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

Prachi Bagadia1,13, Xiao Huang1,13, Tian-Tian Liu1,2,13, Vivek Durai1, Gary E. Grajales-Reyes1, Maximilian Nitschke2, Zora Modrusan4, Jeffrey M. Granja5,6,7, Ansuman T. Satpathy5,8, Carlos G. Briseño9, Marco Gargaro10, Arifumi Iwata11, Sunkyung Kim1, Howard Y. Chang9,5,12, Andrey S. Shaw3, Theresa L. Murphy1 and Kenneth M. Murphy1,2,*

Classical type 1 dendritic cells (cDCs) 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 Zeb20 and Id22 CDPs to Zeb24 and Id22 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 therapy. Dendritic cells (DCs) are an immune lineage encompassing classical DCs (cDCs) and plasmacytoid DCs (pDCs). The cDCs comprise two branches, cDC1 and cDC2, which exert distinct functions in vivo and rely on different transcriptional programs. pDCs and cDCs can both arise from the CDP1, cDC progenitors (pre-cDCs) include clonogenic populations separately committed to cDC1 or cDC2 lineages. Similar progenitors have been confirmed in human DC development10–12. 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 development13,14, 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 expression, 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 vivo. 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 16,17). Nfil3, a basic leucine zipper transcriptional repressor18, is expressed in cDC1s and is required for cDC1 development19,20, but how it functions is unknown4,15. Id2 is a known inhibitor of E-proteins, is expressed in both cDC1 and cDC2, but is required only for cDC1 development21,12. ID2 may exclude pDC fate by blocking the activity of E-proteins, particularly E2-2 (Tcfa), required for pDCs22–25. 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 transcription19,26. How these factors precisely interact and at what stage they influence cDC1 specification is unknown.

In the present study, single-cell RNA-sequencing (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...
expression from being Batf3 independent to being Batf3 dependent. The CDP originates in a Zeb2lo and Id2lo 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 Zeb2lo and Id2lo state. A circuit of mutual Zeb2–Id2 repression serves to stabilize states before and after this transition. Id2 expression in the specified pre-cDC inhibits E-proteins, block- ing 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–CD117intCD135–CD115+ BM population and was observed to be, although not defined as, largely negative for major histocompatibility complex II (MHC-II) and CD11c expression. Subsequently, pre-cDC1 and pre-cDC2 progenitors were identified as arising from the CDP but were not contained within the CDPP3. Pre-cDC1s were defined as Lin CD117+CD135+CD11c–MHC-II–lo cells and were largely CD115+. They can be defined using two methods, relying either on Zbtb46–GFP expression in Zbtb46GFP+ reporter mice, or on conventional surface markers (Fig. 1a)45. 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, 3–10% of CDPPs, defined on the strict exclusion of CD11c- and MHC-II-expressing cells, are Zbtb46–GFP+ (Fig. 1b). These Zbtb46–GFP+ pre-cDCs 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– CDPPs, which produced cells from all three DC lineages (Fig. 1c and see Supplementary Fig. 1a).

The transcriptional profile of these Zbtb46–GFP+ CDPPs suggests that they represent an intermediate population between a non-specified CDP, the Zbtb46–GFP– CDP and the pre-cDC1 (Fig. 1d,e). For example, genes that had their expression changed more than eightfold between the Zbtb46–GFP+ CDPPs and the pre-cDC1 were considered. For such genes, their expression in Zbtb46–GFP+ CDPPs was consistently intermediate between their expression in Zbtb46–GFP– CDPPs and pre-cDC1s (Fig. 1d,e and see Supplementary Tables 1 and 2). Id2 expression in Zbtb46–GFP+ CDPPs was increased by 34-fold in pre-cDC1s, but only by 15-fold in Zbtb46–GFP– CDPPs. Likewise, Zeb2 expression in Zbtb46–GFP+ CDPPs was reduced by 9-fold in pre-cDC1s, but only by 3.6-fold in Zbtb46–GFP– CDPPs. As expected, the Zbtb46–GFP+ CDPPs were segregated away from the pre-cDC2s (Fig. 1e). Thus, these results indicate that Zbtb46–GFP+ CDPPs 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 CDPPs defined as Lin CD127–CD117+CD115+CD135–MHC-II CD11c+ cells (Fig. 2a) on the 10x Genomics platform to assay for unrecognized heterogene-

ity within this population. Uniform Manifold Approximation and Projection (UMAP) analysis21–23 identified eight connected clusters (Fig. 2b,c). Although it was possible to identify genes that were specifically enriched in certain clusters, others such as Tcf4 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 Cc5d 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 coordinate 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 (Zeb2EGFP+) or an Id2–IRES–GFP cassette (Id2EGFP+) 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 Zeb2EGFP mice, 90% of CDPPs 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 Id2EGFP mice, 94% of CDPPs expressed low levels of 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 Zeb2EGFP and Id2EGFP reporter lines confirm the existence of ZEB2–EGFPlo and Id2–GFP+ cells within the CDP, as predicted by scRNA-seq.

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

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

**Nfi3 is required for cDC1 specification within the CDP.** Nfi3 is required for cDC1 development7, but its mechanism and
Fig. 2 | Single-cell RNA transcriptome analysis of CDPs. a, CDPs gated as Live, Lin'CD127'CD117'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.
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\(^{+/-}\) 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\(^{+/-}\) mice, although they do develop normally in Batf3\(^{-/-}\)Zbtb46\(^{+/-}\) mice as previously described (see Supplementary...
Within the CDP, cDC1-specified cells can be identified as Zbtb46–GFPpos cells that comprise 5% of the CDP (Fig. 4a,b). However, these cells are also absent in Nfil3−/− mice.

In Nfil3−/− Zeb2egfp reporter mice, cDC1-specified cells are identified as CD117+ZEB2–EGFP pos cells, which include pre-cDC1s and ZEB2–EGFP pos CDPs, and comprise approximately 6% of Lin–CD135+ cDC1-specified cells.
Fig. 5 | Zeb2 is downstream of Nfil3 in cDC1 development. a, Splenic cDCs from Nfil3−/−Zeb2f/fMx1-cre− (WT), Zeb2f/fMx1-cre− (Zeb2−/−), Nfil3−/− (Nfil3−/−) and Nfil3−/−Zeb2f/fMx1-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 vivo 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 were 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−/− mice. In Nfil3+/−Zeb2+/− reporter mice, cDC1-specified CDPs can be identified as ZEB2−EGFP+ cells, which comprise 7% of CDPs (Fig. 4c,d), which again are absent in Nfil3−/−Zeb2+/− mice. Finally, in Nfil3+/−Id2+/− reporter mice, cDC1-specified cells can be identified as CD117int/Id2−GFP+ cells, which include pre-cDC1s and Id2–GFP+ CDPs, and comprise approximately 2% of Lin−CD135+ BM (Fig. 4e,f). However, these cells are absent in Nfil3−/−Id2+/− mice. Furthermore, cDC1-specified CDPs can be identified as Id2−GFP+ cells, which comprise 7% of CDPs (Fig. 4e,f), but are absent in Nfil3−/−Id2+/− mice. In summary, Nfil3 is required for the appearance of all cDC1-specified progenitors identified by Zebt46−/−GFPR, 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 Zeb2f/fMx1-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+/−Zeb2f/fMx1-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 study. Furthermore, Nfil3−/− mice lacked cDC1s in the spleen, as previously reported. 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, but no mechanism has been established. To evaluate the genetic interaction between...
Zeb2 and Id2, the Zeb2\(^{+/–}\)Id2\(^{+/–}\) strain was crossed with Zeb2\(^{+/–}\), Id2\(^{+/–}\) and Zeb2\(^{+/–}\)Id2\(^{+/–}\) 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 expected\(^{18}\) 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 If4 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−/−Id2−/− 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−/− BM (Id2−/−Zbtb46-GFP−/−) were used. First it was confirmed that splenic pDCs and cDC2s develop normally in Id2−/−Zbtb46-GFP−/− chimeras (see Supplementary Fig. 5a). In addition, Id2−/−Zbtb46-GFP−/− chimeras are essentially identical in gene expression to Id2−/−Zbtb46-GFP−/− cDC2s (see Supplementary Fig. 5b and Supplementary Table 6). Furthermore, unspecified CPs, defined as Lin−CD117 intCD135−CD115−CD135+CD115−CD135−CD135− 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, CPs 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 BATF3. 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 potential18,21, but Id2−/− mice lacked only cDC1s, and did not show the expected loss of all cDCs15. Second, cDC1 develop from CPs 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 that was previously described, which occurs within the cDC itself. scRNA-seq of the cDC identified a cluster of cells defined by the expression pattern of Nfil3, Id2 and Zeb2. Episodic analysis revealed a genetic hierarchy in which Nfil3 induces a transition from CPs that express high levels of Zeb2 and low levels of Id2 to CPs that express...
**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- and Id2-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 repressors, but may activate transcription in other contexts. Likewise, Zeb2 has been suggested to directly repress Id2 expression, although this has not been rigorously tested. Nfil3, Zeb2 and Id2 have also been shown to regulate innate lymphoid cell development, 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-specific 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 lineage. 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 progenitors. However, two recent studies indicated that late pDC progenitors emerge from the common lymphoid progenitor and a ‘pre-pDC’ was described, although this has not been rigorously tested. Nfil3, Zeb2 and Id2 have also been shown to regulate innate lymphoid cell development, but the mechanisms by which these transcription factors act in these cells have not been studied.

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

1. Saxena, M. & Bhardwaj, N. Re-emergence of dendritic cell vaccines for cancer treatment. Trends Cancer 4, 119–137 (2018).
2. 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).
3. 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).
4. Murphy, T. L. et al. Transcriptional control of dendritic cell development. Annu. Rev. Immunol. 34, 93–119 (2016).
5. 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).
6. Omai, 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).
7. Liu, K. et al. In vivo analysis of dendritic cell development and homeostasis. Science 324, 392–397 (2009).
8. 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).
9. Grajeles-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).
10. Lee, J. et al. Restricted dendritic cell and monocyte progenitors in human cord blood and bone marrow. J. Exp. Med. 212, 385–399 (2015).
11. Breton, G. et al. Human dendritic cell (DC) subsets are derived from distinct circulating precursors that are precommitted to become CD1c+ or CD14+DCs. J. Exp. Med. 213, 2861–2870 (2016).
12. See, P. et al. Mapping the human DC lineage through the integration of high-dimensional techniques. Science 356, 1044–1044 (2017).
13. Schmickel, R. 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).
14. 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).
15. Durá, 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).
16. Kashiwada, M. et al. NFIL3/E4BP4 is a key transcription factor for CD8alpha+ dendritic cell development. Blood 117, 6193–6197 (2011).
17. Kusunoki, T. et al. TH2 dominance and defective development of a CD8alpha+ dendritic cell subset in Id2-deficient mice. J. Allergy Clin. Immunol. 111, 136–142 (2003).
18. Hacker, C. et al. Transcriptional profiling identifies Id2 function in dendritic cell development. Nat. Immunol. 4, 380–386 (2003).
19. 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).
20. Wu, X. et al. Transcription factor Zeb2 regulates commitment to plasmacytoid dendritic cell and monocyte fate. Proc. Natl Acad. Sci. USA 113, 14773–14780 (2016).
21. 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).
22. Seillet, C. et al. CD8alpha+ DCs can be induced in the absence of transcription factors Id2, Nfil3, and Batf3. Blood 121, 1574–1583 (2013).
23. 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).
24. Watowich, S. S. & Liu, Y. I. Mechanisms regulating dendritic cell specification and development. Immunol. Rev. 238, 76–92 (2010).
25. Grajewska, L. T. et al. Isoform-specific expression and feedback regulation of E protein TC4 control dendritic cell lineage specification. Immunity 46, 671–677 (2017).
26. 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).
27. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. Preprint at https://arxiv.org/abs/1802.03426 (2018).
28. McInnes, L. et al. UMAP: Uniform Manifold Approximation and Projection. Journal of Open Source Software https://doi.org/10.21105/joss.00861 (2018).
29. Recht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. 37, 38–44 (2019).
30. 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).
31. Chopin, M. et al. Transcriptional factor PU.1 promotes conventional dendritic cell identity and function via induction of transcriptional regulator DC-SCRIPT. Immunity 50, 77–99 (2019).
32. Tussiwand, R. et al. Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. Immunity 42, 916–928 (2015).
33. Briscoe, C. G. et al. Notch2-dependent DC2s mediate splenic germinal center responses. Proc. Natl Acad. Sci. USA 115, 10726–10731 (2018).
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 Institutes 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 consortium. 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.

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.
Methods

Experimental model and subject details. Mice. WT C57BL/6J mice were obtained from the Jackson Laboratory. Zeb2<sup>−/−</sup> mice were as described<sup>1</sup>. Nfil3<sup>−/−</sup> mice were from A. Look and Tak Mak<sup>48</sup>. Mxi1-Cre (B6.Cg-Tg(Mx1-cre)1Cgn/J) (JAX stock no. 003556) and BMDCs (B6.129N(Rosa26Sor)No1J) (stock no. 008463) were obtained from the Jackson Laboratory. B6.SJL (B6. SJL–Ptprc<sup>Pepe</sup>/BoyJ) mice (strain code 56410) were obtained from Charles River. Zeb2–EGFP fusion protein reporter (STOCK Zeb2flbabe<sup>−/−</sup>) 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. Sdp<sup>pm</sup>−/− (Zeb2<sup>−/−</sup>) mice were from Y. Higashiguchi<sup>41</sup>. For experiments shown in Fig. 6f,g, Id2–CreER<sup>E2</sup> mice (JAX stock no. 016224) were bred to Zeb2<sup>−/−</sup> mice to generate Id2<sup>−/−</sup>Zeb2<sup>−/−</sup> mice. These mice were crossed to generate Id2<sup>−/−</sup>Zeb2<sup>−/−</sup>Id2<sup>−</sup>CreER<sup>E2</sup> or Id2<sup>−</sup>CreER<sup>E2</sup>Zeb2<sup>−/−</sup> mice. Livers from 1-day-old Id2<sup>−/−</sup>Zeb2<sup>−/−</sup> mice were dispersed and cells injected into 4- to 6-week-old, lethally irradiated SJL WT mice (Charles River) and chimeras were kept at 4 °C while being stained with antibody cocktails consisting of various antibodies, followed by depletion with MagniSort Streptavidin Negative Selection Beads (Thermo Fisher). All BM cells were then stained with fluorescent antibodies before sorting. MDPs were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup> BM cells; CD11c<sup>−</sup>CD115<sup>−</sup>CD11<sup>−</sup>CD<sup>+</sup>-cell; pre-cDC1s were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell; pre-cDC2s were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell; splenic sorting experiments, spleen was isolated and depleted of Ly6G+ B220- and CD3-expressing cells. The CDc2s were identified as Lin<sup>−</sup>CD145<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell. For splenic sorting experiments, cells were isolated from the spleens. Single-cell suspensions were prepared by manually perforating tissue with 200-μm cell strainers. Sorting was performed on a FACSAria Fusion flow cytometer, using the Seurat R toolkit<sup>51</sup>. The scRNA-seq library was sequenced in one lane of HiSeq4000 (Illumina). Sequencing was filtered and processed using scRNA-seq software (DNASTAR). Unsupervised hierarchical clustering of differentially expressed genes was computed using ArrayStar (DNASTAR) with the Euclidean distance metric and centroid linkage method.

Isolation and culture of BM progenitor cells and splenic DCs. BM progenitors and DCs were isolated as described<sup>49</sup>. For sorting experiments, BM was isolated and depleted of CD3<sup>−</sup>, CD19<sup>−</sup>, CD10<sup>−</sup> and, in some instances, Ly6G<sup>−</sup> and CD45R-expressing cells by staining with the corresponding biotinylated antibodies, followed by depletion with MagniSort Streptavidin Negative Selection Beads (Thermo Fisher). All BM cells were then stained with fluorescent antibodies before sorting. MDPs were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup> BM cells; CD11c<sup>−</sup>CD115<sup>−</sup>CD11<sup>−</sup>CD<sup>+</sup>-cell; pre-cDC1s were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell; pre-cDC2s were identified as Lin<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell; splenic sorting experiments, spleen was isolated and depleted of Ly6G<sup>−</sup> B220- and CD3-expressing cells. The CDc2s were identified as Lin<sup>−</sup>CD145<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup>CD<sup>+</sup>-cell. For splenic sorting experiments, cells were isolated from the spleens. Single-cell suspensions were prepared by manually perforating tissue with 200-μm cell strainers. Sorting was performed on a FACSAria Fusion flow cytometer, using the Seurat R toolkit<sup>51</sup>. The scRNA-seq library was sequenced in one lane of HiSeq4000 (Illumina). Sequencing was filtered and processed using scRNA-seq software (DNASTAR). Unsupervised hierarchical clustering of differentially expressed genes was computed using ArrayStar (DNASTAR) with the Euclidean distance metric and centroid linkage method.

Expression microarray analysis. RNA was extracted using an RNAqueous-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 were 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 multilinear 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<sup>−</sup>CD12<sup>−</sup>CD11<sup>−</sup>CD115<sup>−</sup>CD11c<sup>−</sup> MHC-II<sup>−</sup> CD<sup>+</sup>-cell. Lineage included CD3<sup>+</sup>, CD105, CD19, Ly6G and Ter119. Single-cell genes were measured with the Chromium system using Chromium Single Cell 3 Library and Gel Bead kit v2 (10X Genomics). Cell density and viability of sorted cells were determined by Vi-CELL XR cell detector (Beckman Coulter), and all processed samples had cell viability >90%. The cell density was used to compute 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 mix and clean-up, a total of 12 cycles of PCR amplification were performed to obtain compatible cDNA libraries. Libraries were barcoded and all samples 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 toolkit<sup>22</sup>.

ATAC-seq. ATAC-seq of DC progenitors was performed using the Omni-ATAC protocol as previously described with minor modifications<sup>50</sup>. MDPs, CDPS and pre-cDC1s (10,000 in total) were sorted from BM as described above and lysed in ice-cold ATAC-LSR 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-LSR 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 2x TD buffer, 22.5 μl distilled H2O, 2.5 μl of 10 μM 5′ss 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 5′ss 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 20 μl elution buffer and stored at -20 °C until amplification. Two 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 Ifn8 enhancer. The 454-bp region of the +41-kb Ifn8 enhancer was cloned into hCD4 plasmid pRV2. Each E-box motif (CANNTG) in the enhancer was mutated to a binding-site-free DNA sequence (AACCTAC) determined by SiteOut. The primer sequences for the entire enhancer and the associated mutations are as follows:

For +41 kb Ifn8 enhancer: aaaaagctTGATCTGGGTGGAAC and GAAGAAGATCAGAGGTATG; for segment A: aaaaagctTGATCTGGGTGGAAC and aaaaaaaacctTGTGCTAATTAAAGCAGCAAGG; for segment B: aaaaaagctCTGCCAGACCCTCCCATC and aaaaaagctGGAGAACCCACACTCAAGGCTTCCAGCCCGCAGCT; for mE3/4: AGCTGCGGGCTGGGaactacCCGCACCCCCTCCCCGGAGGGTGCGGg; for mE6: ctacTCTTTGGGaactacGGATGCG; for mE5: GGCTGGAAGCCTTGAGTGGTGGTTCCCCGGGaactacTCTTCACCGTGCGGTCAGG and CGCACGGTGAAGAgtaggtttatcgatagcaagCTTGACACTCTGGGAATAG; for segment A: aaaagatctGATCTGGGGTATGTGGG and aatcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagcttttagc.

Lin–CD117high 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 of polybrene by ‘spin infection’ at 2,250 g for 60 min. Viral supernatant was replaced and the culture was read out 5% Flt3L 1 d after transduction and the culture was read out on day 3. For analysis, the enhancer activity was quantitated using integrated mean fluorescence intensity (iMFI) of variance (ANOVA) when comparing multiple groups. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

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 Ifn8 transcription start site was found with the FIMO 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 pDCs. 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.

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. NfI3/E4bp4 is required for the development and maturation of NK cells in vivo. J. Exp. Med. 206, 2977–2986 (2009).

98. 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).

52. Estiva, J. et al. SiteOut: an online tool to design binding site-free DNA sequences. PLoS ONE 11, e0151740 (2016).

53. Shooshtari, 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).
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.

- 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, Sesequpe, 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 request 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                             | n/a     |
| 🔴 Antibodies                   | 🔴 ChiP-seq |
| ▢ Eukaryotic cell lines         | 🔴 Flow cytometry |
| 🔴 Palaeontology                 | 🔴 MRI-based neuroimaging |
| ▢ Animals and other organisms   |         |
| ▢ Human research participants   |         |
| ¬ Clinical data                 |         |

Antibodies

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 555342). All antibodies were used at a 1:200 dilution vol/vol (v/v), unless otherwise indicated.

The following antibodies were from BD: Brilliant Ultraviolet 395–anti-CD117 (clone 28B, 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 551412, 1:100 v/v), Alexa Fluor 700–anti-Ly6C (clone AL-21, catalog number 561737), Brilliant Violet 471–anti-CD117 (clone SB/199, catalog number 562959, 1:100 v/v), biotin–anti-CD19 (clone 1A3, catalog number 553784), BV510–anti-CD45R (clone RA3-682, catalog number 563103), PE–anti-CD90.2 (clone OX-7, catalog number 554899). The following antibodies were from eBioscience: allophycocyanin–anti-CD117 (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).

The following antibodies were from BioLegend: Brilliant Violet 711–anti-CD115 (clone AF598, catalog number 135515, 1:100 v/v), PE or Brilliant Violet 421–anti-XCR1 (clone ZE2, catalog number 148204 or 148216), Alexa Fluor 700 or APC/Cy7–anti-4-480 (clone IMB8, 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 RA3-682, catalog number 103203 or 103258), biotin–anti-Ly6G (clone 148, catalog number 127609), biotin–anti-Ter119 (clone TER-119, catalog number 116204), biotin–anti-CD105 (clone M1/J18, catalog number 120040), 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 BD Biosciences: FITC–anti-CD45.1 (clone A20, catalog number 55-0453-100), biotin or APC–anti-CD3e (clone 145-2c11, catalog number 30-0031-U500 or 20-0032-U100), violet-Fluor 450–anti-MHC Class II (I-A/II-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).
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 C57BL/6J mice were obtained from The Jackson Laboratory. ZbB4G8gfp/+ mice were described25. Nill3−/− mice were from A. Joksh and Tak Mak42. Mx1-Cre [B6.Cg-Tg(Mx1-Cre)1(Cgn)J] mice [stock no. 003556], and Rosa26Cre/Cre [B6.129-Gt(ROSA26Sortm1CreERT2)1Tyb/J] mice (stock no. 028453) were obtained from The Jackson Laboratory. B6.SIL (B6.SIL-Tptcrca Pepcb/Boyj) mice [strain code 564], were obtained from Charles River. ZEB2-EGFP fusion protein reporter (STOCK Z1hxtbx1;1yhj) mice29 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. Sip1lox(ex7) Zeb2fl/f were from Y. Igashis43. For experiments shown in Fig. 6d,e, Id2-CreERT2 mice (JAX stock #016272)44 were bred to ZbB4G8gfp mice to generate Id2creERT2+/+ZbB4G8gfp/+ mice. These mice were crossed to generate Id2creERT2/creERT2 ZbB4G8gfp/+ or gfp/gfp mice. Livers from day 1 of Id2creERT2/creERT2 pups were dispersed and cells injected into 4-5 week old lethally irradiated S1L WT mice (Charles Rivers) and chimeras used eight weeks after reconstitution. Id2-fox and Id2-IRES-GFP mice30 were generously donated by G. Bele. Tcf3GFP/+ were generated by crossing the Tcf2afl allele (B6.129-Tcf3tm1Mbu/J JAX stock #028184) with Vav-Cre 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 personal 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 isoceiver’s modified Dulbecco’s media (iMDM) + 10% FCS (iMDM) with 250 μg/mL collagenase B (Roche) and 30 U/mL DNase (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-x106 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-µm strainer. Red blood cells were lysed with ACK lysis buffer and cells were subsequently counted on a Vi-CELL analyzer (Beckman Coulter). 3.10x10^6 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. FcεRI (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 µg/ml collagenase B (Roche) and 30 U/ml DNase (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 DNase (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-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 3.5x10^6 cells were used per antibody staining reaction.

For peritoneal cell analysis, 5 ml of MACS buffer (DPBS + 0.5% BSA + 2 mM 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-AVFSW profile as traditionally done. |