Id2*, and the subsequent repression of *Zeb2*, thus forces a new requirement for *Batf3* in maintaining *Irf8* expression during cDC1 development.

Discussion

The present study resolves several long-standing puzzles with regard to cDC1 development. First, *Id2* was proposed to be required for cDC development by excluding pDC fate potential^{23,24}, but *Id2^{-/-}* mice lacked only cDC1s, and did not show the expected loss of all cDCs¹⁷. Second, cDC1s develop from CDPs that express *Irf8* independently of *Batf3*, yet later become dependent on *Batf3* to maintain *Irf8* expression. The basis for this switch from *Batf3*-independent to *Batf3*-dependent *Irf8* expression was unclear. Third, mature cDC1s do not express E-proteins or show +41-kb *Irf8* enhancer activity, yet their development requires both. These apparent inconsistencies all result from a cryptic stage in cDC1 development in which *Irf8* expression relies on the E-protein-dependent +41-kb *Irf8* enhancer. In the present study, this cryptic stage of development was examined to reveal the hierarchy of transcription factors governing cDC1 specification.

The results of the present study define a genetic hierarchy that unifies the actions of known transcription factors required for cDC1 development. cDC1s were known to require *Irf8*, *Batf3*, *Id2* and *Nfil3*, but how these factors interacted was unknown. *Zbtb46-GFP* was used to identify an earlier stage of cDC1 specification than previously described, which occurs within the CDP itself⁹. scRNA-seq of the CDP identified a cluster of cells defined by the expression pattern of *Nfil3*, *Id2* and *Zeb2*. Epistatic analysis revealed a genetic hierarchy in which *Nfil3* induces a transition from CDPs that express high levels of *Zeb2* and low levels of *Id2* to CDPs that express

***Id2* induction imposes a switch in *Irf8* enhancer usage during cDC1 development.** Data have revealed that E-proteins may be necessary for the sufficient induction of *Irf8* in the CDP by

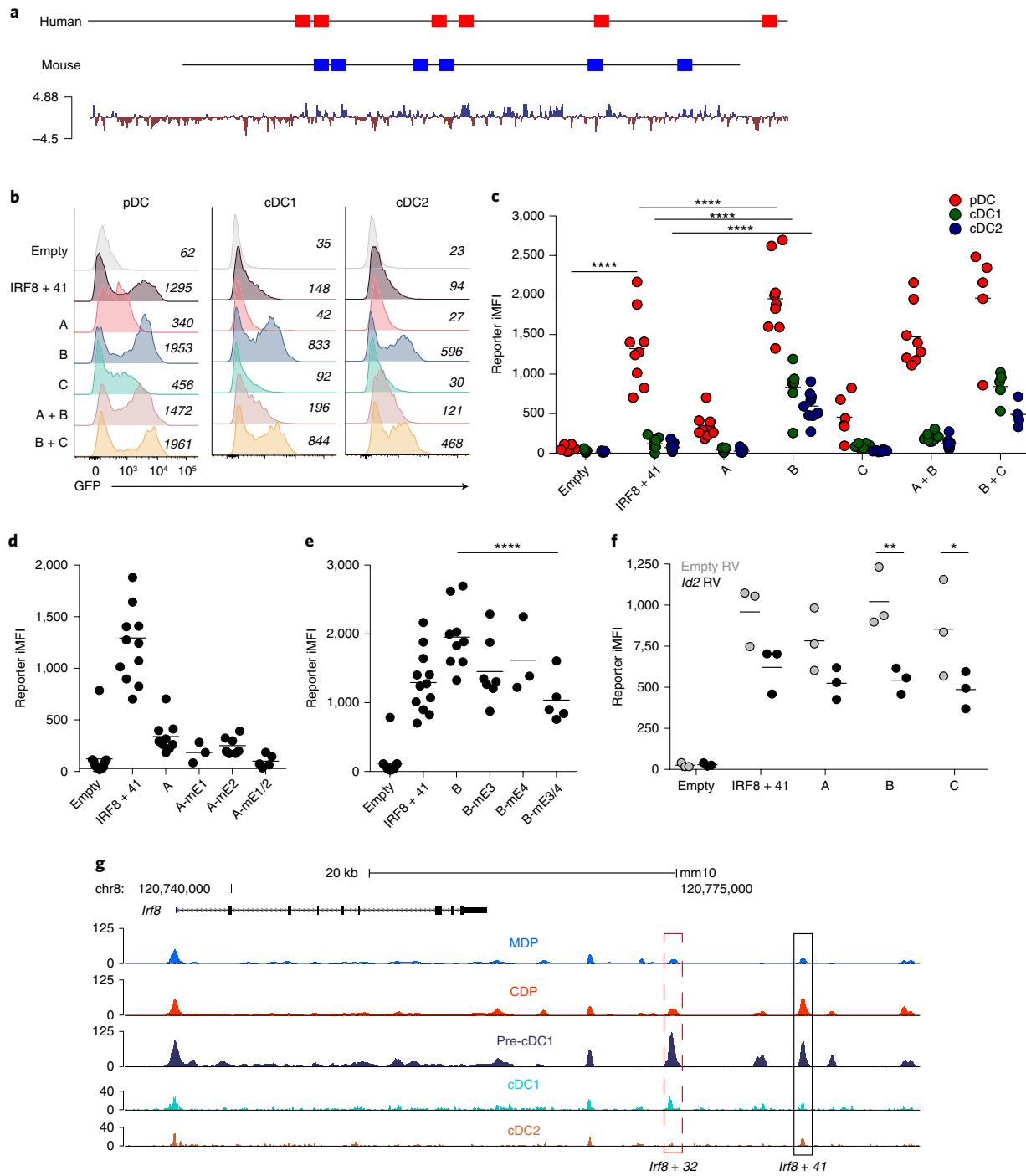


Fig. 7 | Id2 imposes a switch from the +41-kb *Irf8* enhancer to the +32-kb *Irf8* enhancer by reducing E-protein activity. **a**, Conservation of E-box motifs between human (red) and mouse (blue) loci within the +41-kb *Irf8* enhancer. **b**, GFP expression from RV reporters with (IRF8 + 41) or without (empty) the 454-bp, +41-kb enhancer, or with intact segment A (A), intact segment B (B), intact segment C (C) or intact segments A and B (A + B), or intact segments B and C (B + C), in pDCs, cDC1s and cDC2s, shown as histograms (data pooled from more than five independent experiments, $n > 5$). **c**, Data shown in **b** as integrated MFI (iMFI); data pooled from more than five independent experiments, $n > 5$. Small horizontal lines indicate the mean. **d**, GFP expression in pDCs of RV reporters without (empty) or with the 454-bp, +41-kb enhancer (IRF8 + 41), or with intact segment A (A), or with mutations in E-box 1 (A-m1), E-box 2 (A-m2) or both (A-m1/m2), shown as iMFI (data pooled from more than five independent experiments, $n > 5$). Small horizontal lines indicate the mean. **e**, GFP expression in pDCs of RV reporters without (empty) or with the 454-bp, +41-kb enhancer (IRF8 + 41), or with intact segment B (B), or with mutations in E-box 3 (B-m3), E-box 4 (B-m4) or both (B-m3/m4), shown as iMFI (data pooled from more than five independent experiments, $n > 5$). Small horizontal lines indicate the mean. **f**, GFP expression in WEHI-231 cells of RV reporters with (IRF8 + 41) or without (empty) the 454-bp, +41-kb enhancer, or with intact segment A (A), intact segment B (B), intact segment C (C) and cotransduced with either empty RV (gray) or *Id2* RV (purple), shown as iMFI (data pooled from three independent experiments, $n = 3$). Small horizontal lines indicate the mean. **g**, ATAC-seq was performed on the indicated progenitor or DC populations. The *Irf8* locus is shown, with the *Irf8*+41-kb enhancer region (black box) and the +32-kb enhancer region (dotted box) (representing three independent experiments and the Immunological Genome Project Open Chromatin Regions, $n=1$ biological replicate per population). Data are presented as the mean, and a one-way or two-way ANOVA was used to compare groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.0001$.

high levels of *Id2* and low levels of *Zeb2*. A circuit of mutual repression between *Zeb2* and *Id2* stabilizes these distinct states, such that repression of *Zeb2* by *Nfil3* is required to induce this transition. In *Zeb2^{hi}* and *Id2^{lo}* CDPs, *Irf8* expression is maintained by the +41-kb *Irf8* enhancer, which is dependent on E-proteins for activity. Upon *Id2* induction, E-protein activity is lost and *Irf8* expression becomes dependent on *Batf3*, acting at the +32-kb *Irf8* enhancer. It is currently unclear whether *Nfil3* directly represses *Zeb2* and whether *Zeb2* directly represses *Id2*, because there may be other factors in this proposed genetic circuit. *Nfil3* acts largely as a repressor^{21,39}, but may activate transcription in other contexts⁴⁰. Likewise, *Zeb2* has been suggested to directly repress *Id2* expression^{19,20}, although this has not been rigorously tested. *Nfil3*, *Zeb2* and *Id2* have also been shown to regulate innate lymphoid cell development⁴¹, but the mechanisms by which these transcription factors act in these cells have not been studied.

Although the present study seems to clarify several outstanding questions in cDC1 development, it may raise the possible necessity of a revised DC development scheme. A cDC1-specified stage was identified that occurs before the development of the pre-cDC1s. The cells in this stage express a high level of *Irf8*, consistent with the high level of *Irf8* in the CDP. Early expression of *Irf8* seems to correlate with commitment to the cDC1 lineage, as shown recently in a report in which IRF8 expression in human hemopoietic stem cells specifies the cDC1 lineage⁴². cDC1 specification may occur even earlier than this report suggests, but may rely on a minimum threshold of *Irf8* expression, and not simply early expression in the BM. The requirement of the +41-kb *Irf8* enhancer to increase IRF8 levels during the transition from the MDP to the CDP for subsequent cDC1 specification is consistent with this idea. A revised DC development model may require a deeper understanding of the relationship between IRF8 expression level and activity.

The present results also suggest that cDC1 development may be more closely related to pDC development than previously appreciated. The actions of the proposed genetic circuit on the +41-kb *Irf8* enhancer suggest that *Id2* extinguishes E-protein activity at the +41-kb *Irf8* enhancer and imposes a requirement for *Batf3* at the +32-kb *Irf8* enhancer. It is possible that pDCs and cDC1s share a common progenitor. The emergence of pDCs from myeloid or lymphoid BM progenitors is debated, because early studies suggested that pDCs can arise from both lymphoid and myeloid BM progenitors⁴³. However, two recent studies indicated that late pDC progenitors emerge from the common lymphoid progenitor and a ‘pre-pDC’ was described^{44,45}. As these studies did not perform lineage tracing for prior expression of myeloid markers, such as CD115, pDC progenitors conceivably could emerge in a series of stages, which include both myeloid and lymphoid features, as recently suggested⁴⁶. Resolution of whether pDCs and cDC1s share a common progenitor that has segregated from the cDC2 lineage, or simply share molecular transcriptional requirements, will require additional studies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at <https://doi.org/10.1038/s41590-019-0449-3>.

Received: 3 December 2018; Accepted: 17 June 2019;

Published online: 12 August 2019

References

- Saxena, M. & Bhardwaj, N. Re-emergence of dendritic cell vaccines for cancer treatment. *Trends Cancer* **4**, 119–137 (2018).
- Steinman, R. M. & Cohn, Z. A. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *J. Exp. Med.* **137**, 1142–1162 (1973).
- Cella, M. et al. Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon. *Nat. Med.* **5**, 919–923 (1999).
- Murphy, T. L. et al. Transcriptional control of dendritic cell development. *Annu. Rev. Immunol.* **34**, 93–119 (2016).
- Naik, S. H. et al. Development of plasmacytoid and conventional dendritic cell subtypes from single precursor cells derived in vitro and in vivo. *Nat. Immunol.* **8**, 1217–1226 (2007).
- Onai, N. et al. Identification of clonogenic common Flt3⁺ M-CSFR⁺ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. *Nat. Immunol.* **8**, 1207–1216 (2007).
- Liu, K. et al. In vivo analysis of dendritic cell development and homeostasis. *Science* **324**, 392–397 (2009).
- Schlitzer, A. et al. Identification of cDC1- and cDC2-committed DC progenitors reveals early lineage priming at the common DC progenitor stage in the bone marrow. *Nat. Immunol.* **16**, 718–728 (2015).
- Grajales-Reyes, G. E. et al. Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha⁺ conventional DC clonogenic progenitor. *Nat. Immunol.* **16**, 708–717 (2015).
- Lee, J. et al. Restricted dendritic cell and monocyte progenitors in human cord blood and bone marrow. *J. Exp. Med.* **212**, 385–399 (2015).
- Breton, G. et al. Human dendritic cells (DCs) are derived from distinct circulating precursors that are precommitted to become CD1c+ or CD141+DCs. *J. Exp. Med.* **213**, 2861–2870 (2016).
- See, P. et al. Mapping the human DC lineage through the integration of high-dimensional techniques. *Science* **356**, 1044–1044 (2017).
- Schiavoni, G. et al. ICSBP is essential for the development of mouse type I interferon-producing cells and for the generation and activation of CD8alpha⁺ dendritic cells. *J. Exp. Med.* **196**, 1415–1425 (2002).
- Tamura, T. et al. IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. *J. Immunol.* **174**, 2573–2581 (2005).
- Durai, V. et al. Cryptic activation of an *Irf8* enhancer governs cDC1 fate specification. *Nat. Immunol.* <https://doi.org/10.1038/s41590-019-0450-x> (2019).
- Kashiwada, M. et al. NFIL3/E4BP4 is a key transcription factor for CD8alpha⁺ dendritic cell development. *Blood* **117**, 6193–6197 (2011).
- Kusunoki, T. et al. TH2 dominance and defective development of a CD8+ dendritic cell subset in Id2-deficient mice. *J. Allergy Clin. Immunol.* **111**, 136–142 (2003).
- Hacker, C. et al. Transcriptional profiling identifies Id2 function in dendritic cell development. *Nat. Immunol.* **4**, 380–386 (2003).
- Scott, C. L. et al. The transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing Id2. *J. Exp. Med.* **213**, 897–911 (2016).
- Wu, X. et al. Transcription factor Zeb2 regulates commitment to plasmacytoid dendritic cell and monocyte fate. *Proc. Natl Acad. Sci. USA* **113**, 14775–14780 (2016).
- Cowell, I. G., Skinner, A. & Hurst, H. C. Transcriptional repression by a novel member of the bZIP family of transcription factors. *Mol. Cell Biol.* **12**, 3070–3077 (1992).
- Seillet, C. et al. CD8alpha⁺ DCs can be induced in the absence of transcription factors Id2, Nfil3, and Batf3. *Blood* **121**, 1574–1583 (2013).
- Ghosh, H. S. et al. Continuous expression of the transcription factor e2-2 maintains the cell fate of mature plasmacytoid dendritic cells. *Immunity* **33**, 905–916 (2010).
- Watowich, S. S. & Liu, Y. J. Mechanisms regulating dendritic cell specification and development. *Immunol. Rev.* **238**, 76–92 (2010).
- Grajowska, L. T. et al. Isoform-specific expression and feedback regulation of E protein TCF4 control dendritic cell lineage specification. *Immunity* **46**, 65–77 (2017).
- Satpathy, A. T. et al. Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *J. Exp. Med.* **209**, 1135–1152 (2012).
- McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. Preprint at <https://arxiv.org/abs/1802.03426> (2018).
- McInnes, L. et al. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* <https://doi.org/10.21105/JOSS.00861> (2018).
- Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
- Wu, X. et al. Bcl11a controls Flt3 expression in early hematopoietic progenitors and is required for pDC development in vivo. *PLoS ONE* **8**, e64800 (2013).
- Chopin, M. et al. Transcription factor PU.1 promotes conventional dendritic cell identity and function via induction of transcriptional regulator DC-SCRIPT. *Immunity* **50**, 77–90 (2019).
- Tussiwand, R. et al. Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. *Immunity* **42**, 916–928 (2015).
- Briseno, C. G. et al. Notch2-dependent DC2s mediate splenic germinal center responses. *Proc. Natl Acad. Sci. USA* **115**, 10726–10731 (2018).

34. Nishizaki, Y. et al. SIP1 expression patterns in brain investigated by generating a SIP1-EGFP reporter knock-in mouse. *Genesis* **52**, 56–67 (2014).
35. Jackson, J. T. et al. Id2 expression delineates differential checkpoints in the genetic program of CD8alpha+ and CD103+ dendritic cell lineages. *EMBO J.* **30**, 2690–2704 (2011).
36. Calero-Nieto, F. J. et al. Key regulators control distinct transcriptional programmes in blood progenitor and mast cells. *EMBO J.* **33**, 1212–1226 (2014).
37. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
38. Cisse, B. et al. Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135**, 37–48 (2008).
39. Cowell, I. G. & Hurst, H. C. Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. *Nucleic Acids Res.* **22**, 59–65 (1994).
40. Zhang, W. et al. Molecular cloning and characterization of NF-IL3A, a transcriptional activator of the human interleukin-3 promoter. *Mol. Cell Biol.* **15**, 6055–6063 (1995).
41. Ishizuka, I. E. et al. The innate lymphoid cell precursor. *Annu. Rev. Immunol.* **34**, 299–316 (2016).
42. Lee, J. et al. Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors. *Nat. Immunol.* **18**, 877–888 (2017).
43. Sathe, P. et al. Convergent differentiation: myeloid and lymphoid pathways to murine plasmacytoid dendritic cells. *Blood* **121**, 11–19 (2013).
44. Rodrigues, P. F. et al. Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. *Nat. Immunol.* **19**, 711–722 (2018).
45. Herman, J. S., Sagar & Grun, D. Fate ID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 (2018).
46. Ghosh, H. S. et al. ETO family protein Mtg16 regulates the balance of dendritic cell subsets by repressing Id2. *J. Exp. Med.* **211**, 1623–1635 (2014).
47. Heng, T. S. & Painter, M. W. Immunological genome project consortium the immunological genome project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
55. Bailey, T. L. et al. MEME suite: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).

Acknowledgements

We thank J. Chen and L. Goldstein for technical assistance. We thank the Genome Technology Access Center in the Department of Genetics at Washington University

School of Medicine for help with genomic analysis. The Center is partially supported by the National Cancer Institute's Cancer Center Support grant no. P30 CA91842 to the Siteman Cancer Center and by the Institute of Clinical and Translational Sciences/Clinical and Translational Science Award grant no. UL1TR000448 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. This publication is solely the responsibility of the authors and does not necessarily represent the official view of the NCRR or NIH. This work benefitted from data assembled by the ImmGen consortium⁴⁷. This work was supported by the Howard Hughes Medical Institute (K.M.M. and H.Y.C.), the National Science Foundation (grant no. DGE-1745038 to P.B.), the US NIH (grant nos. F30DK108498 to V.D., K08 CA23188-01 to A.T.S. and P50-HG007735 to H.Y.C.) and the Parker Institute for Cancer Immunotherapy (A.T.S. and H.Y.C.). A.T.S. was supported by a Career Award for Medical Scientists from the Burroughs Wellcome Fund.

Author contributions

P.B., X.H., T.T.L., T.L.M. and K.M.M. designed the study. P.B., X.H. and T.T.L. performed experiments related to analysis of immune populations, cell sorting and culture, gene microarray and generation of mice, with advice from C.G.B., G.E.G.-R., M.G. and S.K. P.B., M.N., Z.M. and A.S.S. performed and analyzed scRNA-seq data. V.D., J.M.G., A.T.S. and H.Y.C. performed ATAC-seq of DC progenitors. J.M.G. and A.T.S. performed computational analysis of ATAC-seq data. A.I. assisted with analysis of E-box motifs. P.B. performed all retroviral and reporter assays. P.B., X.H., T.T.L. and K.M.M. wrote the manuscript with advice from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41590-019-0449-3>.

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

Correspondence and requests for materials should be addressed to K.M.M.

Peer review information: Z. Fehervari was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Experimental model and subject details. Mice. WT C57BL6/J mice were obtained from the Jackson laboratory. *Zbtb46^{GFP/+}* mice were as described²⁶. *Nfil3^{-/-}* mice were from A. Look and Tak Mak⁴⁸. *Mx1-Cre*(B6.Cg-Tg(Mx1-cre)1Cgn/J) mice (stock no. 003556) and *Rosa26^{Cre-ERT2/cre-ERT2}*(B6.129-Gt(ROSA)26Sor^{tm1(Cre/ERT2)Tg/J}) mice (stock no. 008463) were obtained from the Jackson Laboratory. B6.SJL (B6.SJL-*Ptprc^aPep^c*/BoyJ) mice (strain code 564) were obtained from Charles River. ZEB2-EGFP fusion protein reporter (STOCK *Zfhxlb^{tm2.1YH}*) mice were derived from biological material provided by the RIKEN BioResource Center through the National BioResource Project of the Ministry of Education, Culture, Sports, Science and Technology, Japan. *SIP1^{fl/fl}*(*Zeb2^{fl/fl}*) mice were from Y. Higashi⁴⁹. For experiments shown in Fig. 6f,g, *Id2-CreERT2* mice (JAX stock no. 016222)⁵⁰ were bred to *Zbtb46^{GFP}* mice to generate *Id2^{creERT2/+}Zbtb46^{GFP/+}* mice. These mice were crossed to generate *Id2^{creERT2/creERT2}Zbtb46^{GFP/+}* or *Id2^{creERT2/creERT2}Zbtb46^{GFP/GFP}* mice. Livers from 1-day-old *Id2^{creERT2/creERT2}* pups were dispersed and cells injected into 4- to 6-week-old, lethally irradiated SJL WT mice (Charles River) and chimeras used 8 weeks after reconstitution. *Id2*-flox and *Id2*-IRES-GFP mice⁵¹ were generously donated by G. Belz. *Tcf3^{GFP/+}* mice were generated by crossing the *Tcf2e^a* allele (B6.129-Tcf3tm1Mbu/J JAX stock no. 028184) with Vav-iCre mice (JAX stock no. 008610).

All mice were generated, bred and maintained on the C57BL/6 background in the Washington University St Louis School of Medicine specific pathogen-free animal facility. Animals were housed in individually ventilated cages covered with autoclaved bedding and provided with nesting material for environmental enrichment. Up to five mice were housed per cage. Cages were changed once a week, and irradiated food and water in autoclaved bottles were provided freely. Animal manipulation was performed using standard protective procedures, including filtered air exchange systems, chlorine-based disinfection and personnel protective equipment, including gloves, gowns, shoe covers, facemasks and head caps. All animal studies followed institutional guidelines with protocols approved by the Animal Studies Committee at Washington University in St Louis.

Unless otherwise specified, experiments were performed with mice aged between 6 and 10 weeks. No differences were observed between male and female mice in any assays performed, and so mice of both genders were used interchangeably throughout the present study. Within individual experiments, mice used were age- and sex-matched littermates whenever possible.

Antibodies and flow cytometry. Cells were kept at 4 °C while being stained in phosphate-buffered saline supplemented with 0.5% BSA and 2 mM ethylenediaminetetraacetic acid in the presence of antibody-blocking CD16/32 (clone 2.4G2; BD 553142). All antibodies were used at a 1:200 dilution v/v, unless otherwise indicated.

The following antibodies were from BD: Brilliant Ultraviolet 395-anti-CD117 (clone 2B8, catalog no. 564011, 1:100 v/v), PE-CF594-anti-CD135 (clone A2F10.1, catalog no. 562537, 1:100 v/v), V500-anti-MHC-II (clone M5/114.15.2, catalog no. 742893), Brilliant Violet 421-anti-CCR9 (clone CW-1.2, catalog no. 565412, 1:100 v/v), Alexa Fluor 700-anti-Ly6C (clone AL-21, catalog no. 561237), Brilliant Violet 421-anti-CD127 (clone SB/199, catalog no. 562959, 1:100 v/v), biotin-anti-CD19 (clone 1D3, catalog no. 553784), BV510-anti-CD45R (clone RA3-6B2, catalog no. 563103) and PE-anti-CD90.1 (clone OX-7, catalog no. 554898). The following antibodies were from eBioscience: allophycocyanin-anti-CD317 (clone eBio927, catalog no. 17-3172-82, 1:100 v/v), PE-Cy7-anti-CD24 (clone M1/69, catalog no. 25-0242-82), peridinin chlorophyll protein-eFluor 710-anti-CD172a (clone P84, catalog no. 46-1721-82), peridinin chlorophyll protein-Cy5.5-anti-Siglec-H (clone eBio-440c, catalog no. 46-0333-82) and PE-anti-CD11c (clone N418, catalog no. 12-0114-82).

The following antibodies were from BioLegend: Brilliant Violet 711-anti-CD115 (clone AFS98, catalog no. 135515, 1:100 v/v), PE or Brilliant Violet 421-anti-XCR1 (clone ZET, catalog no. 148204 or 148216), Alexa Fluor 700 or APC/Cy7-anti-F4/80 (clone BM8, catalog no. 123130 or 123118, 1:100 v/v), PE-anti-CD45.2 (clone 104, catalog no. 109808), biotin or PE/Dazzle 594-anti-CD45R (clone RA3-6B2, catalog no. 103203 or 103258), biotin-anti-Ly6G (clone 1A8, catalog no. 127603), biotin-anti-Ter119 (clone TER-119, catalog no. 116204), biotin-anti-CD105 (clone MJ/718, catalog no. 120404), biotin-anti-NK1.1 (clone PK136, catalog no. 108704), biotin-anti-CD127 (clone A7R34, catalog no. 135006, 1:100 v/v), biotin-anti-Ly-6A/E (clone D7, catalog no. 108104) and PE-anti-human-CD4 (clone RPA-T4, catalog no. 300550, 1:50 v/v). The following antibodies were from Tonbo Bioscience: FITC-anti-CD45.1 (clone A20, catalog no. 35-0453-U500), biotin or APC-anti-CD3e (clone 145-2c11, catalog no. 30-0031-U500 or 20-0032-U100) and violetFluor 450-anti-MHC-II (I-A/I-E) (clone M5/114.15.2, catalog no. 75-5321-U100). The following antibody was from Invitrogen: allophycocyanin-eFluor 780-anti-CD11c (clone N418, catalog no. 47-0114-82). Cells were analyzed on a FACSCanto II or FACSAria Fusion flow cytometer (BD) and data were analyzed using FlowJo v.10 software (TreeStar).

Induced gene deletion. Conditional gene deletion in *Nfil3^{-/-}Zeb2^{fl/fl}Mx1-cre* (*Nfil3^{-/-}Zeb2^{fl/fl}*), *Zeb2^{fl/fl}Mx1-cre*-*Nfil3^{+/+}* (WT), *Zeb2^{fl/fl}Mx1-cre*-*Nfil3^{-/-}* (*Nfil3^{-/-}*) and *Zeb2^{fl/fl}Mx1-cre***Nfil3^{+/+}* (*Zeb2^{fl/fl}*) mice was induced by intraperitoneal injection of 150 µg poly(I:C) (Sigma-Aldrich; 1.0 mg ml⁻¹ stock solution dissolved in saline)

twice within 36–72 h. Gene deletion in WT, *Zeb2^{fl/fl}Rosa26^{Cre-ERT2}* (*Zeb2^{-/-}*), *Id2^{fl/fl}*-*Rosa26^{Cre-ERT2}* (*Id2^{-/-}*) and *Zeb2^{fl/fl}Id2^{fl/fl}Rosa26^{Cre-ERT2}* (*Zeb2^{-/-}Id2^{-/-}*) mice was induced by administration of tamoxifen citrate chow (Envigo) for 4–5 weeks. Mice were given up to 2 d of regular chow per week if notable weight loss was observed. After treatment, mice were rested on regular chow for 1 week before analysis.

Isolation and culture of BM progenitor cells and splenic DCs. BM progenitors and DCs were isolated as described⁵. For BM sorting experiments, BM was isolated and depleted of CD3-, CD19-, CD105-, Ter119- and, in some instances, Ly6G- and CD45R-expressing cells by staining with the corresponding biotinylated antibodies, followed by depletion with MagniSort Streptavidin Negative Selection Beads (Thermo Fisher). All remaining BM cells were then stained with fluorescent antibodies before sorting. MDPs were identified as Lin-CD117^{hi}CD135*CD115⁺ BM cells; CDPs were Lin-CD117^{int}CD135*CD115⁺MHC-II-CD11c⁺; pre-cDC1s were Lin-CD117^{int}CD135*CD115⁺MHC-II^{lo-int}CD11c⁺CD24⁺Siglec-H⁺ or Lin-CD117^{int}CD135*CD115⁺MHC-II^{lo-int}CD11c⁺Siglec-H-Zbtb46-GFP^{pos}; and pre-cDC2s were Lin-CD117^{int}CD135*CD115⁺MHC-II⁺CD11c⁺. For splenic sorting experiments, spleen was isolated and depleted of Ly6G-, B220- and CD3-expressing cells. The cDC2s were identified as Lin-CD45R-CD317-MHC-II⁺CD11c⁺CD172a⁺ cells. Cells were purified on a FACSaria Fusion into Iscove's modified Dulbecco's medium (IMDM) plus 10% FBS with 5% Flt3L conditioned medium. Sort purity of >95% was confirmed by postsort analysis before cells were used for further experiments. For experiments that included Flt3L cultures, sorted cells (1 × 10³ to 10 × 10³ cells per 200 µl complete IMDM) were cultured for 5 or 7 d at 37 °C with 5% Flt3L conditioned medium.

Expression microarray analysis. RNA was extracted using an RNaseq Micro Kit (Ambion) or a NucleoSpin RNA XS Kit (Machery-Nagel), and then was amplified using Ovation Pico WTA System (NuGEN) or WT Pico System (Affymetrix), and hybridized to GeneChip Mouse Gene 1.0 ST microarrays (Affymetrix) for 18 h at 45 °C in a GeneChip Hybridization Oven 640. The data were analyzed using the Affymetrix GeneChip Command Console. Microarray expression data was processed using Command Console (Affymetrix, Inc.) and the raw (.CEL) files generated were analyzed using Expression Console software with Affymetrix default Robust Multichip Analysis Gene analysis settings (Affymetrix, Inc.). Probe summarization (Robust Multichip Analysis), quality control analysis and probe annotation were performed according to recommended guidelines (Expression Console Software, Affymetrix, Inc.). Data were normalized by robust multiarray average summarization and underwent quartile normalization with ArrayStar software (DNASTAR). Unsupervised hierarchical clustering of differentially expressed genes was computed using ArrayStar (DNASTAR) with the Euclidean distance metric and centroid linkage method.

ScRNA-seq. One hundred thousand CDPs were sort purified as Live, Lin-CD127-CD117^{int}CD115⁺CD135⁺MHC-II⁺CD11c⁺ cells. Lineage included CD3, CD105, CD19, Ly6G and Ter119. Single-cell genes were measured with the Chromium system using Chromium Single Cell 3' Library and Gel Bead Kit v.2 (10x Genomics). Cell density and viability of sorted cells were determined by Vi-CELL XR cell counter (Beckman Coulter), and all processed samples had cell viability at >90%. The cell density was used to impute the volume of single-cell suspension needed in the reverse transcription master mix, to achieve ~6,000 cells per sample. After Gel Bead-in-Emulsion reverse transcription reaction and clean-up, a total of 12 cycles of PCR amplification was performed to obtain complementary DNAs. Libraries for RNA-seq were prepared following the manufacturer's user guide (10x Genomics), profiled using Bioanalyzer High Sensitivity DNA kit (Agilent Technologies) and quantified with Kapa Library Quantification Kit (Kapa Biosystems). Each scRNA-seq library was sequenced in one lane of HiSeq4000 (Illumina). Sequencing data were pooled from two runs of 4,796 and 4,758 individual cells. Run 1 had 2,354 median genes and 85,247 mean reads per cell. Run 2 had 2,247 median genes and 85,265 mean reads per cell. Sequencing was filtered and processed using the Seurat R toolkit⁵¹.

ATAC-seq. ATAC-seq of DC progenitors was performed using the Omni-ATAC protocol as previously described with minor modifications⁵². MDPs, CDPs and pre-cDC1s (10,000 in total) were sorted from BM as described above and lysed in ice-cold ATAC-RSB buffer containing 0.1% NP40, 0.1% Tween-20 and 0.01% digitonin. Cells were incubated at 4 °C for 3 min, then washed with ATAC-RSB buffer containing only 0.1% Tween-20. Nuclei were spun down by centrifugation and then incubated in 50 µl of transposition buffer (25 µl of 2× TD buffer, 22.5 µl distilled H₂O, 2.5 µl of Tn5 transposase (Nextera DNA Library Prep Kit, Illumina)) and incubated at 37 °C for 30 min. If 10,000 cells could not be obtained for a certain population, then the quantity of Tn5 transposase was titrated down proportionately to the number of cells obtained, but cells were still incubated in 50 µl total. Transposed DNA was purified with a DNA Clean & Concentrator kit (Zymo Research), eluted in 21 µl of elution buffer and stored at -20 °C until amplification. Three biological replicates for each cell population were obtained and sequenced. ATAC-seq libraries were prepared as previously described, barcoded and sequenced on an Illumina Nextseq.

Retroviral analysis of murine +41-kb *Irf8* enhancer. The 454-bp region of the +41-kb *Irf8* enhancer was cloned into hCD4 pA GFP-RV⁹. Each E-box motif (CANNTG) in the enhancer was mutated to a binding-site-free DNA sequence (AACTAC) determined by SiteOut⁵².

The primer sequences for the entire enhancer and the associated mutations are as follows:

For +41 kb *Irf8* enhancer: aaaagatctGATCTGGGTATGTGGGAAC and GA AAGAAGATCTGGGTATGT; for segment A: aaaagatctGATCTGGGTATGT GGGAAC and aaaaaggctTGTGCTAATTAAAGCCAAGGG; for segment B: a aaaggatccCTGATCCCAGATCCCATC and aaaaagcttGAGGAACCAACCAACT CAAGG; for segment C: aaaaggatccTCAAGTTTGAGGAAGAAG and aatctttt attttatcgatcaagCTTGACACTCTGGGAATAG; for segment A + B: GCGACGG TCGCGCAGGCTagaaaagatctGATCTGGGTATGTGGG and aatcttttatttcgata aaaaaggcttGAGGAACCAACCACT; for segment B + C: aaaaggatccCTGTACC CCAGATCCCATC and aatcttttatttcgatcaagCTTGACACTCTGGGAATAG; for mE1: GTGTCCTCACactacGGATCCCATATAAGGTTTATTTTAC and CCTTATATGGATCgttagtGTGAGAGACACAAAGGGTC; for mE2: GCC CAGGCCaactacTTCCCCCTGTACCCAG and GTACAGGGGGGAAGta gttGGGCCTGGCGATGTTCTG; for mE3: TCCTCCTCTGGTAGAGAAGAA GCTGCGGGCTGGGaactacCCGCACCCCTCCC and GGGGAGGGTGCAGGgt agttCCGAGGGCAGCTTCTACCAAGGAGG; for mE4: GCACCCCT CCCGGaactacTCTTCACCGTGCAGGTAG and CGCACGGTAAGAgtag ttCCGGGGAGGGTGCAGG; for mE5: GGCTGGAAGCCTTGAGTGGTGGTT CCTCaactacTCTTGGCACCTG and CAGGTGCCAAAGAgtagttGAGGAAC CACCACTCAAGGCTTCCAGCC; for mE6: ctacTCTTGGGaactacGGATGCG TCCTGTTAGGACC and CTTAACAGGACGCATCCgttagtCCCCAAAGAgtag ttGAGG; and for mE3/4: AGCTGCGGGCTGGGaactacCCGCACCCCTCCCCG GaactaCTTCACCGT and ACGGTGAAGAgtagttCCGGGGAGGGTGCAGGgt agttCCCGAGCCCGCAGCT.

RV vectors were transfected into Plat-E cells with TransIT-LTI (Mirus Bio) and viral supernatants were collected 2 d later. For retroviral analysis in Flt3L cultures, Lin⁻CD117^{high} BM cells were infected on day 1 after plating with the supernatants of transfected packaging cells, and concentrated by centrifugation with 2 µg ml⁻¹ polybrene by 'spin infection' at 2,250g for 60 min. Viral supernatant was replaced by complete IMDM + 5% Flt3L 1 d after transduction and the culture was read out on day 8. For analysis, the enhancer activity was quantitated using integrated mean fluorescence intensity (iMFI)^{53,54}.

For RV analysis in WEHI-231 cultures, WEHI-231 cells were infected on day 1, after plating with supernatants of transfected packaging cells, with the reporter constructs and either empty or *Id2* RV, and concentrated by centrifugation with 2 µg ml⁻¹ of polybrene by 'spin infection' at 2,250g for 60 min. Viral supernatant was replaced by complete IMDM 1 d after transduction and the culture was read out on day 3. For analysis, the enhancer activity was quantitated using iMFI in cells that were cointfected with either empty or *Id2* RV^{53,54}.

Analysis of E-box motifs in human +58-kb IRF8 enhancer. The occurrence of E-box motifs in the element +41-kb relative to the *Irf8* transcription start site was found

with the FIMO⁵⁵ motif-identification program at a *P*-value threshold of 1×10^{-3} with the E-box position weight matrix obtained for the E2-2 peaks of human pDCs³⁸. Human and mouse elements were aligned via Clustal Omega W.

Quantification and statistical analysis. Statistical analysis for scRNA-seq data is described above. Horizontal lines in figures indicate the mean. Results from independent experiments were pooled as indicated in figure legends. Data were analyzed using Prism (GraphPad), and either unpaired, two-tailed, Student's *t*-tests were used when comparing two groups or ordinary one-way or two-way analysis of variance (ANOVA) when comparing multiple groups. **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the corresponding author upon request. Microarrays are available on the Gene Expression Omnibus (GEO) database with the SuperSeries accession no. GSE123800. Data from Fig. 1 are available with accession number GSE123747, from Fig. 3 with accession numbers GSE123794 and GSE123796, and from Fig. 6 with accession numbers GSE123797 and GSE123799. Data from Supplementary Fig. 4 are available with accession number GSE123797 and from Supplementary Fig. 5 with accession numbers GSE123798 and GSE123799. The scRNA-seq data are available with the accession number GSE132770, and are used in Fig. 2 and Supplementary Fig. 1b. The ATAC-seq data of DC progenitors are available with the accession number GSE132240 and are used in Fig. 7. Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, K.M.M.

References

48. Kamizono, S. et al. Nfil3/E4bp4 is required for the development and maturation of NK cells in vivo. *J. Exp. Med.* **206**, 2977–2986 (2009).
49. Higashi, Y. et al. Generation of the floxed allele of the SIP1 (Smad-interacting protein 1) gene for Cre-mediated conditional knockout in the mouse. *Genesis* **32**, 82–84 (2002).
50. Rawlins, E. L. et al. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development* **136**, 3741–3745 (2009).
51. Butler, A. et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
52. Estrada, J. et al. SiteOut: an online tool to design binding site-free DNA sequences. *PLoS ONE* **11**, e0151740 (2016).
53. Shoohtari, P. et al. Correlation analysis of intracellular and secreted cytokines via the generalized integrated mean fluorescence intensity. *Cytometry A* **77**, 873–880 (2010).
54. Darrah, P. A. et al. Multifunctional TH1 cells define a correlate of vaccine-mediated protection against *Leishmania major*. *Nat. Med.* **13**, 843–850 (2007).

Corresponding author(s): Kenneth M. Murphy

Last updated by author(s): Jun 4, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Flow cytometry data was collected using FACSDiva software and FlowJo v.10. Single-cell RNA sequencing was analyzed using the R package Seurat. Microarray data was analyzed using ArrayStar14 (DNASTAR). ATAC-seq was analyzed using UCSC Genome Browser. Statistical analysis was performed using PRISM (Graphpad).

Data analysis

FlowJo v10, ArrayStar 14, Seqpurge, Bowtie2, ChromVARMotifs, Homer

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Microarrays are available on the GEO database with the SuperSeries accession number GSE123800. Data from Fig. 1 is available with accession number GSE123747; from Fig. 3 are GSE123794 and GSE123796; from Fig. 6 are GSE123797 and GSE123799. Data from Supplementary Fig. 4 is GSE123797 and from Supplementary Fig. 5 are GSE123798 and GSE123799. The single-cell RNA-sequencing data is available on the GEO database with the following accession number: and is utilized in Fig. 2 and Supplementary Fig. 1b. The ATAC-seq data of DC progenitors is available on the GEO database with the following accession number: and is utilized in Fig. 7. All the data will be available June 23, 2019. We have submitted the GEO submissions for the single-cell RNA-sequencing and ATAC-seq as on June 3, 2019, but have not yet received accession numbers, but will provide them as soon as possible.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No formal power calculations were performed, however, the sample size of each group was robust and data was routinely collected across at least three independent replicates for each assay with an end result of never having fewer than three mice per group.
Data exclusions	No data was excluded from analyses.
Replication	All experiments were replicated at least three different times with completely independent sets of mice that were the result of independent crosses. All replicates confirmed the data.
Randomization	No formal randomization was performed as comparisons were done across mice of different genotypes, not across mice of the same genotypes receiving different treatments.
Blinding	Investigators were blinded to the genotype of the mice during sample preparation and data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

Antibodies

Antibodies used	<p>Cells were kept at 4°C while being stained in PBS supplemented with 0.5% BSA and 2mM EDTA in the presence of antibody blocking CD16/32 (clone 2.4G2; BD 553142). All antibodies were used at a 1:200 dilution vol/vol (v/v), unless otherwise indicated.</p> <p>The following antibodies were from BD: Brilliant Ultraviolet 395–anti-CD117 (clone 2B8, catalog number 564011, 1:100 v/v), PE-CF594–anti-CD135 (clone A2F10.1, catalog number 562537, 1:100 v/v), V500–anti-MHC-II (clone M5/114.15.2, catalog number 742893), Brilliant Violet 421–anti-CCR9 (clone CW-1.2, catalog number 565412, 1:100 v/v), Alexa Fluor 700–anti-Ly6C (clone AL-21, catalog number 561237), Brilliant Violet 421–anti-CD127 (clone SB/199, catalog number 562959, 1:100 v/v), biotin–anti-CD19 (clone 1D3, catalog number 553784), BV510–anti-CD45R (clone RA3-6B2, catalog number 563103), PE–anti-CD90.1 (clone OX-7, catalog number 554898). The following antibodies were from eBioscience: allophycocyanin–anti-CD317 (clone eBio927, catalog number 17-3172-82, 1:100 v/v), PE–Cy7–anti-CD24 (clone M1/69, catalog number 25-0242-82), peridinin chlorophyll protein (PerCP)–eFluor 710–anti-CD172a (clone P84, catalog number 46-1721-82), PerCP–Cy5.5–anti-SiglecH (clone eBio-440c, catalog number 46-0333-82), PE–anti-CD11c (clone N418, catalog number 12-0114-82).</p> <p>The following antibodies were from BioLegend: Brilliant Violet 711–anti-CD115 (clone AFS98, catalog number 135515, 1:100 v/v), PE or Brilliant Violet 421–anti-XCR1 (clone ZET, catalog number 148204 or 148216), Alexa Flour 700 or APC/Cy7–anti-F4/80 (clone BM8, catalog number 123130 or 123118, 1:100 v/v), PE–anti-CD45.2 (clone 104, catalog number 109808), biotin or PE/Dazzle 594–anti-CD45R (clone RA3-6B2, catalog number 103203 or 103258), biotin–anti-Ly6G (clone 1A8, catalog number 127603), biotin–anti-Ter119 (clone TER-119, catalog number 116204), biotin–anti-CD105 (clone MJ/718, catalog number 120404), biotin–anti-NK1.1 (clone PK136, catalog number 108704), biotin–anti-CD127 (clone A7R34, catalog number 135006, 1:100 v/v), biotin–anti-Ly-6A/E (clone D7, catalog number 108104), PE–anti-human-CD4 (clone RPA-T4, catalog number 300550, 1:50 v/v). The following antibodies were from Tonbo Bioscience: FITC–anti-CD45.1 (clone A20, catalog number 35-0453-U500), biotin or APC–anti-CD3e (clone 145-2c11, catalog number 30-0031-U500 or 20-0032-U100), violetFluor 450–anti-MHC Class II (I-A/I-E) (clone M5/114.15.2, catalog number 75-5321-U100). The following antibodies were from Invitrogen: allophycocyanin–eFluor 780–anti-CD11c (clone N418, catalog number 47-0114-82).</p>
-----------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Validation

All these antibodies are reactive against mouse and have been routinely used and cited by numerous publications in our field. These antibodies have been validated by the manufacturer, and the catalogs of each antibody lists the specific citations in which these antibodies were used and tested.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	ATCC (WEHI-231 and PlatE)
Authentication	Microscopic inspections (these cells are easily distinguished based on morphology)
Mycoplasma contamination	These cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Unless otherwise specified, experiments were performed with mice between 6 and 10 weeks of age. No differences were observed between male and female mice in any assays performed and so mice of both genders were used interchangeably throughout this study. Within individual experiments, mice used were age- and sex-matched littermates whenever possible. WT C57BL6/J mice were obtained from The Jackson laboratory. Zbtb46gfp/+ mice were described ²⁵ . Nfil3−/− mice were from A. Look and Tak Mak ⁴² . Mx1-Cre [B6.Cg-Tg(Mx1-cre)1Cgn/J] mice (stock no. 003556), and Rosa26Cre/Cre [B6.129-Gt(ROSA)26Sortm1(creERT2)Tyr/J] mice (stock no. 008463) were obtained from The Jackson Laboratory. B6.SJL (B6.SJL-Ptpca Pepcb /BoyJ) mice (strain code 564), were obtained from Charles River. ZEB2-EGFP fusion protein reporter (STOCK Zfhxlbtm2.1Yhi) mice ²⁹ were derived from biological material provided by the RIKEN BioResource Center through the National BioResource Project of the Ministry of Education, Culture, Sports, Science and Technology, Japan. SIP1flox(ex7) (Zeb2f/f) were from Y. Higashi ⁴³ . For experiments shown in Fig. 6d,e, Id2-CreERT2 mice (JAX stock #016222)44 were bred to Zbtb46gfp mice to generate Id2creERT2/+Zbtb46gfp/+ mice. These mice were crossed to generate Id2creERT2/creERT2 Zbtb46gfp/+ or gfp/gfp mice. Livers from day 1 old Id2creERT2/creERT2 pups were dispersed and cells injected into 4-6 week old lethally irradiated SJL WT mice (Charles Rivers) and chimeras used eight weeks after reconstitution. Id2-flox and Id2-IRES-GFP mice ³⁰ were generously donated by G. Belz. Tcf3GFP/+ were generated by crossing the Tcf3tm1Mbu/J JAX stock #028184) with Vav-iCre mice (JAX stock #008610).
Wild animals	Study did not involve wild animals.
Field-collected samples	Study did not involve field-collected animals.
Ethics oversight	All mice were generated, bred, and maintained on the C57BL/6 background in the Washington University in St. Louis School of Medicine specific pathogen-free animal facility. Animals were housed in individually ventilated cages covered with autoclaved bedding and provided with nesting material for environmental enrichment. Up to five mice were housed per cage. Cages were changed once a week, and irradiated food and water in autoclaved bottles were provided ad libitum. Animal manipulation was performed using standard protective procedures, including filtered air exchange systems, chlorine-based disinfection, and personnel protective equipment including gloves, gowns, shoe covers, face masks, and head caps. All animal studies followed institutional guidelines with protocols approved by the Animal Studies Committee at Washington University in St. Louis.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- 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).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Briefly, spleens and inguinal skin-draining LNs were minced and digested in 5 mL of Iscove's modified Dulbecco's media (IMDM) + 10% FCS (cIMDM) with 250 µg/mL collagenase B (Roche) and 30 U/mL DNase I (Sigma-Aldrich) for 45 min at 37°C with stirring. After digestion was complete, single cell suspensions from all organs were passed through 70-µm strainers and red blood cells were lysed with ammonium chloride-potassium bicarbonate (ACK) lysis buffer. Cells were subsequently counted with a Vi-CELL analyzer (Beckman Coulter) and 3-5x10⁶ cells were used per antibody staining reaction. Bone marrow (BM) was harvested from the femur, tibia, and pelvis of mice. Bones were collected and fragmented by mortar and

pestle in MACS buffer, and debris was removed by passing cells through a 70- μ m strainer. Red blood cells were lysed with ACK lysis buffer and cells were subsequently counted on a Vi-CELL analyzer (Beckman Coulter). 3-10x10⁶ were used per antibody staining reaction. For BM culture experiments, bulk BM cells were cultured at 37°C in 4 mL total volume of cIMDM supplemented with 100 ng/mL Flt3L (Peprotech) for eight days before further analysis. Briefly, spleens and inguinal skin-draining LNs were minced and digested in 5 mL of Iscove's modified Dulbecco's media (IMDM) + 10% FCS (cIMDM) with 250 μ g/mL collagenase B (Roche) and 30 U/mL DNaseI (Sigma-Aldrich) for 45 min at 37°C with stirring. Lungs were minced and digested in 5 mL of cIMDM with 4 mg/mL collagenase D (Roche) and 30 U/mL DNaseI (Sigma-Aldrich) for 1.5 hours at 37°C with stirring. After digestion was complete, single cell suspensions from all organs were passed through 70- μ m strainers and red blood cells were lysed with ammonium chloride-potassium bicarbonate (ACK) lysis buffer. Cells were subsequently counted with a Vi-CELL analyzer (Beckman Coulter) and 3-5x10⁶ cells were used per antibody staining reaction. For peritoneal cell analysis, 5 mL of MACS buffer (DPBS + 0.5% BSA + 2mM EDTA) was injected into the peritoneum of mice using a 27 g needle. After injection the mice were shaken gently to dislodge peritoneal cells. A 25 g needle was then used to collect the peritoneal fluid. Cells were ACK lysed and counted as described above.

Instrument

BD FACSCanto II or BD FACSaria Fusion

Software

FlowJo v10

Cell population abundance

A FACSaria Fusion was used for sorting and cells were sorted into cIMDM. Sort purity of >95% was confirmed by post-sort analysis before cells were used for experiments.

Gating strategy

Gating strategies for all cell populations are depicted within the paper. For FSC/SSC populations were gated as within the lymphocyte gate as traditionally has been done. Singlets were gated based on FSC-A/FSC-W profile as traditionally done.

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