Published in final edited form as:

Nat Immunol. 2019 September ; 20(9): 1174–1185. doi:10.1038/s41590-019-0449-3.

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

Prachi Bagadia[1,13], **Xiao Huang**[1,13], **Tiantian Liu**[1,2,13], **Vivek Durai**[1], **Gary E. Grajales-Reyes**[1], **Maximilian Nitschke**[3], **Zora Modrusan**[4], **Jeffrey M. Granja**[5,6,7], **Ansuman T. Satpathy**[5,8], **Carlos G. Briseño**[9], **Marco Gargaro**[10], **Arifumi Iwata**[11], **Sunkyung Kim**[1], **Howard Y. Chang**[5,12], **Andrey S. Shaw**[3], **Theresa L. Murphy**[1], **Kenneth M. Murphy**[1,2,*]

[1]Department of Pathology and Immunology, Washington University in St. Louis, School of Medicine, St. Louis, Missouri

[2]Howard Hughes Medical Institute, Washington University in St. Louis, School of Medicine, St. Louis, Missouri

[3]Research Biology, Genentech, South San Francisco, California

[4]Department of Molecular Biology, Genentech, South San Francisco, California

[5]Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, CA, USA

[6]Deparment of Genetics Stanford University School of Medicine, Stanford, CA, USA

[7]Biophysics Program, Stanford University School of Medicine, Stanford, CA, USA

[8]Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA

[9]Department of Oncology, Amgen Inc., South San Francisco, California

[10]Department of Experimental Medicine, University of Perugia, Italy

[11]Department of Allergy and Clinical Immunology, Graduate School of Medicine, Chiba University, Chiba

[12]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA

Abstract

*To whom correspondence should be addressed. Phone: 314-362-2009, kmurphy@wustl.edu.

[13]Co-first author;

AUTHOR CONTRIBUTIONS

P.B., X.H., T.L., T.L.M., and K.M.M. designed the study; P.B., X.H., and 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 single-cell RNA-sequencing; 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.L., and K.M.M. wrote the manuscript with advice from all authors.

COMPETING INTERESTS

The authors declare no competing interests.

Classical type 1 dendritic cells (cDC1s) are required for anti-viral and anti-tumor 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 DC potential and explains the switch in *Irf8* enhancer usage during cDC1 development.

Introduction

Development of classical type 1 dendritic cells (cDC1s) has become a topic of interest because of the critical role this lineage plays in anti-tumor immunity and checkpoint blockade therapy[1]. DCs are an immune lineage encompassing classical DCs (cDCs) and plasmacytoid DCs (pDCs)[2,3]. cDCs comprise two branches, cDC1 and cDC2, that exert distinct functions *in vivo* and rely on different transcriptional programs[4]. pDCs and cDCs can both arise from the common DC progenitor (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 *Irf8* and *Batf3* are required for cDC1 development[9,13,14], but cDC1 develop from CDP progenitors 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[9], causing it to divert into cells that are transcriptionally similar to cDC2 (XXXX). An enhancer located at +32 kb of the IRF8 transcription start site contained several AP1-IRF composite elements (AICEs) that bind IRF8 and BATF3 in cDC1s *in vivo*[9]. CRISPR-mediated deletion of the +32 kb *Irf8* enhancer in mice (*Irf8 +32*$^{-/-}$) suggests that *Batf3* supports *Irf8* autoactivation using this enhancer (XXXX). Like *Batf3*$^{-/-}$ mice, *Irf8 +32*$^{-/-}$ mice lack mature cDC1 but maintain pre-cDC1 development *in vivo*. Development of this progenitor instead depends upon 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 *Nfil3, Id2*, and *Zeb2*[15–19]. *Nfil3*, a basic leucine zipper (bZIP) transcriptional repressor[20], is expressed in cDC1s and is required for cDC1 development[15,21], but how it functions is unknown[4,15].

Id2 is a known inhibitor of E proteins, is expressed in both cDC1 and cDC2, and is required only for cDC1 development[16,17]. *Id2* may exclude pDC fate by blocking activity of E proteins, particularly E2-2 (*Tcf4*), required for pDCs[22–24]. However, this model predicts that *Id2*$^{-/-}$ mice should lack both cDC1 and cDC2 lineages, since 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[18,19]. How these factors precisely interact and at what stage they influence cDC1 specification is unknown.

Here, we used single-cell RNA-sequencing and genetic epistasis to determine the functional hierarchy of transcription factors involved in cDC1 specification. We organized a transcriptional circuit that explains the switch in *Irf8* 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. Single-cell RNA-sequencing 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, and thereby imposing a new requirement for *Batf3* for maintaining *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 MHC-II and CD11c expression[6]. Subsequently, pre-cDC1 and pre-cDC2 progenitors were identified to arise from the CDP but were not contained within the CDP[8,9]. Pre-cDC1s were defined as Lin^{-}CD117^{int}CD135^{+}CD11c^{+}MHC-II$^{lo\text{-}int}$ 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)[9,25]. In each case, we 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 nearly exclusive cDC1 potential *in vitro*, comparable to pre-cDC1, 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, Supplementary Fig. 1a).

The transcriptional profile of these *Zbtb46*-GFP^{pos} CDPs suggests they represent an intermediate population between a non-specified CDP, the *Zbtb46*-GFP^{neg} CDP, and the pre-cDC1 (Fig. 1d,e). For example, we considered genes whose expression changed more than 8-fold between the *Zbtb46*-GFP^{neg} CDP and the pre-cDC1. 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 Supplementary Tables 1,2). *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-cDC2 (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.

Single-cell RNA-sequencing of the CDP identifies factors associated with cDC1 specification

The identification of *Zbtb46*-GFP expressing cells in the CDP that had nearly exclusive cDC1 potential suggested that the CDP might contain cells that have already specified to cDC1 fate. Single-cell RNA-sequencing (scRNA-seq) was performed on 9,554 CDPs defined as Lin^{-}CD127^{-}CD117^{int}CD115^{+}CD135^{+}MHC-II^{-}CD11c^{-} (Fig. 2a) on the 10X Genomics platform to assay for unrecognized heterogeneity within this population. Uniform Manifold Approximation and Projection (UMAP) analysis[26–28] identified 8 closely connected clusters (Fig. 2b,c). Although we were able to identify genes that were specifically enriched in certain clusters, others such as *Klf4* and *Ly6d* were not specifically enriched in one cluster (Supplementary Fig. 1b). However, scRNA-seq was able to identify a cluster that was enriched in *Zbtb46* expression, corroborating our 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* (E2-2) and *Zeb2* (Fig. 2d,e). Cluster 3 also showed reduced *Csf1r* expression (Fig. 2d), consistent with lower CD115 expression in pre-cDC1 and incongruent with the higher CD115 expression 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, likely contained macrophage or neutrophil contamination as 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*[29,30,31,32] were not differentially expressed across the CDP, perhaps suggesting that specification of cDC2s and pDCs occurs after the CDP (Fig. 2d, Supplementary Fig. 1b). In addition, the CDP appeared homogenous 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 these genes may regulate cDC1 specification at an earlier stage than previously recognized.

cDC1 specification is functionally characterized by low *Zeb2* and high *Id2* expression

To test the functional importance of these genes for cDC1 specification, we first analyzed two reporter mouse lines expressing a ZEB2-EGFP fusion protein (*Zeb2*^{egfp})[33] or an Id2-IRES-GFP cassette (*Id2*^{gfp})[34]. Both reporters exhibit a GFP expression pattern consistent with the level of *Zeb2* and *Id2* gene expression across many immune lineages (Supplementary Fig. 2a,b). In *Zeb2*^{egfp} mice, 90% of CDPs expressed high levels of ZEB2-EGFP, but 10% expressed low levels of ZEB2-EGFP, similar to 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*^{egfp} 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.

We next analyzed the developmental potential of CDPs expressing high or low levels of ZEB2-EGFP, *Id2*-GFP, and *Zbtb46*-GFP 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 *in vitro* Flt3L culture (Fig. 3d,e, Supplementary Fig. 2c,d). Finally, CDPs expressing *Zbtb46*-GFP developed nearly exclusively into cDC1 (96%), while CDPs lacking *Zbtb46*-GFP developed into both cDC1 (30%) and cDC2 (70%) (Fig. 1c, 3e). In all three cases, pDCs developed exclusively from CDPs that were either *Zbtb46*-GFP^{neg}, ZEB2-EGFP^{hi}, or *Id2*-GFP^{lo} (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 an intermediate population between non-specified CDPs and the pre-cDC1 (Fig. 3f–i). We considered genes whose expression differed more than 5-fold between the pre-cDC1 and either ZEB2-EGFP^{hi} CDPs (Fig. 3f,g) or *Id2*-GFP^{lo} CDPs (Fig. 3h,i). 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 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 pre-cDC1s (Fig. 3h,i and Supplementary Table 4). Additionally, the cells that are ZEB2-EGFP^{lo} within the CDP have induced *Id2*, and cells that are *Id2*-GFP^{hi} within the CDP have downregulated *Zeb2* (Fig. 3f–i). Both of these populations also show increasing *Zbtb46* expression compared to the non-specified CDPs. Although these three cDC1-specificed 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 to the previously identified pre-cDC1.

Nfil3 is required for cDC1 specification within the CDP

Nfil3 is required for cDC1 development[15], but its mechanism and timing of action remain obscure. To determine the stage where *Nfil3* acts in cDC1 development, we crossed *Nfil3*$^{-/-}$ mice with ZEB2-EGFP, *Id2*-GFP and *Zbtb46*-GFP reporter mice, and assayed whether cDC1-specified progenitors developed in 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, but do develop normally in *Batf3*$^{-/-}$*Zbtb46*$^{gfp/+}$ mice as previously described (Supplementary 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*^{egfp} reporter mice, cDC1-specified cells are identified as CD117^{int} ZEB2-EGFP ^{lo} cells that includes pre-cDC1s and ZEB2-EGFP^{lo} CDPs and comprise

approximately 6% of Lin^{-}CD135^{+} 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 that 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 that include pre-cDC1s and *Id2*-GFP^{hi} CDPs and comprises approximately 2% of Lin^{-}CD135^{+} BM (Fig. 4e,f). However, these cells are absent in *Nfil3*$^{-/-}$*Id2*^{gfp} mice. Further, cDC1-specified CDPs can be identified as *Id2*-GFP^{hi} cells that comprise 7% of the CDP (Fig. 4e,f), but which 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

We next evaluated the interactions between *Nfil3* and other factors using genetic mutants rather than GFP reporters. We first examined interactions between *Nfil3* and *Zeb2*. We crossed *Nfil3*$^{-/-}$ mice to *Zeb2*$^{f/f}$*Mx1*-Cre mice in which ZEB2 can be inactivated by poly(I:C) treatment (*Zeb2*$^{-/-}$). We compared cDC1 development and the presence of cDC1-specified progenitors in *Nfil3*$^{+/+}$*Zeb2*$^{f/f}$*Mx1-cre*$^{-}$ (wildtype), *Nfil3*$^{-/-}$, *Zeb2*$^{-/-}$, mice as well as *Nfil3*$^{-/-}$*Zeb2*$^{-/-}$ mice (Fig. 5). First, *Zeb2*$^{-/-}$ mice have more than a 2-fold increase in splenic cDC1s compared with wildtype mice (Fig. 5a,b), consistent with our previous study[19]. Further, *Nfil3*$^{-/-}$ mice lacked cDC1s in spleen, as previously reported[15]. However, *Nfil3* $^{-/-}$ *Zeb2*$^{-/-}$ DKO mice had a splenic cDC1 population that, like *Zeb2*$^{-/-}$ mice, is about 2-fold greater than 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*$^{-/-}$ DKO BM was increased compared to *Nfil3*$^{-/-}$ BM. Finally, we directly examined pre-cDC1 development in these mice. *Zeb2*$^{-/-}$ mice have increased numbers of pre-cDC1 compared to wildtype mice, while *Nfil3*$^{-/-}$ mice have greatly reduced numbers of pre-cDC1 (Fig. 5e,f). However, *Nfil3*$^{-/-}$*Zeb2*$^{-/-}$ DKO mice have markedly restored pre-cDC1 development compared to *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[18,19], but no mechanism has been established. To evaluate the genetic interaction between *Zeb2* and *Id2*, we crossed the *Rosa26*$^{Cre-ERT2}$ strain with *Zeb2*$^{f/f}$, *Id2*$^{f/f}$, and *Zeb2*$^{f/f}$ *Id2*$^{f/f}$ mice to produce mice in which tamoxifen administration can conditionally inactivate ZEB2 (*Zeb2*$^{-/-}$), ID2 (*Id2*$^{-/-}$), or both (*Zeb2*$^{-/-}$*Id2*$^{-/-}$) , respectively. We first evaluated pre-cDC1 specification and cDC1 development in these mice (Fig. 6a–d). *Zeb2*$^{-/-}$ mice show a 2-fold increase in cDC1 and pre-cDC1 compared with wildtype mice, similar to mice with ZEB2 deficiency generated using poly(I:C) and *Mx1*-Cre (Fig. 5). *Id2*$^{-/-}$ mice lack splenic cDC1, as expected[17], and also lack pre-cDC1 in BM. However, *Zeb2*$^{-/-}$*Id2*$^{-/-}$ mice showed a restored development of splenic cDC1 and BM pre-cDC1 (Fig. 6a–d). Moreover, similar results were obtained from *in vitro* Flt3L cultures of BM cells from these

mice (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

We next compared the transcriptional profiles of splenic cDC1 in wildtype $Zeb2^{-/-}$, $Zeb2^{-/-}Id2^{-/-}$, and $Nfil3^{-/-}Zeb2^{-/-}$ mice using gene expression microarrays (Fig. 6e, Supplementary Fig. 4c, Supplementary Table 5). cDC1 from all genotypes expressed high *Irf8* and *Batf3*, and low *Irf4* and *Tcf4*, levels, as expected. *Nfil3* was highly expressed in cDC1 isolated from wildtype, $Zeb2^{-/-}$, and $Zeb2^{-/-}Id2^{-/-}$ mice and was absent in cDC1 isolated from $Nfil3^{-/-}Zeb2^{-/-}$ mice, consistent with *Nfil3* genetically functioning upstream of both *Zeb2* and *Id2*. Further, *Id2* was expressed at the expected high levels in cDC1 from wildtype and $Zeb2^{-/-}$ mice, and absent in cDC1 from $Zeb2^{-/-}Id2^{-/-}$mice, in agreement with *Id2* genetically functioning upstream of *Zeb2*. Unexpectedly, *Id2* gene expression remained high in cDC1 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 *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, as ZEB2 deficiency can restore cDC1 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-cDC1 would maintain *Zeb2* expression, unlike $Id2^{+/+}$ pre-cDC1. To test this, we used chimeric mice reconstituted with $Id2^{-/-}Zbtb46^{gfp/gfp}$ BM ($Id2^{-/-}Zbtb46^{gfp/gfp}$). We first confirmed that splenic pDCs and cDC2s develop normally in $Id2^{-/-}Zbtb46^{gfp/gfp}$ chimeras (Supplementary Fig. 5a). We also showed that $Id2^{-/-}Zbtb46^{gfp/gfp}$ cDC2s are transcriptionally essentially identical to $Id2^{+/+}Zbtb46^{gfp/+}$cDC2s (Supplementary Fig. 5b and Supplementary Table 6). Further, unspecified CDPs, defined as $Lin^{-}CD117^{int}$*Zbtb46*-GFP^{-} CDPs, in $Id2^{-/-}$ mice are similar to $Id2^{+/+}$and $Batf3^{-/-}$ CDPs, both in frequency, expression of CD115 and CD135 (Fig. 6f), and transcriptional profile (Fig. 6g, Supplementary Fig. 5c, and Supplementary Table 7). However, in $Id2^{-/-}Zbtb46^{gfp/gfp}$ chimeras, cDC1-specified cells ($Lin^{-}CD117^{int}$*Zbtb46*-GFP^{pos}) were present but were reduced in frequency by 3-fold. The cDC1-specified cells in $Id2^{-/-}Zbtb46^{gfp/gfp}$ chimeras maintained CD135 expression but had higher expression of CD115 compared to $Id2^{+/+}Zbtb46^{gfp/+}$, implying a partial block in development of specified cDC1s. In addition, these cells failed to induce *Batf3* but maintained expression of *Zeb2* compared to $Id2^{+/+}$ *Zbtb46*-GFP^{pos} cells (Fig. 6g and Supplementary Fig. 5c). These results confirm a role for *Id2* in inducing *Batf3* and repressing *Zeb2* expression during cDC1 specification. Since *Id2* inhibits E protein transcription factors, *Id2* might indirectly repress *Zeb2* if E proteins supported *Zeb2* expression. In agreement, E2A is expressed in CDPs and

binds to E-box motifs in the *Zeb2* locus based on ChIP-seq analysis (Supplementary Fig. 5d,e,f)[35].

Id2 induction imposes a switch in *Irf8* enhancer usage during cDC1 development

Data has revealed that E proteins may be necessary for the sufficient induction of *Irf8* in the CDP by activating the +41 kb *Irf8* enhancer (XXXX). This enhancer is transiently active during cDC1 progenitor development, but is required for the development of both pre-cDC1 and cDC1 *in vivo*[9,36](XXXX). This 454 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 (Supplementary Fig. 6)[37]. Using the 454 bp region in a retroviral reporter system[9], we found robust activity that was specific for pDCs, but not cDC1s or cDC2 (Fig. 7b,c). We also examined the activity of three individual enhancer segments each containing 2 E-box motifs. Segments A and C showed reduced overall activity compared with the 454 bp enhancer, but retained pDC specificity, while 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 significantly reduced enhancer activity in pDCs (Supplementary Fig. 7a,b). Within segment A, mutation of either E-box alone reduced overall activity, while mutation of both E-boxes together completely extinguished activity (Fig. 7d, 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, 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-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 the pre-cDC1 to maintain *Irf8* expression necessary for cDC1 development. To identify a potential enhancer, we performed ATAC-seq on MDP, CDP, and pre-cDC1 progenitors and found a peak that indicated accessibility within the *Irf8* region only in the pre-cDC1 and in mature cDC1, but not in the earlier MDP or CDP or mature cDC2 (Fig. 7g, red dashed line). This peak was located at +32 kb of the *Irf8* TSS and was shown to be BATF3-dependent[9] (XXXX). 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

This study resolves several long-standing puzzles regarding cDC1 development. First, *Id2* was proposed to be required for cDC development by excluding pDC fate potential[22,23], but *Id2*$^{-/-}$ mice lacked only cDC1, and did not show the expected loss of all cDCs[16]. Second, cDC1 develop from CDP progenitors 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 cDC1 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 this study,

we examined this cryptic stage of development to reveal the hierarchy of transcription factors governing cDC1 specification.

Our results define a genetic hierarchy that unifies the actions of the known transcription factors required for cDC1 development. cDC1s were known to require *Irf8*, *Batf3*, *Id2*, and *Nfil3*, but how these factors interacted was unknown. We used *Zbtb46*-GFP to identify an earlier stage of cDC1 specification than previously described that occurs within the CDP itself[9]. Single-cell RNA-sequencing 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 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*, as there may be other factors in this proposed genetic circuit. *Nfil3* acts largely as a repressor [20,38],but may activate transcription in contexts[39]. Likewise, *Zeb2* has been suggested to directly repress *Id2* expression[18,19], although this has not been rigorously tested. *Nfil3*, *Zeb2*, and *Id2* have also been shown to regulate ILC development[40], but the mechanisms by which these transcription factors act in these cells has not been studied.

Although our study seems to clarify several outstanding questions in cDC1 development, it may raise the possible necessity of a revised DC development scheme. We identified a cDC1-specified stage that occurs before the development of the pre-cDC1. 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 hematopoietic stem cells specifies to the DC1 lineage[41]. cDC1 specification may occur even earlier than our 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 during the transition from the MDP to the CDP for subsequent cDC1 specification is consistent with this idea of a minimum threshold for *Irf8* expression. A revised DC development model may require a deeper understanding of the relationship between IRF8 expression level and activity.

Our 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, as early studies suggested that pDCs can arise from both lymphoid and myeloid BM progenitors[42]. However, two recent studies indicated that late pDC progenitors emerge from the common lymphoid progenitor and a “pre-pDC” was described[43,44]. Since these studies did not perform lineage tracing for prior expression of myeloid markers, such as CD115-Cre, pDCs progenitors conceivably could emerge in a

series of stages that include both myeloid and lymphoid features, as recently suggested[45]. Resolving whether pDC and cDC1 share a common progenitor that has segregated from the cDC2 lineage, or simply share molecular transcriptional requirements will require additional studies.

Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Kenneth Murphy (kmurphy@wustl.edu)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice—WT C57BL6/J mice were obtained from The Jackson laboratory. *Zbtb46*$^{gfp/+}$ mice were described[25]. *Nfil3*$^{-/-}$ mice were from A. Look and Tak Mak[46]. *Mx1*-Cre [B6.Cg-Tg(Mx1-cre)1Cgn/J] mice (stock no. 003556), and *Rosa26*$^{Cre/Cre}$ [B6.129-*Gt(ROSA)26Sor*$^{tm1(cre/ERT2)Tyj/J}$] mice (stock no. 008463) were obtained from The Jackson Laboratory. B6.SJL (B6.SJL-*Ptprc*^{a} *Pepc*^{b} /BoyJ) mice (strain code 564), were obtained from Charles River. ZEB2-EGFP fusion protein reporter (STOCK *Zfhxlb*$^{tm2.1Yhi}$) mice[33] 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*$^{flox(ex7)}$ (*Zeb2*$^{f/f}$) were from Y. Higashi[47]. For experiments shown in Fig. 6f,g, *Id2-CreER*^{T2} mice (JAX stock #016222)[48] were bred to *Zbtb46*^{gfp} mice to generate *Id2*$^{creERT2/+}$*Zbtb46*$^{gfp/+}$ mice. These mice were crossed to generate *Id2*$^{creERT2/creERT2}$ *Zbtb46*$^{gfp/+ or gfp/gfp}$ mice. Livers from day 1 old *Id2*$^{creERT2/creERT2}$ 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[34] were generously donated by G. Belz. *Tcf3*$^{GFP/+}$ were generated by crossing the Tcfe2a^{fl} allele (B6.129-Tcf3tm1Mbu/J JAX stock #028184) with Vav-iCre mice (JAX stock #008610).

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.

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.

Antibodies and flow cytometry.—Cells were kept at 4ºC 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.

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).

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). Cells were analyzed on a FACSCanto II or FACSAria Fusion flow cytometer (BD), and data were analyzed with FlowJo v10 software (TreeStar).

Induced Gene Deletion.—Conditional gene deletion in *Nfil3*$^{-/-}$*Zeb2*$^{f/f}$*Mx1-cre* (*Nfil3*$^{-/-}$ *Zeb2*$^{-/-}$), *Zeb2*$^{f/f}$ *Mx1-cre*$^{-}$ *Nfil3*$^{+/+}$ (WT), *Zeb2*$^{f/f}$ *Mx1-cre*$^{-}$ *Nfil3*$^{-/-}$ (*Nfil3*$^{-/-}$) and *Zeb2*$^{f/f}$ *Mx1-cre*$^{+}$ *Nfil3*$^{+/+}$ (*Zeb2*$^{-/-}$) mice was induced by *i.p.* injection of 150 μg poly(I:C) (SigmaAldrich; 1.0 mg/mL stock solution dissolved in saline) twice within 36–72 h. Gene deletion in WT, *Zeb2*$^{f/f}$ *Rosa26*$^{Cre-ERT2}$ (*Zeb2*$^{-/-}$), *Id2*$^{f/f}$ *Rosa26*$^{Cre-ERT2}$ (*Id2*$^{-/-}$) and *Zeb2*$^{f/f}$ *Id2*$^{f/f}$ *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 significant weight loss was observed. After treatment, mice were rested on regular chow for one week before analysis.

Isolation and culture of BM progenitor cells and splenic DCs.—Bone marrow progenitors and DCs were isolated as described[9]. 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 prior to sorting. MDPs were identified as $Lin^{-}CD117^{hi}CD135^{+}CD115^{+}$ BM cells; CDPs were $Lin^{-}CD117^{int}CD135^{+}CD115^{+}MHC\text{-}II^{-}CD11c^{+}$; pre-cDC1s are $Lin^{-}CD117^{int}CD135^{+}CD115^{-}MHC\text{-}II^{lo\text{-}int}CD11c^{+}CD24^{+}Siglec\text{-}H^{-}$ or as $Lin^{-}CD117^{int}CD135^{+}MHC\text{-}II^{lo\text{-}int}CD11c^{+}Siglec\text{-}H^{-}Zbtb46\text{-}GFP^{+}$, and pre-cDC2s as $Lin^{-}CD117^{lo}CD135^{+}CD115^{+}MHC\text{-}II^{-}CD11c^{+}$. For splenic sorting experiments, spleen was isolated and depleted of Ly6G-, B220-, and CD3-expressing cells. cDC2 were identified as $Lin^{-}CD45R^{-}CD317^{-}MHC\text{-}II^{+}CD11c^{+}CD172a^{+}$ cells. Cells were purified on a FACSAria Fusion into IMDM plus 10% FBS with 5% Flt3L conditioned media. Sort purity of >95% was confirmed by post-sort analysis before cells were used for further experiments. For experiments that included Flt3L cultures, sorted cells (1×10^3 to 10×10^3 cells per 200 μl complete IMDM) were cultured for 5 or 7 d at 37 °C with 5% Flt3L conditioned media.

Expression microarray analysis.—RNA was extracted with a RNAqueous-Micro Kit (Ambion) or a NucleoSpin RNA XS Kit (Machery-Nagel), then was amplified with 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 was analyzed with 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 RMA Gene analysis settings (Affymetrix, Inc). Probe summarization (Robust Multichip Analysis, RMA), 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 with ArrayStar (DNASTAR) with the Euclidean distance metric and centroid linkage method.

Single-cell RNA-sequencing.—100,000 CDPs were sort purified as Live,[CD105, CD3, CD19, Ly6G, Ter119]$^{-}CD127^{-}CD117^{int}CD115^{+}CD135^{+}MHC\text{-}II^{-}CD11c^{-}$ cells and single-cell gene 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 counter (Beckman Coulter), and all processed samples had cell viability at >90%. The cell density was used to impute volume of single cell suspension needed in the reverse transcription (RT) master mix, to achieve ~6,000 cells per sample. After Gel Bead-in-Emulsion reverse transcription (GEM-RT) reaction and clean-up, a total

of 12 cycles of PCR amplification was performed to obtain cDNAs. 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 single-cell RNA-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 means 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 toolkit[49].

ATAC-Seq.—ATAC-Seq of DC progenitors was performed using the Omni-ATAC protocol as previously described with minor modifications[36]. 10,000 MDPs, CDPs, and pre-cDC1s were sorted from bone marrow 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 2X TD buffer, 22.5 μL dH_2O, 2.5 μL 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-RV[9]. Each E-box motif (CANNTG) in the enhancer was mutated to a binding site-free DNA sequence (AACTAC) determined by SiteOut[50].

The primer sequences for the entire enhancer and the associated mutations are as follows: *for +41 kb Irf8 enhancer*: aaaagatctGATCTGGGGTATGTGGGAAC and GAAAGAAGATCTGGGGTATGT; *for segment A:* aaaagatctGATCTGGGGTATGTGGGAAC and aaaaaagcttTGTGCTAATTAAAGCCAAGAGG; *for segment B*: aaaaggatccCTGTACCCCAGATCCCATC and aaaaaagcttGAGGAACCACCACTCAAGG; *for segment C:* aaaaggatccTCAGGTTTGGGGAAGAAG and aatcttttattttatcgatagcaagCTTGACACTCTGGGAATAG; *for segment A+B*: GCGACGGTCGCGCGAGCtagaaaagatctGATCTGGGGTATGTGGG and aatcttttattttatcgataaaaaaagcttGAGGAACCACCACT; *for segment B+C*: aaaaggatccCTGTACCCCAGATCCCATC and aatcttttattttatcgatagcaagCTTGACACTCTGGGAATAG; *for mE1*: GTGTCTCTCACaactacGGATCCCATATAAGGTTTATTTTTAC and CCTTATATGGGATCCgtagttGTGAGAGACACAAAGGGTTC; *for mE2*: GCCCAGGCCCaactacTTCCCCCCTGTACCCCAG and GTACAGGGGGGAAgtagttGGGCCTGGGCGATGTTCTG; *for mE3*:

TCCTCCTCTGGTAGAGAAGAAGCTGCGGGCTGGGaactacCCGCACCCTCCCC and GGGGAGGGTGCGGgtagttCCCAGCCCGCAGCTTCTTCTCTACCAGAGGAGG; *for mE4*: GCACCCTCCCCGGaactacTCTTCACCGTGCGGTCAGG and CGCACGGTGAAGAgtagttCCGGGGAGGGTGCGGg; *for mE5*: GGCTGGAAGCCTTGAGTGGTGGTTCCTCaactacTCTTTGGGCACCTG and CAGGTGCCCAAAGAgtagttGAGGAACCACCACTCAAGGCTTCCAGCC; *for mE6*: ctacTCTTTGGGaactacGGATGCGTCCTGTTAGGACC and CCTAACAGGACGCATCCgtagttCCCAAAGAgtagttGAGG; and *for mE¾*: AGCTGCGGGCTGGGaactacCCGCACCCTCCCCGGaactacTCTTCACCGT and ACGGTGAAGAgtagttCCGGGGAGGGTGCGGgtagttCCCAGCCCGCAGCT.

Retroviral vectors were transfected into Plat-E cells with TransIT-LTI (Mirus Bio) and viral supernatants were collected two days 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 ug/ml polybrene by ‘spin infection’ at 2,250 r.p.m. for 60 min. Viral supernatant was replaced by complete IMDM + 5% Flt3L one day after transduction and the culture was read out on day 8. For analysis, the enhancer activity was quantitated using integrated MFI[51,52].

For retroviral 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 retrovirus and concentrated by centrifugation with 2 ug/ml polybrene by ‘spin infection’ at 2,250 r.p.m. for 60 min. Viral supernatant was replaced by complete IMDM one day after transduction and the culture was read out on day 3.For analysis, the enhancer activity was quantitated using integrated MFI in cells that were co-infected with either empty or ID2 retrovirus[51,52].

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* TSS was found with FIMO[53] 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[37].

Human and mouse elements were aligned via Clustal Omega W.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis for single cell RNA-sequencing 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), using unpaired two-tailed Student’s t tests when comparing two groups or ordinary one-way or two-way ANOVA when comparing multiple groups. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

LIFE SCIENCES REPORTING SUMMARY

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

DATA AND SOFTWARE AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon request. 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: GSE132240 and is utilized in Fig. 7.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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 NCI Cancer Center Support Grant #P30 CA91842 to the Siteman Cancer Center and by ICTS/CTSA Grant# 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 NCRR or NIH. This work benefitted from data assembled by the ImmGen consortium[54]. This work was supported by the Howard Hughes Medical Institute (K.M.M. and H.Y.C.), the National Science Foundation (DGE-1745038 to P.B.), the US National Institutes of Health (F30DK108498 to V.D.; K08 CA23188-01 to A.T.S.; 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.

REFERENCES

1. Saxena M and Bhardwaj N Re-Emergence of Dendritic Cell Vaccines for Cancer Treatment. Trends Cancer 4, 119–137 (2018). [PubMed: 29458962]
2. Steinman RM and Cohn ZA Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. J Exp.Med 137, 1142–1162 (1973). [PubMed: 4573839]
3. Cella M et al., Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon [see comments]. Nature Medicine 5, 919–923 (1999).
4. Murphy TL et al., Transcriptional Control of Dendritic Cell Development. Annu.Rev Immunol 34, 93–119 (2016). [PubMed: 26735697]
5. Naik SH 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). [PubMed: 17922015]
6. Onai N et al., Identification of clonogenic common Flt3(+) M-CSFR+ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. Nature Immunology 8, 1207–1216 (2007). [PubMed: 17922016]
7. Liu K et al., In vivo analysis of dendritic cell development and homeostasis. Science 324, 392–397 (2009). [PubMed: 19286519]
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). [PubMed: 26054720]
9. Grajales-Reyes GE et al., Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha(+) conventional DC clonogenic progenitor. Nat Immunol 16, 708–717 (2015). [PubMed: 26054719]

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). [PubMed: 25687283]
11. 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). [PubMed: 27864467]
12. See P et al., Mapping the human DC lineage through the integration of high-dimensional techniques. Science 356, 1044–1044 (2017).
13. 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). [PubMed: 12461077]
14. Tamura T et al., IFN regulatory factor-4 and −8 govern dendritic cell subset development and their functional diversity. The Journal of Immunology 174, 2573–2581 (2005). [PubMed: 15728463]
15. Kashiwada M et al., NFIL3/E4BP4 is a key transcription factor for CD8{alpha}+ dendritic cell development. Blood 117, 6193–6197 (2011). [PubMed: 21474667]
16. 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). [PubMed: 12532109]
17. Hacker C et al., Transcriptional profiling identifies Id2 function in dendritic cell development. Nat Immunol 4, 380–386 (2003). [PubMed: 12598895]
18. Scott CL et al., The transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing Id2. J Exp.Med 213, 897–911 (2016). [PubMed: 27185854]
19. Wu X et al., Transcription factor Zeb2 regulates commitment to plasmacytoid dendritic cell and monocyte fate. Proc.Natl Acad.Sci.U S A 113, 14775–14780 (2016). [PubMed: 27930303]
20. Cowell IG, Skinner A, and Hurst HC Transcriptional repression by a novel member of the bZIP family of transcription factors. Mol Cell Biol 12, 3070–3077 (1992). [PubMed: 1620116]
21. Seillet C et al., CD8alpha+ DCs can be induced in the absence of transcription factors Id2, Nfil3, and Batf3. Blood 121, 1574–1583 (2013). [PubMed: 23297132]
22. Ghosh HS et al., Continuous expression of the transcription factor e2–2 maintains the cell fate of mature plasmacytoid dendritic cells. Immunity 33, 905–916 (2010). [PubMed: 21145760]
23. Watowich SS and Liu YJ Mechanisms regulating dendritic cell specification and development. Immunol.Rev 238, 76–92 (2010). [PubMed: 20969586]
24. Grajkowska LT et al., Isoform-Specific Expression and Feedback Regulation of E Protein TCF4 Control Dendritic Cell Lineage Specification. Immunity 46, 65–77 (2017). [PubMed: 27986456]
25. Satpathy AT et al., Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. Journal of Experimental Medicine 209, 1135–1152 (2012). [PubMed: 22615127]
26. McInnes L and Healy J UMAP: uniform manifold approximation and projection for dimension reduction.(2018).
27. McInnes L et al., UMAP: Uniform Manifold Approximation and Projection. The Journal of Open Source Software(2018).
28. Becht E et al., Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. (2018).
29. 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). [PubMed: 23741395]
30. 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). [PubMed: 30611612]
31. Tussiwand R et al., Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. Immunity 42, 916–928 (2015). [PubMed: 25992862]
32. Briseno CG et al., Notch2-dependent DC2s mediate splenic germinal center responses. Proc.Natl Acad.Sci.U S A 115, 10726–10731 (2018). [PubMed: 30279176]
33. Nishizaki Y et al., SIP1 expression patterns in brain investigated by generating a SIP1-EGFP reporter knock-in mouse. Genesis. 52, 56–67 (2014). [PubMed: 24243579]

34. Jackson JT et al., Id2 expression delineates differential checkpoints in the genetic program of CD8alpha+ and CD103+ dendritic cell lineages. The EMBO Journal 30, 2690–2704 (2011). [PubMed: 21587207]
35. Calero-Nieto FJ et al., Key regulators control distinct transcriptional programmes in blood progenitor and mast cells. EMBO J 33, 1212–1226 (2014). [PubMed: 24760698]
36. Corces MR et al., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14, 959–962 (2017). [PubMed: 28846090]
37. Cisse B et al., Transcription factor E2–2 is an essential and specific regulator of plasmacytoid dendritic cell development. Cell 135, 37–48 (2008). [PubMed: 18854153]
38. Cowell IG and Hurst HC Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. Nucleic Acids Res. 22, 59–65 (1994). [PubMed: 8127655]
39. 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). [PubMed: 7565758]
40. Ishizuka IE et al., The Innate Lymphoid Cell Precursor. Annu.Rev Immunol 34, 299–316 (2016). [PubMed: 27168240]
41. 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). [PubMed: 28650480]
42. Sathe P et al., Convergent differentiation: myeloid and lymphoid pathways to murine plasmacytoid dendritic cells. Blood 121, 11–19 (2013). [PubMed: 23053574]
43. Rodrigues PF et al., Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. Nat Immunol 19, 711–722 (2018). [PubMed: 29925996]
44. Herman JS, Sagar, and Grun, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. Nat Methods 15, 379–386 (2018). [PubMed: 29630061]
45. Ghosh HS et al., ETO family protein Mtg16 regulates the balance of dendritic cell subsets by repressing Id2. Journal of Experimental Medicine 211, 1623–1635 (2014). [PubMed: 24980046]
46. 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). [PubMed: 19995955]
47. 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). [PubMed: 11857784]
48. Rawlins EL et al., The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development 136, 3741–3745 (2009). [PubMed: 19855016]
49. Butler A et al., Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 36, 411–420 (2018). [PubMed: 29608179]
50. Estrada J et al., SiteOut: An Online Tool to Design Binding Site-Free DNA Sequences. PLoS One 11, e0151740 (2016). [PubMed: 26987123]
51. Shooshtari P et al., Correlation analysis of intracellular and secreted cytokines via the generalized integrated mean fluorescence intensity. Cytometry A 77, 873–880 (2010). [PubMed: 20629196]
52. Darrah PA et al., Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. Nat.Med 13, 843–850 (2007). [PubMed: 17558415]
53. Bailey TL et al., MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208 (2009). [PubMed: 19458158]
54. Heng TS, Painter MW, and Immunological Genome Project Consortium The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol 9, 1091–1094 (2008). [PubMed: 18800157]

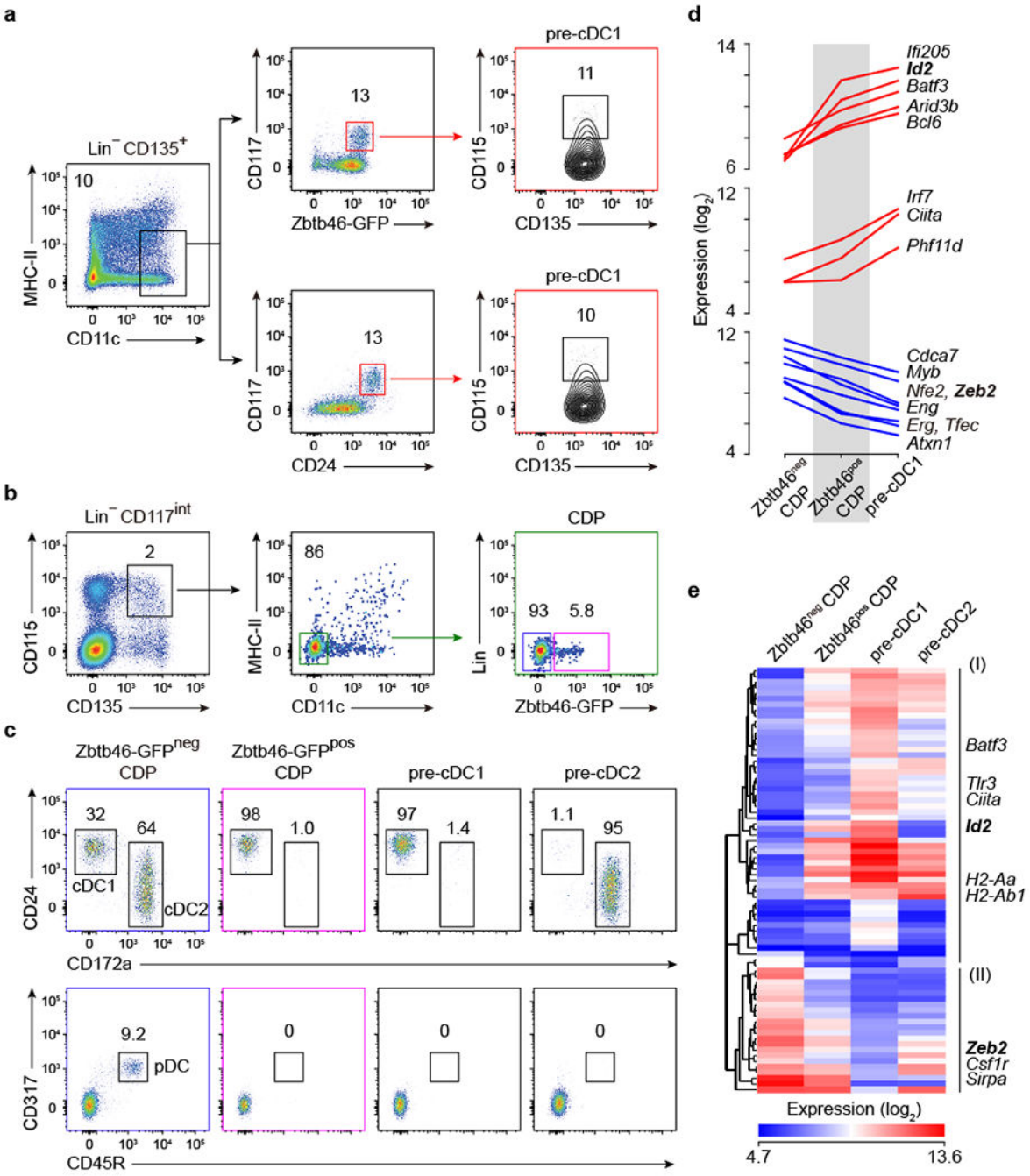

Fig. 1: *Zbtb46*-GFP Expression in CDPs Identifies the Earliest Committed cDC1 Progenitor.

a, BM from *Zbtb46*$^{gfp/+}$ mice was analyzed by flow cytometry to identify pre-cDC1 as defined by *Zbtb46*-GFP or by CD24 expression. Lineage (Lin) included CD3, CD19, NK1.1, Ly-6G, TER-119, CD105, CD127 and Siglec-H. Numbers are the percent of cells in the indicated gates (representative of 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**) (representative of three independent experiments, n = 3 mice).**c,** *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs,

pre-cDC1 and pre-cDC2 were sort purified from *Zbtb46*$^{gfp/+}$ mice, cultured for 5 d in Flt3L, and analyzed by flow cytometry for development of pDCs and cDC1 (representative of three independent experiments, n = 4 for *Zbtb46*-GFP^{pos}, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1 and n = 3 for pre-cDC2) **d-e,** *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1 and pre-cDC2 were purified as in (**c**) and analyzed using gene expression microarrays. Shown is expression of transcription factors with at least 4-fold differences between *Zbtb46*-GFP^{neg} CDP and pre-cDC1s (**d**) or hierarchical clustering for genes with a least 8-fold differences between *Zbtb46*-GFP^{neg} CDP and pre- cDC1s (**e**) (results averaged from biological triplicates for *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs, and pre-cDC1 or biological replicate for pre-cDC2).

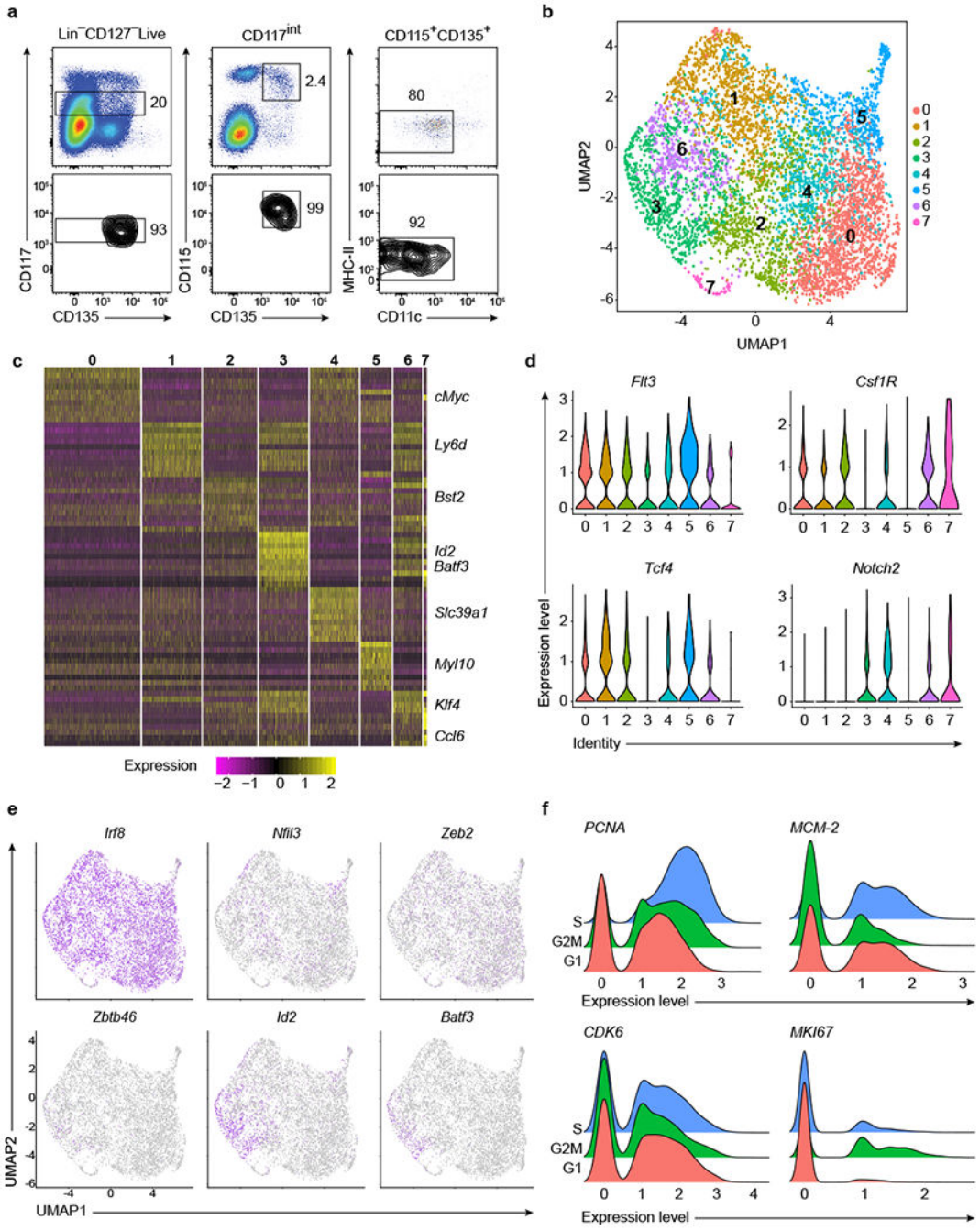

Fig. 2: Single-cell RNA Transcriptome Analysis of CDPs.

a, CDPs gated as Live,[CD105, CD3, CD19, Ly6G, Ter119]$^{-}$CD127^{-}CD117^{int}CD115^{+}CD135^{+}MHC-II^{-}CD11c^{-} cells were purified by sorting from C57BL/6J mice. Shown are pre-sort (**top**) and post-sort (**bottom**) for cells collected for single-cell RNA-sequencing. **b,** UMAP clustering of CDPs from Seurat analysis (data represents combined analysis of two independent sequencing runs)**c,** Heatmap of 9,954 cells for the top ten genes of each cluster from Seurat analysis. Shown are names of representative genes within each cluster.**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.

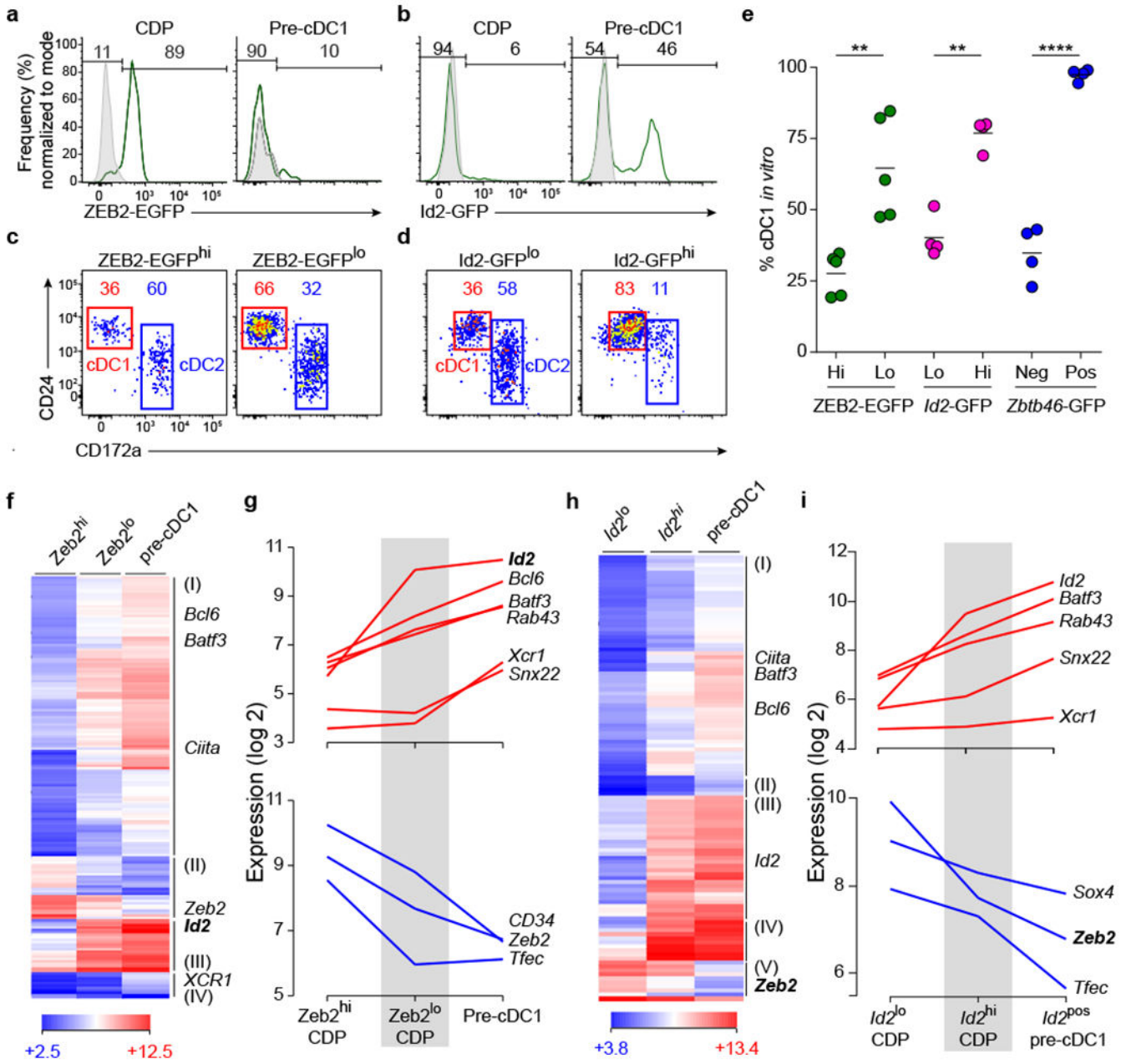

Fig. 3: *Zeb2* and *Id2* Heterogeneity Identifies cDC1 Specification in CDPs.

BM from *Zeb2*$^{egfp/egfp}$ (**a**) and *Id2*^{gfp} (**b**) mice were analyzed by flow cytometry to identify GFP expression in CDPs and pre-cDC1s. WT mice (*Zeb2*$^{+/+}$ and *Id2*$^{+/+}$) are shown as gray histograms. Numbers indicate the percentage of cells in the indicated gates. (representative of three independent experiments, n = 3 mice).**c-d,** ZEB2-EGFP^{lo} and ZEB2-EGFP^{hi} CDPs (**c**), and *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) (representative of 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 (CD45R $^{-}$CD317^{-}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} or *Id2*-GFP^{lo} CDPs and *Zbtb46*-GFP^{pos} or *Zbtb46-GFP*^{neg} CDPs). Small horizontal lines indicate the mean. **f,** Hierarchical clustering of genes expressed at least 5-fold differently between pre-cDC1 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 5-fold differently between pre-cDC1 and *Id2*-GFP^{lo} CDPs (results averaged from two independent experiments). **i,** Expression of the indicated genes described in (**h**). Data are presented as mean and two-tailed unpaired Student's t test was used to compare groups. *p < 0.05, **p < 0.01, ****p < 0.0001.

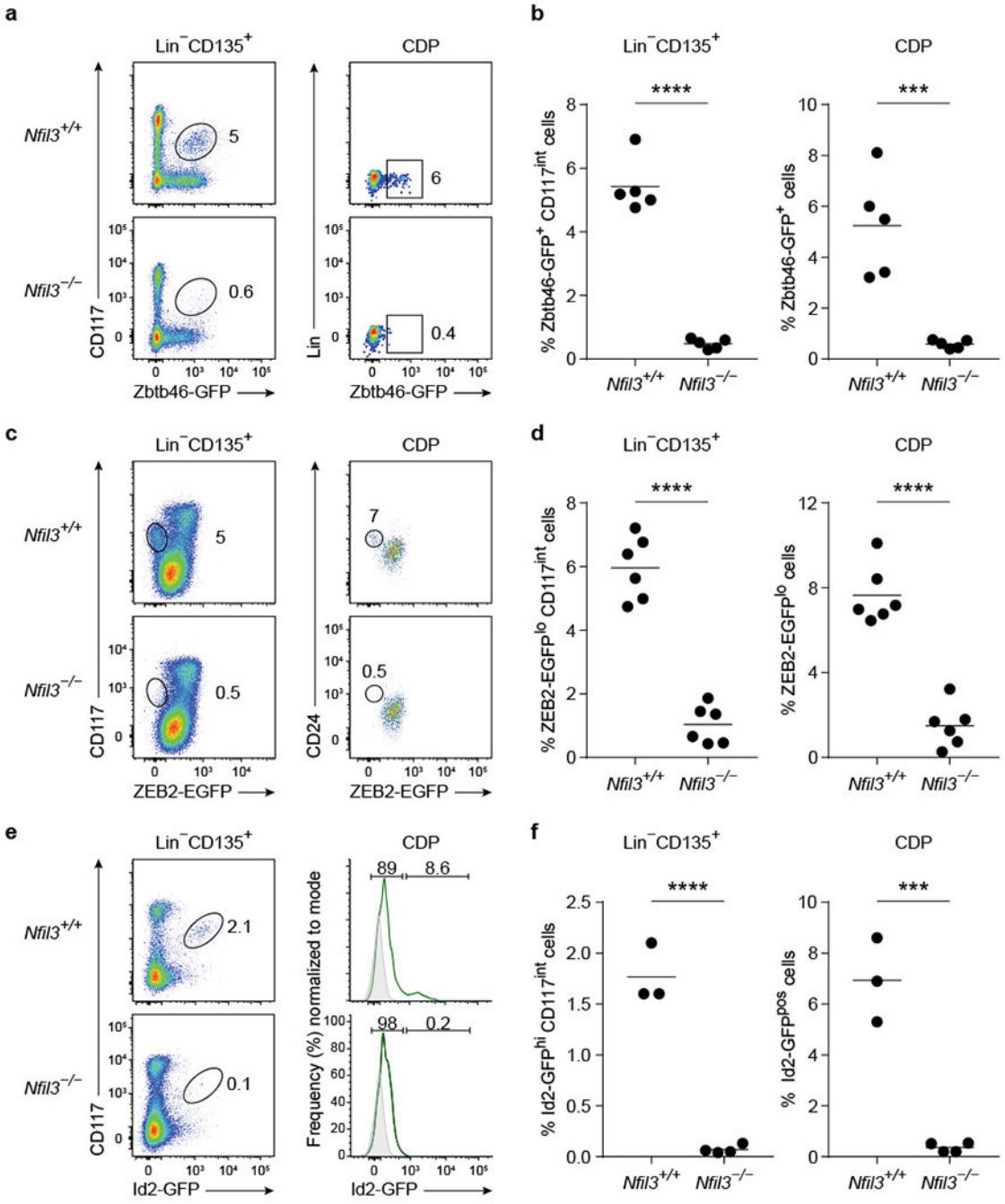

Fig. 4: Nfil3 is Required for cDC1 Specification.

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 percent of cells in the indicated gates (representative of five independent experiments, n = 5 mice). **b,** Cells from (**a**) are shown as a percentage of Lin^{-}CD135^{+} (**left**) or CDPs (**right**). Small horizontal lines indicate the mean.**c,** BM from *Nfil3*$^{+/+}$*Zeb2*$^{egfp/+}$ and *Nfil3*$^{-/-}$*Zeb2*$^{egfp/+}$ mice was analyzed for Lin^{-}CD135^{+}CD117^{int} ZEB2-EGFP^{lo} cells (**left**) or ZEB2-EGFP^{lo} CDPs (**right**) (representative of three

independent experiments, n = 6 mice). **d,** Cells from (**c**) are shown as a percentage of Lin^- $CD135^+$ (**left**) or 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**) (representative of 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^- $CD135^+$ (**left**) or CDPs (**right**). Small horizontal lines indicate the mean .Data in b, d, and f are presented as mean and two-tailed unpaired Student's t test was used to compare groups. ***p < 0.001; ****p < 0.0001.

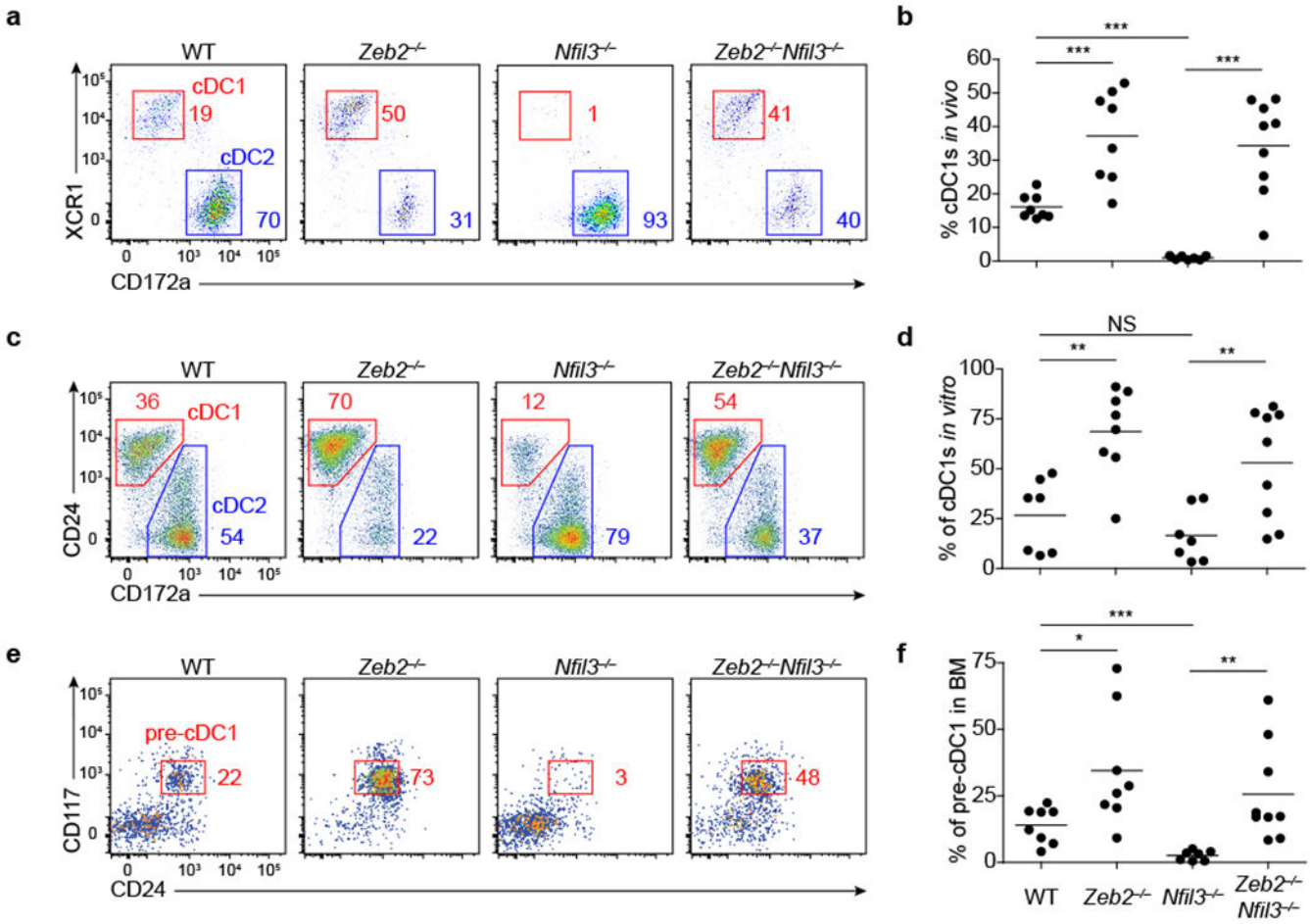

Fig. 5: *Zeb2* is Downstream of *Nfil3* in cDC1 Development.

a, Splenic cDCs from *Nfil3*$^{+/+}$*Zeb2*$^{f/f}$ *Mx1-cre*$^{-}$ (WT), *Zeb2*$^{f/f}$ *Mx1-cre*$^{+}$ (*Zeb2*$^{-/-}$), *Nfil3*$^{-/-}$ (*Nfil3*$^{-/-}$), and *Nfil3*$^{-/-}$ *Zeb2*$^{f/f}$ *Mx1-cre*$^{+}$ (*Nfil3*$^{-/-}$ *Zeb2*$^{-/-}$) mice, gated as in Fig. 3e, were analyzed for cDC1 (**red**) and cDC2 (**blue**) frequency. Numbers are the percent of cells in the indicated gates (data representative of 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)** are presented as individual mice. Small horizontal lines indicate the mean. **c,** cDCs derived *in vitro* from Flt3L-treated BM cultures from mice in (**a**) were analyzed for cDC1 (**red**) and cDC2 (**blue**) frequency as in (**a**) (data representative of 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)** are presented for individual mice. Small horizontal lines indicate the mean. **e,** BM from mice in (**a**) was analyzed for the frequency of pre-cDC1 (**red**). BM cells are pre-gated as Lin^{-} SiglecH^{-}CD135^{+} (data representative of 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)** are presented for individual mice. Small horizontal lines indicate the mean. Mean and two-tailed unpaired Student's t test was used to compare groups. *p < 0.05; ****p < 0.0001; ns, not significant.

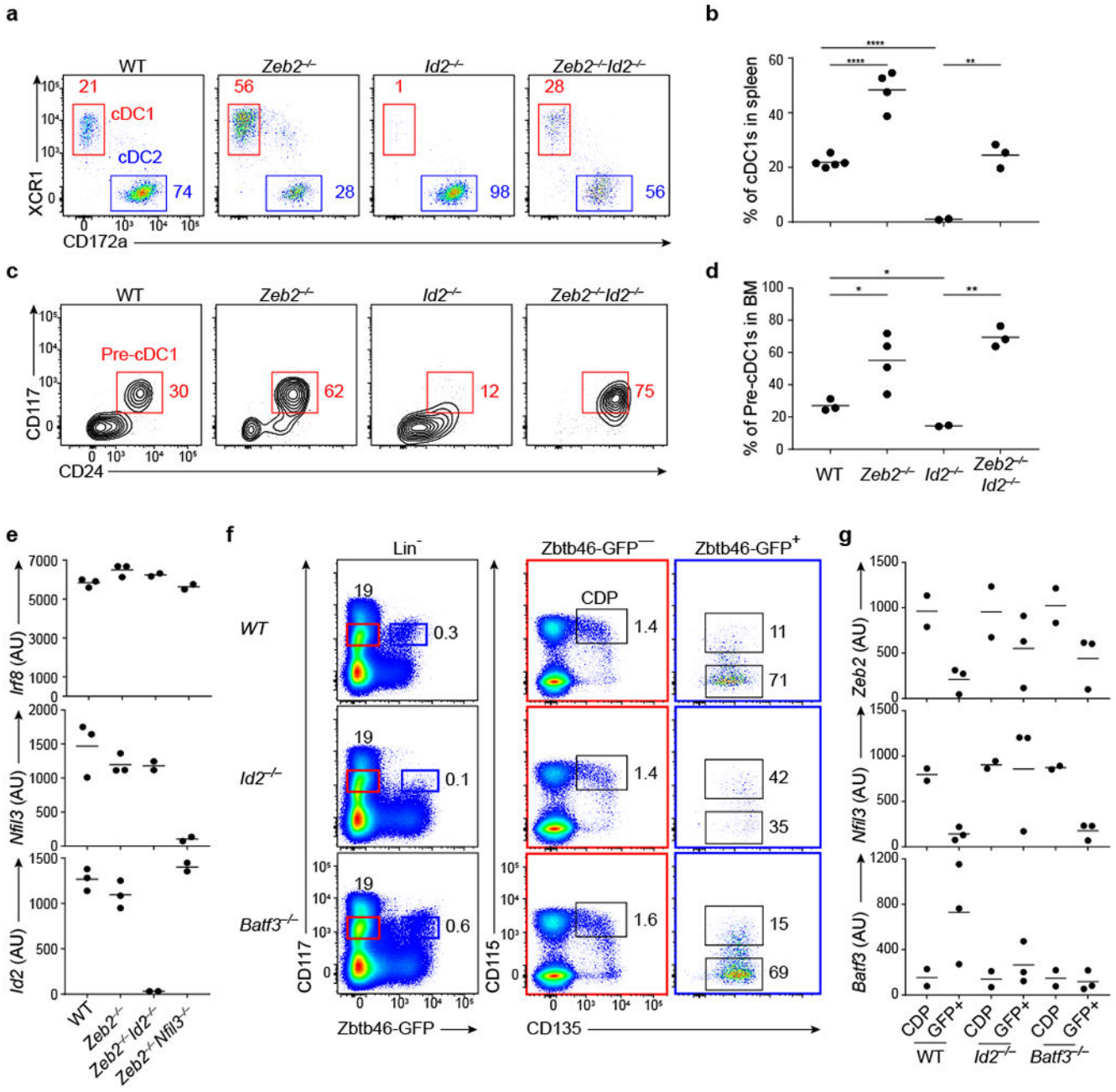

Fig. 6: Expression of *Id2* and *Zeb2* is Mutually Repressive in the CDP.

a, Splenic cDCs harvested from WT, *Zeb2*$^{f/f}$ *Rosa26*$^{(cre\text{-}ERT2/+)}$ (*Zeb2*$^{-/-}$), *Id2*$^{f/f}$ *Rosa26*$^{(cre\text{-}ERT2/+)}$ (*Id2*$^{-/-}$), and *Id2*$^{f/f}$ *Zeb2*$^{f/f}$ *Rosa26* $^{(cre\text{-}ERT2/cre\text{-}ERT2)}$ (*Zeb2*$^{-/-}$ *Id2*$^{-/-}$) were analyzed for cDC1 (**red**) and cDC2 (**blue**) frequency, gated as in Fig. 3e. Numbers are the percent of cells in the indicated gates (data representative of 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-cDC1 (**red**). BM cells are pre-gated as Lin^{-} SiglecH^{-}CD135^{+} (data representative of 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,** Shown is the expression of *Irf8, Nfil3*, and *Id2* in splenic cDC1 sorted 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^{--} (**red**) or CD117^{int}*Zbtb46*-GFP^{+} (**blue**) cells were separately analyzed for CD115 and CD135 expression (data representative of five independent experiments, n = 5 mice) **g,** CDPs and *Zbtb46*-GFP^{pos} cells in (**f**) were sort purified and analyzed by gene expression microarray. Shown are gene expression levels for *Zeb2, Nfil3, and Batf3* (data representative of 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 mean and two-tailed unpaired Student's t test was used to compare groups. *p < 0.05, ***p < 0.001, ns, not significant.

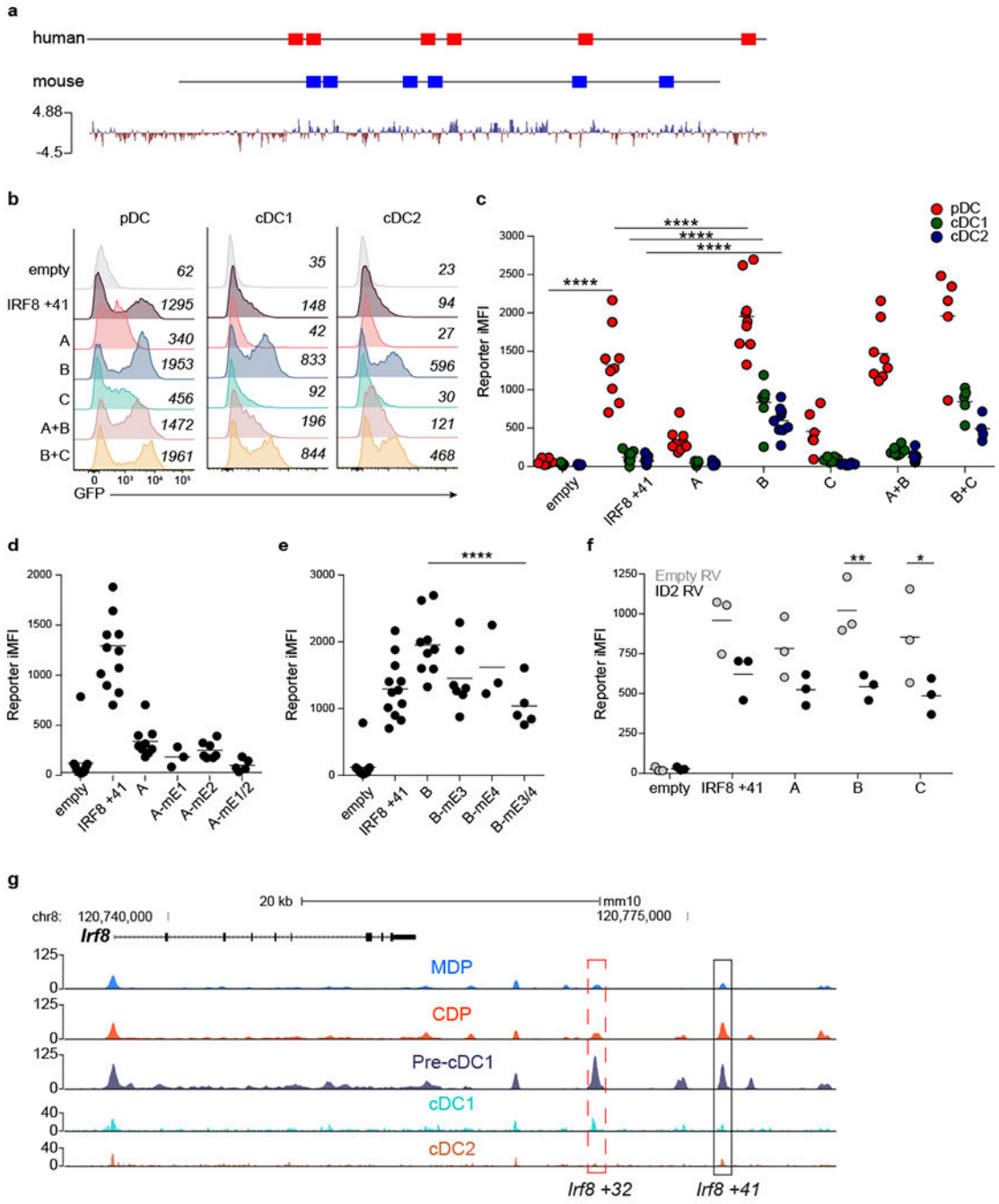

Fig. 7: *Id2* imposes a switch from the +41 kb *Irf8* enhancer to the +32 kb *Irf8* enhancer by Reducing E protein activity,

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 >5 independent experiments, n > 5).**c,** Data shown in (**b**) shown as integrated MFI (data pooled from >5

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 integrated MFI (data pooled from >5 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 integrated MFI (data pooled from >5 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 co-transduced with either empty RV (gray) or ID2 RV (purple), shown as integrated MFI (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. Shown is the *Irf8* locus, with the *Irf8* +41 kb enhancer region (black box) and the +32 kb enhancer region (dotted box). (representative of three independent experiments and the Immunological Genome Project Open Chromatin Regions, n= 1 biological replicate per population). Data are presented as mean and one-way or two-way ANOVA was used to compare groups. *p < 0.05, **p < 0.01, ****p < 0.0001.